HOPF ALGEBRAS IN COMBINATORICS (VERSION CONTAINING SOLUTIONS)

DARIJ GRINBERG AND VICTOR REINER

Contents

Date: July 27, 2020 (with minor corrections November 17, 2020).

Key words and phrases. Hopf algebra, combinatorics, symmetric functions, quasisymmetric functions.

Introduction

The concept of a Hopf algebra crystallized out of algebraic topology and the study of algebraic groups in the 1940s and 1950s (see [8] and [35] for its history). Being a fairly elementary algebraic notion itself, it subsequently found applications in other mathematical disciplines, and is now particularly commonplace in representation theory[1].

These notes concern themselves (after a brief introduction into the algebraic foundations of Hopf algebra theory in Chapter 1) with the Hopf algebras that appear in combinatorics. These Hopf algebras tend to have bases naturally parametrized by combinatorial objects (partitions, compositions, permutations, tableaux, graphs, trees, posets, polytopes, etc.), and their Hopf-algebraic operations often encode basic operations on these objects[2]. Combinatorial results can then be seen as particular cases of general algebraic properties of Hopf algebras (e.g., the multiplicativity of the Möbius function can be recovered from the fact that the antipode of a Hopf algebra is an algebra anti-endomorphism), and many interesting invariants of combinatorial objects turn out to be evaluations of Hopf morphisms. In some cases (particularly that of symmetric functions), the rigidity in the structure of a Hopf algebra can lead to enlightening proofs.

One of the most elementary interesting examples of a combinatorial Hopf algebra is that of the symmetric functions. We will devote all of Chapter 2 to studying it, deviating from the usual treatments (such as in Stanley [206, Ch. 7], Sagan [186] and Macdonald [142]) by introducing the Hopf-algebraic structure early on and using it to obtain combinatorial results. Chapter 3 will underpin the importance of this algebra by proving Zelevinsky's main theorem of PSH theory, which (roughly) claims that a Hopf algebra over $\mathbb{Z}$ satisfying a certain set of axioms must be a tensor product of copies of the Hopf algebra of symmetric functions. These axioms are fairly restrictive, so this result is far from curtailing the diversity of combinatorial Hopf algebras; but they are natural enough that, as we will see in Chapter 4, they are satisfied for a Hopf algebra of representations of symmetric groups. As a consequence, this Hopf algebra will be revealed isomorphic to the symmetric functions – this is the famous Frobenius correspondence between symmetric functions and characters of symmetric groups, usually obtained through other ways ([73, §7.3], [186, §4.7]). We will further elaborate on the representation theories of wreath products and general linear groups over finite fields; while Zelevinsky's PSH theory does not fully explain the latter, it illuminates it significantly.

In the next chapters, we will study further examples of combinatorial Hopf algebras: the quasisymmetric functions and the noncommutative symmetric functions in Chapter 5, various other algebras (of graphs, posets, matroids, etc.) in Chapter 7, and the Malvenuto-Reutenauer Hopf algebra of permutations in Chapter 8.

The main prerequisite for reading these notes is a good understanding of graduate algebra[3], in particular multilinear algebra (tensor products, symmetric powers and exterior powers)[4] and basic categorical language[5]. In Chapter 4, familiarity with representation theory of finite groups (over $\mathbb{C}$) is assumed, along with the theory of finite fields and (at some places) the rational canonical form of a matrix. Only basic knowledge of combinatorics is required (except for a few spots in Chapter 7), and familiarity with geometry and topology is needed only to understand some tangential remarks. The concepts of Hopf algebras and coalgebras and the basics of symmetric function theory will be introduced as needed. We will work over a commutative base ring most of the time, but no commutative algebra (besides, occasionally, properties of modules over a PID) will be used.

These notes began as an accompanying text for Fall 2012 Math 8680 Topics in Combinatorics, a graduate class taught by the second author at the University of Minnesota. The first author has since added many exercises (and solutions), as well as Chapter 6 on Lyndon words and the polynomiality of QSym. The notes might still grow, and any comments, corrections and complaints are welcome!

[1] where it provides explanations for similarities between group representations and Lie algebra representations

[2] such as concatenating two compositions, or taking the disjoint union of two graphs – but, more often, operations which return a multiset of results, such as cutting a composition into two pieces at all possible places, or partitioning a poset into two subposets in every way that satisfies a certain axiom

[3] William Schmitt's expositions [193] are tailored to a reader interested in combinatorial Hopf algebras; his notes on modules and algebras cover a significant part of what we need from abstract algebra, whereas those on categories cover all category theory we will use and much more.

[4] Keith Conrad's expository notes [40] are useful, even if not comprehensive, sources for the latter.

[5] We also will use a few nonstandard notions from linear algebra that are explained in the Appendix (Chapter 11).

The course was an attempt to focus on examples that we find interesting, but which are hard to find fully explained currently in books or in one paper. Much of the subject of combinatorial Hopf algebras is fairly recent (1990s onwards) and still spread over research papers, although sets of lecture notes do exist, such as Foissy's [70]. A reference which we discovered late, having a great deal of overlap with these notes is Hazewinkel, Gubareni, and Kirichenko [93]. References for the purely algebraic theory of Hopf algebras are much more frequent (see the beginning of Chapter 1 for a list). Another recent text that has a significant amount of material in common with ours (but focuses on representation theory and probability applications) is Méliot's [153].

Be warned that our notes are highly idiosyncratic in choice of topics, and they steal heavily from the sources in the bibliography.

Warnings: Unless otherwise specified ...

- $\mathbf{k}$ here usually denotes a commutative ring[6].
- all maps between $\mathbf{k}$-modules are $\mathbf{k}$-linear.
- every ring or $\mathbf{k}$-algebra is associative and has a 1, and every ring morphism or $\mathbf{k}$-algebra morphism preserves the 1's.
- all $\mathbf{k}$-algebras A have the property that $(\lambda 1_A)\, a = a\,(\lambda 1_A) = \lambda a$ for all $\lambda \in \mathbf{k}$ and $a \in A$.
- all tensor products are over $\mathbf{k}$ (unless a subscript specifies a different base ring).
- 1 will denote the multiplicative identity in some ring like $\mathbf{k}$ or in some $\mathbf{k}$-algebra (sometimes also the identity of a group written multiplicatively).
- for any set S, we denote by id_S (or by id) the identity map on S.
- The symbols $\subset$ (for "subset") and $<$ (for "subgroup") don't imply properness (so $\mathbb{Z} \subset \mathbb{Z}$ and $\mathbb{Z} < \mathbb{Z}$).
- the n-th symmetric group (i.e., the group of all permutations of $\{1, 2, \ldots, n\}$) is denoted $\mathfrak{S}_n$.
- A permutation $\sigma \in \mathfrak{S}_n$ will often be identified with the n-tuple $(\sigma(1), \sigma(2), \ldots, \sigma(n))$, which will occasionally be written without commas and parentheses (i.e., as follows: $\sigma(1)\,\sigma(2)\cdots\sigma(n)$). This is called the *one-line notation* for permutations.
- The product of permutations $a \in \mathfrak{S}_n$ and $b \in \mathfrak{S}_n$ is defined by $(ab)(i) = a(b(i))$ for all i.
- *Words* over (or in) an *alphabet* I simply mean finite tuples of elements of a set I. It is customary to write such a word $(a_1, a_2, \ldots, a_k)$ as $a_1 a_2 \ldots a_k$ when this is not likely to be confused for multiplication.
- $\mathbb{N} := \{0, 1, 2, \ldots\}$.
- if i and j are any two objects, then $\delta_{i,j}$ denotes the *Kronecker delta* of i and j; this is the integer 1 if $i = j$ and 0 otherwise.
- a *family* of objects indexed by a set I means a choice of an object f_i for each element $i \in I$; this family will be denoted either by $(f_i)_{i \in I}$ or by $\{f_i\}_{i \in I}$ (and sometimes the "$i \in I$" will be omitted when the context makes it obvious – so we just write $\{f_i\}$).
- several objects $s_1, s_2, \ldots, s_k$ are said to be *distinct* if every $i \neq j$ satisfy $s_i \neq s_j$.
- similarly, several sets $S_1, S_2, \ldots, S_k$ are said to be *disjoint* if every $i \neq j$ satisfy $S_i \cap S_j = \varnothing$.
- the symbol $\sqcup$ (and the corresponding quantifier $\bigsqcup$) denotes a disjoint union of sets or posets. For example, if $S_1, S_2, \ldots, S_k$ are k sets, then $\bigsqcup_{i=1}^k S_i$ is their disjoint union. This disjoint union can mean either of the following two things:
  - It can mean the union $\bigcup_{i=1}^k S_i$ in the case when the sets $S_1, S_2, \ldots, S_k$ are disjoint. This is called an "internal disjoint union", and is simply a way to refer to the union of sets while simultaneously claiming that these sets are disjoint. Thus, of course, it is only well-defined if the sets are disjoint.
  - It can also mean the union $\bigcup_{i=1}^k \{i\} \times S_i$. This is called an "external disjoint union", and is well-defined whether or not the sets $S_1, S_2, \ldots, S_k$ are disjoint; it is a way to assemble the sets $S_1, S_2, \ldots, S_k$ into a larger set which contains a copy of each of their elements that "remembers" which set this element comes from.

  The two meanings are different, but in the case when $S_1, S_2, \ldots, S_k$ are disjoint, they are isomorphic. We hope the reader will not have a hard time telling which of them we are trying to evoke.

[6] As explained below, "ring" means "associative ring with 1". The most important cases are when $\mathbf{k}$ is a field or when $\mathbf{k} = \mathbb{Z}$.

Similarly, the notion of a direct sum of **k**-modules has two meanings ("internal direct sum" and "external direct sum").

- A sequence $(w_1, w_2, \ldots, w_k)$ of numbers (or, more generally, of elements of a poset) is said to be *strictly increasing* (or, for short, *increasing*) if it satisfies $w_1 < w_2 < \cdots < w_k$. A sequence $(w_1, w_2, \ldots, w_k)$ of numbers (or, more generally, of elements of a poset) is said to be *weakly increasing* (or *nondecreasing*) if it satisfies $w_1 \leq w_2 \leq \cdots \leq w_k$. Reversing the inequalities, we obtain the definitions of a *strictly decreasing* (a.k.a. *decreasing*) and of a *weakly decreasing* (a.k.a. *nonincreasing*) sequence. All these definitions extend in an obvious way to infinite sequences. Note that "nondecreasing" is not the same as "not decreasing"; for example, any sequence having at most one entry is both decreasing and nondecreasing, whereas the sequence $(1, 3, 1)$ is neither.

Hopefully context will resolve some of the ambiguities.

1. What is a Hopf algebra?

The standard references for Hopf algebras are Abe [1] and Sweedler [213], and some other good ones are [33, 36, 47, 93, 107, 118, 157, 176, 196, 225]. See also Foissy [70] and Manchon [149] for introductions to Hopf algebras tailored to combinatorial applications. Most texts only study Hopf algebras over fields (with exceptions such as [36, 33, 225]). We will work over arbitrary commutative rings[7], which requires some more care at certain points (but we will not go deep enough into the algebraic theory to witness the situation over commutative rings diverge seriously from that over fields).

Let's build up the definition of Hopf algebra structure bit-by-bit, starting with the more familiar definition of algebras.

1.1. Algebras.

Recall that an *associative* $\mathbf{k}$*-algebra* is defined to be a $\mathbf{k}$-module A equipped with an associative $\mathbf{k}$-bilinear map $\operatorname{mult} : A \times A \to A$ (the *multiplication map* of A) and an element $1 \in A$ (the *(multiplicative) unity* or *identity* of A) that is neutral for this map mult (that is, it satisfies $\operatorname{mult}(a, 1) = \operatorname{mult}(1, a) = a$ for all $a \in A$). If we recall that

- $\mathbf{k}$-bilinear maps $A \times A \to A$ are in 1-to-1 correspondence with $\mathbf{k}$-linear maps $A \otimes A \to A$ (by the universal property of the tensor product), and
- elements of A are in 1-to-1 correspondence with $\mathbf{k}$-linear maps $\mathbf{k} \to A$,

then we can restate this classical definition of associative $\mathbf{k}$-algebras as follows in terms of $\mathbf{k}$-linear maps[8]:

Definition 1.1.1. An *associative* $\mathbf{k}$*-algebra* is a $\mathbf{k}$-module A equipped with a $\mathbf{k}$-linear *associative operation* $A \otimes A \overset{m}{\to} A$, and a $\mathbf{k}$-linear *unit* $\mathbf{k} \overset{u}{\to} A$, for which the following two diagrams are commutative:

$$\begin{tikzcd}
& A \otimes A \otimes A \arrow[dl, "m\otimes \mathrm{id}"'] \arrow[dr, "\mathrm{id} \otimes m"] & \\
A \otimes A \arrow[dr, "m"] & & A \otimes A \arrow[dl, "m"'] \\
& A &
\end{tikzcd} \tag{1.1.1}$$

$$\begin{tikzcd}
A \otimes \mathbf{k} \arrow[d, "\mathrm{id} \otimes u"'] & A \arrow[l] \arrow[r] \arrow[d, "\mathrm{id}"'] & \mathbf{k} \otimes A \arrow[d, "u\otimes \mathrm{id}"] \\
A \otimes A \arrow[r, "m"] & A & A \otimes A \arrow[l, "m"']
\end{tikzcd} \tag{1.1.2}$$

where the maps $A \to A \otimes \mathbf{k}$ and $A \to \mathbf{k} \otimes A$ are the isomorphisms sending $a \mapsto a \otimes 1$ and $a \mapsto 1 \otimes a$.

We abbreviate "associative $\mathbf{k}$-algebra" as "$\mathbf{k}$-algebra" (associativity is assumed unless otherwise specified) or as "algebra" (when $\mathbf{k}$ is clear from the context). We sometimes refer to m as the "multiplication map" of A as well.

As we said, the multiplication map $m : A \otimes A \to A$ sends each $a \otimes b$ to the product ab, and the unit map $u : \mathbf{k} \to A$ sends the identity $1_{\mathbf{k}}$ of $\mathbf{k}$ to the identity 1_A of A.

Well-known examples of $\mathbf{k}$-algebras are *tensor* and *symmetric algebras*, which we can think of as algebras of *words* and *multisets*, respectively.

Example 1.1.2. If V is a $\mathbf{k}$-module and $n \in \mathbb{N}$, then the *n-fold tensor power* $V^{\otimes n}$ of V is the $\mathbf{k}$-module $\underbrace{V \otimes V \otimes \cdots \otimes V}_{n \text{ times}}$. (For $n = 0$, this is the $\mathbf{k}$-module $\mathbf{k}$, spanned by the "empty tensor" $1_{\mathbf{k}}$.)

The *tensor algebra* $T(V) = \bigoplus_{n \geq 0} V^{\otimes n}$ on a $\mathbf{k}$-module V is an associative $\mathbf{k}$-algebra spanned (as $\mathbf{k}$-module) by decomposable tensors $v_1 v_2 \cdots v_k := v_1 \otimes v_2 \otimes \cdots \otimes v_k$ with $k \in \mathbb{N}$ and $v_1, v_2, \ldots, v_k \in V$. Its multiplication

[7] and we will profit from this generality in Chapters 3 and 4, where we will be applying the theory of Hopf algebras to $\mathbf{k} = \mathbb{Z}$ in a way that would not be possible over $\mathbf{k} = \mathbb{Q}$

[8] Explicitly speaking, we are replacing the $\mathbf{k}$-bilinear multiplication map $\operatorname{mult} : A \times A \to A$ by the $\mathbf{k}$-linear map $m : A \otimes A \to A,\ a \otimes b \mapsto \operatorname{mult}(a, b)$, and we are replacing the element $1 \in A$ by the $\mathbf{k}$-linear map $u : \mathbf{k} \to A,\ 1_{\mathbf{k}} \mapsto 1$.

is defined $\mathbf{k}$-linearly by

$$m\left(v_1 v_2 \cdots v_k \otimes w_1 w_2 \cdots w_\ell\right) := v_1 v_2 \cdots v_k w_1 w_2 \cdots w_\ell$$

[9] for all $k, \ell \in \mathbb{N}$ and $v_1, v_2, \ldots, v_k, w_1, w_2, \ldots, w_\ell$ in V. The unit map $u : \mathbf{k} \to T\left(V\right)$ sends $1_{\mathbf{k}}$ to the empty tensor $1_{T(V)} = 1_{\mathbf{k}} \in \mathbf{k} = V^{\otimes 0}$.

If V is a free $\mathbf{k}$-module, say with $\mathbf{k}$-basis $\{x_i\}_{i \in I}$, then $T(V)$ has a $\mathbf{k}$-basis of decomposable tensors $x_{i_1} \cdots x_{i_k} := x_{i_1} \otimes \cdots \otimes x_{i_k}$ indexed by *words* $(i_1, \ldots, i_k)$ in the alphabet I, and the multiplication on this basis is given by concatenation of words:

$$m(x_{i_1} \cdots x_{i_k} \otimes x_{j_1} \cdots x_{j_\ell}) = x_{i_1} \cdots x_{i_k} x_{j_1} \cdots x_{j_\ell}.$$

Recall that a *two-sided ideal* of a $\mathbf{k}$-algebra A is defined to be a $\mathbf{k}$-submodule J of A such that all $j \in J$ and $a \in A$ satisfy $ja \in J$ and $aj \in J$. Using tensors, we can restate this as follows: A *two-sided ideal* of a $\mathbf{k}$-algebra A means a $\mathbf{k}$-submodule J of A satisfying $m(J \otimes A) \subset J$ and $m(A \otimes J) \subset J$. Often, the word "two-sided" is omitted and one just speaks of an ideal.

It is well-known that if J is a two-sided ideal of a $\mathbf{k}$-algebra A, then one can form a *quotient algebra* A/J.

Example 1.1.3. Let V be a $\mathbf{k}$-module. The *symmetric algebra* $\operatorname{Sym}(V) = \bigoplus_{n \geq 0} \operatorname{Sym}^n(V)$ is the quotient of $T(V)$ by the two-sided ideal generated by all elements $xy - yx$ with x, y in V. When V is a free $\mathbf{k}$-module with basis $\{x_i\}_{i \in I}$, this symmetric algebra $S\left(V\right)$ can be identified with a (commutative) polynomial algebra $\mathbf{k}[x_i]_{i \in I}$, having a $\mathbf{k}$-basis of (commutative) monomials $x_{i_1} \cdots x_{i_k}$ as $\{i_1, \ldots, i_k\}_{\text{multiset}}$ runs through all finite multisubsets[10] of I, and with multiplication defined $\mathbf{k}$-linearly via multiset union[11].

Note that the $\mathbf{k}$-module $\mathbf{k}$ itself canonically becomes a $\mathbf{k}$-algebra. Its associative operation $m : \mathbf{k} \otimes \mathbf{k} \to \mathbf{k}$ is the canonical isomorphism $\mathbf{k} \otimes \mathbf{k} \to \mathbf{k}$, and its unit $u : \mathbf{k} \to \mathbf{k}$ is the identity map.

Topology and group theory give more examples.

Example 1.1.4. The *cohomology algebra* $H^*(X; \mathbf{k}) = \bigoplus_{i \geq 0} H^i(X; \mathbf{k})$ with coefficients in $\mathbf{k}$ for a topological space X has an associative *cup product*. Its unit $\mathbf{k} = H^*(\mathbf{pt}; \mathbf{k}) \overset{u}{\to} H^*(X; \mathbf{k})$ is induced from the unique (continuous) map $X \to \mathbf{pt}$, where $\mathbf{pt}$ is a one-point space.

Example 1.1.5. For a group G, the *group algebra* $\mathbf{k}G$ has $\mathbf{k}$-basis $\{t_g\}_{g \in G}$ and multiplication defined $\mathbf{k}$-linearly by $t_g t_h = t_{gh}$, and unit defined by $u(1) = t_e$, where e is the identity element of G.

1.2. **Coalgebras.** In Definition 1.1.1, we have defined the notion of an algebra entirely in terms of linear maps; thus, by reversing all arrows, we can define a dual notion, which is called a *coalgebra*. If we are to think of the multiplication $A \otimes A \to A$ in an algebra as *putting together* two basis elements of A to get a sum of basis elements of A, then coalgebra structure should be thought of as *taking basis elements apart*.

Definition 1.2.1. A *co-associative* $\mathbf{k}$*-coalgebra* is a $\mathbf{k}$-module C equipped with a *comultiplication*, that is, a $\mathbf{k}$-linear map $C \overset{\Delta}{\to} C \otimes C$, and a $\mathbf{k}$-linear *counit* $C \overset{\epsilon}{\to} \mathbf{k}$ for which the following diagrams (which are exactly the diagrams in (1.1.1) and (1.1.2) but with *all arrows reversed*) are commutative:

[9] Some remarks about our notation (which we are using here and throughout these notes) are in order.

Since we are working with tensor products of $\mathbf{k}$-modules like $T\left(V\right)$ – which themselves are made of tensors – here, we must specify what the $\otimes$ sign means in expressions like $a \otimes b$ where a and b are elements of $T\left(V\right)$. Our convention is the following: When a and b are elements of a tensor algebra $T\left(V\right)$, we always understand $a \otimes b$ to mean the pure tensor $a \otimes b \in T\left(V\right) \otimes T\left(V\right)$ rather than the product of a and b inside the tensor algebra $T\left(V\right)$. The latter product will plainly be written ab.

The operator precedence between $\otimes$ and multiplication in $T\left(V\right)$ is such that multiplication in $T\left(V\right)$ binds more tightly than the $\otimes$ sign; e.g., the term $ab \otimes cd$ means $(ab) \otimes (cd)$. The same convention applies to any algebra instead of $T\left(V\right)$.

[10] By a *multisubset* of a set S, we mean a multiset each of whose elements belongs to S (but can appear arbitrarily often).

[11] The *multiset union* of two finite multisets A and B is defined to be the multiset C with the property that every x satisfies

$$(\text{multiplicity of } x \text{ in } C) = (\text{multiplicity of } x \text{ in } A) + (\text{multiplicity of } x \text{ in } B).$$

Equivalently, the multiset union of $\{a_1, a_2, \ldots, a_k\}_{\text{multiset}}$ and $\{b_1, b_2, \ldots, b_\ell\}_{\text{multiset}}$ is $\{a_1, a_2, \ldots, a_k, b_1, b_2, \ldots, b_\ell\}_{\text{multiset}}$. The multiset union is also known as the *disjoint union* of multisets.

$$\begin{tikzcd}
 & C \otimes C \otimes C & \\
C \otimes C \arrow[ur, "\Delta \otimes \mathrm{id}"] & & C \otimes C \arrow[ul, "\mathrm{id} \otimes \Delta"'] \\
 & C \arrow[ul, "\Delta"] \arrow[ur, "\Delta"'] &
\end{tikzcd} \tag{1.2.1}$$

$$\begin{tikzcd}
C \otimes \mathbf{k} \arrow[r] & C & \mathbf{k} \otimes C \arrow[l] \\
C \otimes C \arrow[u, "\mathrm{id} \otimes \epsilon"] & C \arrow[l, "\Delta"] \arrow[r, "\Delta"'] \arrow[u, "\mathrm{id}"] & C \otimes C \arrow[u, "\epsilon \otimes \mathrm{id}"']
\end{tikzcd} \tag{1.2.2}$$

Here the maps $C \otimes \mathbf{k} \to C$ and $\mathbf{k} \otimes C \to C$ are the isomorphisms sending $c \otimes 1 \mapsto c$ and $1 \otimes c \mapsto c$.

We abbreviate "co-associative $\mathbf{k}$-coalgebra" as "$\mathbf{k}$-coalgebra" (co-associativity, i.e., the commutativity of the diagram (1.2.1), is assumed unless otherwise specified) or as "coalgebra" (when $\mathbf{k}$ is clear from the context).

Sometimes, the word "coproduct" is used as a synonym for "comultiplication"[12].

One often uses the *Sweedler notation*

$$\Delta(c) = \sum_{(c)} c_1 \otimes c_2 = \sum c_1 \otimes c_2 \tag{1.2.3}$$

to abbreviate formulas involving Δ. This means that an expression of the form $\sum_{(c)} f(c_1, c_2)$ (where $f : C \times C \to M$ is some $\mathbf{k}$-bilinear map from $C \times C$ to some $\mathbf{k}$-module M) has to be understood to mean $\sum_{k=1}^{m} f(d_k, e_k)$, where $k \in \mathbb{N}$ and $d_1, d_2, \ldots, d_k \in C$ and $e_1, e_2, \ldots, e_k \in C$ are chosen such that $\Delta(c) = \sum_{k=1}^{m} d_k \otimes e_k$. (There are many ways to choose such k, d_i and e_i, but they all produce the same result $\sum_{k=1}^{m} f(d_k, e_k)$. Indeed, the result they produce is $F(\Delta(c))$, where $F : C \otimes C \to M$ is the $\mathbf{k}$-linear map induced by the bilinear map f.) For example, commutativity of the left square in (1.2.2) asserts that $\sum_{(c)} c_1 \epsilon(c_2) = c$ for each $c \in C$. Likewise, commutativity of the right square in (1.2.2) asserts that $\sum_{(c)} \epsilon(c_1) c_2 = c$ for each $c \in C$. The commutativity of (1.2.1) can be written as $\sum_{(c)} \Delta(c_1) \otimes c_2 = \sum_{(c)} c_1 \otimes \Delta(c_2)$, or (using nested Sweedler notation to unravel the two remaining Δ's) as

$$\sum_{(c)} \sum_{(c_1)} (c_1)_1 \otimes (c_1)_2 \otimes c_2 = \sum_{(c)} \sum_{(c_2)} c_1 \otimes (c_2)_1 \otimes (c_2)_2.$$

The $\mathbf{k}$-module $\mathbf{k}$ itself canonically becomes a $\mathbf{k}$-coalgebra, with its comultiplication $\Delta : \mathbf{k} \to \mathbf{k} \otimes \mathbf{k}$ being the canonical isomorphism $\mathbf{k} \to \mathbf{k} \otimes \mathbf{k}$, and its counit $\epsilon : \mathbf{k} \to \mathbf{k}$ being the identity map.

Example 1.2.2. Let $\mathbf{k}$ be a field. The *homology* $H_*(X; \mathbf{k}) = \bigoplus_{i \geq 0} H_i(X; \mathbf{k})$ for a topological space X is naturally a coalgebra: the (continuous) *diagonal embedding* $X \to X \times X$ sending $x \mapsto (x, x)$ induces a coassociative map

$$H_*(X; \mathbf{k}) \to H_*(X \times X; \mathbf{k}) \cong H_*(X; \mathbf{k}) \otimes H_*(X; \mathbf{k})$$

in which the last isomorphism comes from the *Künneth theorem* with field coefficients $\mathbf{k}$. As before, the unique (continuous) map $X \to \mathbf{pt}$ induces the counit $H_*(X; \mathbf{k}) \xrightarrow{\epsilon} H_*(\mathbf{pt}; \mathbf{k}) \cong \mathbf{k}$.

Exercise 1.2.3. Let C be a $\mathbf{k}$-module, and let $\Delta : C \to C \otimes C$ be a $\mathbf{k}$-linear map. Prove that there exists *at most one* $\mathbf{k}$-linear map $\epsilon : C \to \mathbf{k}$ such that the diagram (1.2.2) commutes.

For us, the notion of a coalgebra serves mostly as a stepping stone towards that of a Hopf algebra, which will be the focus of these notes. However, coalgebras have interesting properties of their own (see, e.g., [150]).

[12] although the word "coproduct" already has a different meaning in algebra

1.3. **Morphisms, tensor products, and bialgebras.** Just as we rewrote the definition of an algebra in terms of linear maps (in Definition 1.1.1), we can likewise rephrase the standard definition of a morphism of algebras:

Definition 1.3.1. A *morphism of algebras* is a $\mathbf{k}$-linear map $A \overset{\varphi}{\to} B$ between two $\mathbf{k}$-algebras A and B that makes the following two diagrams commute:

(1.3.1)
$$\begin{tikzcd} A \arrow[r, "\varphi"] & B \\ A \otimes A \arrow[u, "m_A"] \arrow[r, "\varphi\otimes\varphi"] & B \otimes B \arrow[u, "m_B"'] \end{tikzcd} \qquad \begin{tikzcd} A \arrow[rr, "\varphi"] & & B \\ & \mathbf{k} \arrow[ul, "u_A"] \arrow[ur, "u_B"'] & \end{tikzcd}$$

Here the subscripts on m_A, m_B, u_A, u_B indicate for which algebra they are part of the structure (e.g., the map u_A is the map u of the algebra A); we will occasionally use such conventions from now on.

Similarly, a *morphism of coalgebras* is a $\mathbf{k}$-linear map $C \overset{\varphi}{\to} D$ between two $\mathbf{k}$-coalgebras C and D that makes the reverse diagrams commute:

(1.3.2)
$$\begin{tikzcd} C \arrow[r, "\varphi"] \arrow[d, "\Delta_C"'] & D \arrow[d, "\Delta_D"] \\ C \otimes C \arrow[r, "\varphi\otimes\varphi"] & D \otimes D \end{tikzcd} \qquad \begin{tikzcd} C \arrow[rr, "\varphi"] \arrow[dr, "\epsilon_C"'] & & D \arrow[dl, "\epsilon_D"] \\ & \mathbf{k} & \end{tikzcd}$$

As usual, we shall use the word "*homomorphism*" as a synonym for "morphism", and we will say "$\mathbf{k}$-coalgebra homomorphism" for "homomorphism of coalgebras" (and similarly for algebras and other structures).

As usual, the word "*isomorphism*" (of algebras, of coalgebras, or of other structures that we will define further below) means "invertible morphism whose inverse is a morphism as well". Two algebras (or coalgebras, or other structures) are said to be *isomorphic* if there exists an isomorphism between them.

Example 1.3.2. Let $\mathbf{k}$ be a field. Continuous maps $X \overset{f}{\to} Y$ of topological spaces induce algebra morphisms $H^*(Y; \mathbf{k}) \to H^*(X; \mathbf{k})$, and coalgebra morphisms $H_*(X; \mathbf{k}) \to H_*(Y; \mathbf{k})$.

Coalgebra morphisms behave similarly to algebra morphisms in many regards: For example, the inverse of an invertible coalgebra morphism is again a coalgebra morphism[13]. Thus, the invertible coalgebra morphisms are precisely the coalgebra isomorphisms.

Definition 1.3.3. Given two $\mathbf{k}$-algebras A, B, their tensor product $A \otimes B$ also becomes a $\mathbf{k}$-algebra defining the multiplication bilinearly via

$$m((a \otimes b) \otimes (a' \otimes b')) := aa' \otimes bb',$$

or, in other words, $m_{A\otimes B}$ is the composite map

$$A \otimes B \otimes A \otimes B \xrightarrow{\mathrm{id} \otimes T \otimes \mathrm{id}} A \otimes A \otimes B \otimes B \xrightarrow{m_A \otimes m_B} A \otimes B$$

where T is the *twist map* $B \otimes A \to A \otimes B$ that sends $b \otimes a \mapsto a \otimes b$. (See Exercise 1.3.4(a) below for a proof that this $\mathbf{k}$-algebra $A \otimes B$ is well-defined.)

Here we are omitting the topologist's sign in the twist map which should be present for graded algebras and coalgebras that come from cohomology and homology: For homogeneous elements a and b of two graded modules A and B, the topologist's twist map $T : B \otimes A \to A \otimes B$ sends

$$b \otimes a \longmapsto (-1)^{\deg(b)\deg(a)} a \otimes b \tag{1.3.3}$$

instead of $b \otimes a \mapsto a \otimes b$. This means that, if one is using the topologists' convention, most of our examples which we later call *graded* should actually be considered to live in only *even* degrees (which can be achieved, e.g., by artificially doubling all degrees in their grading). We will, however, keep to our own definitions (so that our twist map T will always send $b \otimes a \mapsto a \otimes b$) unless otherwise noted. Only in parts of Exercise 1.6.5 will we use the topologist's sign. Readers interested in the wide world of algebras defined using the topologist's

[13] The easy proof of this fact is left to the reader.

sign convention (which is also known as the *Koszul sign rule*) can consult [65, Appendix A2]; see also [87] for applications to algebraic combinatorics[14].

The unit element of $A \otimes B$ is $1_A \otimes 1_B$, meaning that the unit map $\mathbf{k} \overset{u_{A\otimes B}}{\to} A \otimes B$ is the composite

$$\begin{tikzcd} \mathbf{k} \arrow[r] & \mathbf{k} \otimes \mathbf{k} \arrow[r, "u_A \otimes u_B"] & A \otimes B \end{tikzcd} .$$

Similarly, given two coalgebras C, D, one can make $C \otimes D$ a coalgebra in which the comultiplication and counit maps are the composites of

$$\begin{tikzcd} C \otimes D \arrow[r, "\Delta_C \otimes \Delta_D"] & C \otimes C \otimes D \otimes D \arrow[r, "\mathrm{id} \otimes T \otimes \mathrm{id}"] & C \otimes D \otimes C \otimes D \end{tikzcd}$$

and

$$\begin{tikzcd} C \otimes D \arrow[r, "\epsilon_C \otimes \epsilon_D"] & \mathbf{k} \otimes \mathbf{k} \arrow[r] & \mathbf{k} \end{tikzcd} .$$

(See Exercise 1.3.4(b) below for a proof that this $\mathbf{k}$-coalgebra $C \otimes D$ is well-defined.)

Exercise 1.3.4. (a) Let A and B be two $\mathbf{k}$-algebras. Show that the $\mathbf{k}$-algebra $A \otimes B$ introduced in Definition 1.3.3 is actually well-defined (i.e., its multiplication and unit satisfy the axioms of a $\mathbf{k}$-algebra).

(b) Let C and D be two $\mathbf{k}$-coalgebras. Show that the $\mathbf{k}$-coalgebra $C \otimes D$ introduced in Definition 1.3.3 is actually well-defined (i.e., its comultiplication and counit satisfy the axioms of a $\mathbf{k}$-coalgebra).

It is straightforward to show that the concept of tensor products of algebras and of coalgebras satisfy the properties one would expect:

- For any three $\mathbf{k}$-coalgebras C, D and E, the $\mathbf{k}$-linear map

$$(C \otimes D) \otimes E \to C \otimes (D \otimes E), \qquad (c \otimes d) \otimes e \mapsto c \otimes (d \otimes e)$$

  is a coalgebra isomorphism. This allows us to speak of the $\mathbf{k}$-coalgebra $C \otimes D \otimes E$ without worrying about the parenthesization.
- For any two $\mathbf{k}$-coalgebras C and D, the $\mathbf{k}$-linear map

$$T : C \otimes D \to D \otimes C, \qquad c \otimes d \mapsto d \otimes c$$

  is a coalgebra isomorphism.
- For any $\mathbf{k}$-coalgebra C, the $\mathbf{k}$-linear maps

$$\begin{aligned} C &\to \mathbf{k} \otimes C, \qquad c \mapsto 1 \otimes c \qquad \text{and} \\ C &\to C \otimes \mathbf{k}, \qquad c \mapsto c \otimes 1 \end{aligned}$$

  are coalgebra isomorphisms.
- Similar properties hold for algebras instead of coalgebras.

One of the first signs that these definitions interact nicely is the following straightforward proposition.

Proposition 1.3.5. *When A is both a $\mathbf{k}$-algebra and a $\mathbf{k}$-coalgebra, the following are equivalent:*

- *The maps Δ and ϵ are morphisms for the algebra structure (A, m, u).*
- *The maps m and u are morphisms for the coalgebra structure (A, Δ, ϵ).*

[14] To be precise, [87] works with the related concept of *superalgebras*, which are graded by elements of $\mathbb{Z}/2\mathbb{Z}$ rather than $\mathbb{N}$ but use the same sign convention as the topologists have for algebras.

- *These four diagrams commute:*

(1.3.4)

$$\begin{tikzcd}
& A\otimes A \arrow[ld, "\Delta\otimes\Delta"'] \arrow[rdd, "m"] & \\
A\otimes A\otimes A\otimes A \arrow[dd, "\mathrm{id}\otimes T\otimes\mathrm{id}"'] & & \\
& & A \arrow[ldd, "\Delta"] \\
A\otimes A\otimes A\otimes A \arrow[rd, "m\otimes m"'] & & \\
& A\otimes A &
\end{tikzcd}$$

$$\begin{tikzcd}
A\otimes A \arrow[r, "\epsilon\otimes\epsilon"] \arrow[d, "m"'] & \mathbf{k}\otimes\mathbf{k} \arrow[d, "m"] \\
A \arrow[r, "\epsilon"'] & \mathbf{k}
\end{tikzcd}
\qquad
\begin{tikzcd}
\mathbf{k} \arrow[r, "u"] \arrow[d, "\Delta"'] & A \arrow[d, "\Delta"] \\
\mathbf{k}\otimes\mathbf{k} \arrow[r, "u\otimes u"'] & A\otimes A
\end{tikzcd}$$

$$\begin{tikzcd}
\mathbf{k} \arrow[rr, "\mathrm{id}"] \arrow[rd, "u"'] & & \mathbf{k} \\
& A \arrow[ru, "\epsilon"'] &
\end{tikzcd}$$

Exercise 1.3.6. (a) If A, A', B and B' are four $\mathbf{k}$-algebras, and $f : A \to A'$ and $g : B \to B'$ are two $\mathbf{k}$-algebra homomorphisms, then show that $f \otimes g : A \otimes B \to A' \otimes B'$ is a $\mathbf{k}$-algebra homomorphism.

(b) If C, C', D and D' are four $\mathbf{k}$-coalgebras, and $f : C \to C'$ and $g : D \to D'$ are two $\mathbf{k}$-coalgebra homomorphisms, then show that $f \otimes g : C \otimes D \to C' \otimes D'$ is a $\mathbf{k}$-coalgebra homomorphism.

Definition 1.3.7. Call the $\mathbf{k}$-module A a $\mathbf{k}$-*bialgebra* if it is a $\mathbf{k}$-algebra and $\mathbf{k}$-coalgebra satisfying the three equivalent conditions in Proposition 1.3.5.

Example 1.3.8. For a group G, one can make the group algebra $\mathbf{k}G$ a coalgebra with counit $\mathbf{k}G \overset{\epsilon}{\to} \mathbf{k}$ mapping $t_g \mapsto 1$ for all g in G, and with comultiplication $\mathbf{k}G \overset{\Delta}{\to} \mathbf{k}G \otimes \mathbf{k}G$ given by $\Delta(t_g) := t_g \otimes t_g$. Checking the various diagrams in (1.3.4) commute is easy. For example, one can check the pentagonal diagram on each basis element $t_g \otimes t_h$:

$$\begin{tikzcd}
& t_g\otimes t_h \arrow[ld, "\Delta\otimes\Delta"'] \arrow[rdd, "m"] & \\
t_g\otimes t_g\otimes t_h\otimes t_h \arrow[dd, "\mathrm{id}\otimes T\otimes\mathrm{id}"'] & & \\
& & t_{gh} \arrow[ldd, "\Delta"] \\
t_g\otimes t_h\otimes t_g\otimes t_h \arrow[rd, "m\otimes m"'] & & \\
& t_{gh}\otimes t_{gh} &
\end{tikzcd}$$

Remark 1.3.9. In fact, one can think of adding a bialgebra structure to a $\mathbf{k}$-algebra A as a way of making A-modules M, N have an A-module structure on their tensor product $M \otimes N$: the algebra $A \otimes A$ already acts naturally on $M \otimes N$, so one can let a in A act via $\Delta(a)$ in $A \otimes A$. In the theory of group representations over $\mathbf{k}$, that is, $\mathbf{k}G$-modules M, this is how one defines the *diagonal action* of G on $M \otimes N$, namely t_g acts as $t_g \otimes t_g$.

Definition 1.3.10. An element x in a coalgebra for which $\Delta(x) = x \otimes x$ and $\epsilon(x) = 1$ is called *group-like*.

An element x in a bialgebra for which $\Delta(x) = 1 \otimes x + x \otimes 1$ is called *primitive*. We shall also sometimes abbreviate "primitive element" as "primitive".

Example 1.3.11. Let V be a $\mathbf{k}$-module. The *tensor algebra* $T(V) = \bigoplus_{n \geq 0} V^{\otimes n}$ is a coalgebra, with counit ϵ equal to the identity on $V^{\otimes 0} = \mathbf{k}$ and the zero map on $V^{\otimes n}$ for $n > 0$, and with comultiplication defined to make the elements x in $V^{\otimes 1} = V$ all primitive:

$$\Delta(x) := 1 \otimes x + x \otimes 1 \text{ for } x \in V^{\otimes 1}.$$

Since the elements of V generate $T(V)$ as a $\mathbf{k}$-algebra, and since $T(V) \otimes T(V)$ is also an associative $\mathbf{k}$-algebra, the universal property of $T(V)$ as the free associative $\mathbf{k}$-algebra on the generators V allows one to define $T(V) \overset{\Delta}{\to} T(V) \otimes T(V)$ arbitrarily on V, and extend it as an algebra morphism.

It may not be obvious that this Δ is coassociative, but one can prove this as follows. Note that

$$((\mathrm{id} \otimes \Delta) \circ \Delta)(x) = x \otimes 1 \otimes 1 + 1 \otimes x \otimes 1 + 1 \otimes 1 \otimes x = ((\Delta \otimes \mathrm{id}) \circ \Delta)(x)$$

for every x in V. Hence the two maps $(\mathrm{id} \otimes \Delta) \circ \Delta$ and $(\Delta \otimes \mathrm{id}) \circ \Delta$, considered as algebra morphisms $T(V) \to T(V) \otimes T(V) \otimes T(V)$, must coincide on every element of $T(V)$ since they coincide on V. We leave it as an exercise to check the map ϵ defined as above satisfies the counit axioms (1.2.2).

Here is a sample calculation in $T(V)$ when x, y, z are three elements of V:

$$\begin{aligned}
\Delta(xyz) &= \Delta(x)\Delta(y)\Delta(z) \\
&= (1 \otimes x + x \otimes 1)(1 \otimes y + y \otimes 1)(1 \otimes z + z \otimes 1) \\
&= (1 \otimes xy + x \otimes y + y \otimes x + xy \otimes 1)(1 \otimes z + z \otimes 1) \\
&= 1 \otimes xyz + x \otimes yz + y \otimes xz + z \otimes xy \\
&\qquad + xy \otimes z + xz \otimes y + yz \otimes x + xyz \otimes 1.
\end{aligned}$$

This illustrates the idea that comultiplication "takes basis elements apart" (and, in the case of $T(V)$, not just basis elements, but any decomposable tensors). Here for any $v_1, v_2, \ldots, v_n$ in V one has

$$\Delta(v_1 v_2 \cdots v_n) = \sum v_{j_1} \cdots v_{j_r} \otimes v_{k_1} \cdots v_{k_{n-r}}$$

where the sum is over ordered pairs $(j_1, j_2, \ldots, j_r), (k_1, k_2, \ldots, k_{n-r})$ of complementary subwords of the word $(1, 2, \ldots, n)$. [15] Equivalently (and in a more familiar language),

$$\Delta(v_1 v_2 \cdots v_n) = \sum_{I \subset \{1,2,\ldots,n\}} v_I \otimes v_{\{1,2,\ldots,n\} \setminus I}, \tag{1.3.5}$$

where v_J (for J a subset of $\{1, 2, \ldots, n\}$) denotes the product of all v_j with $j \in J$ in the order of increasing j.

We can rewrite the axioms of a $\mathbf{k}$-bialgebra A using Sweedler notation. Indeed, asking for $\Delta : A \to A \otimes A$ to be a $\mathbf{k}$-algebra morphism is equivalent to requiring that

$$\sum_{(ab)} (ab)_1 \otimes (ab)_2 = \sum_{(a)} \sum_{(b)} a_1 b_1 \otimes a_2 b_2 \qquad \text{for all } a, b \in A \tag{1.3.6}$$

and $\sum_{(1)} 1_1 \otimes 1_2 = 1_A \otimes 1_A$. (The other axioms have already been rewritten or don't need Sweedler notation.)

Recall one can quotient a $\mathbf{k}$-algebra A by a two-sided ideal J to obtain a quotient algebra A/J. An analogous construction can be done for coalgebras using the following concept, which is dual to that of a two-sided ideal:

[15] More formally speaking, the sum is over all permutations $(j_1, j_2, \ldots, j_r, k_1, k_2, \ldots, k_{n-r})$ of $(1, 2, \ldots, n)$ satisfying $j_1 < j_2 < \cdots < j_r$ and $k_1 < k_2 < \cdots < k_{n-r}$.

Definition 1.3.12. In a coalgebra C, a *two-sided coideal* is a $\mathbf{k}$-submodule $J \subset C$ for which

$$\Delta(J) \subset J \otimes C + C \otimes J,$$
$$\epsilon(J) = 0.$$

The quotient $\mathbf{k}$-module C/J then inherits a coalgebra structure[16]. Similarly, in a bialgebra A, a subset $J \subset A$ which is both a two-sided ideal and two-sided coideal gives rise to a quotient bialgebra A/J.

Exercise 1.3.13. Let A and C be two $\mathbf{k}$-coalgebras, and $f : A \to C$ a surjective coalgebra homomorphism.

(a) If f is surjective, then show that $\ker f$ is a two-sided coideal of A.
(b) If $\mathbf{k}$ is a field, then show that $\ker f$ is a two-sided coideal of A.

Example 1.3.14. Let V be a $\mathbf{k}$-module. The *symmetric algebra* $\operatorname{Sym}(V)$ was defined as the quotient of the tensor algebra $T(V)$ by the two-sided ideal J generated by all *commutators* $[x, y] = xy - yx$ for x, y in V (see Example 1.1.3). Note that x, y are primitive elements in $T(V)$, and the following very reusable calculation shows that *the commutator of two primitives is primitive*:

$$\begin{aligned}
\Delta[x,y] &= \Delta(xy - yx) = \Delta(x)\Delta(y) - \Delta(y)\Delta(x) \\
&\qquad \text{(since } \Delta \text{ is an algebra homomorphism)} \\
&= (1 \otimes x + x \otimes 1)(1 \otimes y + y \otimes 1) - (1 \otimes y + y \otimes 1)(1 \otimes x + x \otimes 1) \\
&= 1 \otimes xy - 1 \otimes yx + xy \otimes 1 - yx \otimes 1 \\
&\qquad + x \otimes y + y \otimes x - x \otimes y - y \otimes x \\
&= 1 \otimes (xy - yx) + (xy - yx) \otimes 1 \\
&= 1 \otimes [x,y] + [x,y] \otimes 1. \qquad (1.3.7)
\end{aligned}$$

In particular, the commutators $[x, y]$ have $\Delta[x, y]$ in $J \otimes T(V) + T(V) \otimes J$. They also satisfy $\epsilon([x, y]) = 0$. Since they are generators for J as a two-sided ideal, it is not hard to see this implies $\Delta(J) \subset J \otimes T(V) + T(V) \otimes J$, and $\epsilon(J) = 0$. Thus J is also a two-sided coideal, and $\operatorname{Sym}(V) = T(V)/J$ inherits a bialgebra structure.

In fact we will see in Section 3.1 that symmetric algebras are the universal example of bialgebras which are *graded, connected, commutative, cocommutative.* But first we should define some of these concepts.

Definition 1.3.15. (a) A *graded* $\mathbf{k}$*-module*[17] is a $\mathbf{k}$-module V equipped with a $\mathbf{k}$-module direct sum decomposition $V = \bigoplus_{n \geq 0} V_n$. In this case, the addend V_n (for any given $n \in \mathbb{N}$) is called the *n-th homogeneous component* (or the *n-th graded component*) of the graded $\mathbf{k}$-module V. Furthermore, elements x in V_n are said to be *homogeneous* of degree n; occasionally, the notation $\deg(x) = n$ is used to signify this[18]. The decomposition $\bigoplus_{n \geq 0} V_n$ of V (that is, the family of submodules $(V_n)_{n \in \mathbb{N}}$) is called the *grading* of V.

(b) The tensor product $V \otimes W$ of two graded $\mathbf{k}$-modules V and W is, by default, endowed with the graded module structure in which

$$(V \otimes W)_n := \bigoplus_{i+j=n} V_i \otimes W_j.$$

(c) A $\mathbf{k}$-linear map $V \xrightarrow{\varphi} W$ between two graded $\mathbf{k}$-modules is called *graded* if $\varphi(V_n) \subset W_n$ for all n. Graded $\mathbf{k}$-linear maps are also called *homomorphisms of graded* $\mathbf{k}$*-modules*. An *isomorphism of graded* $\mathbf{k}$*-modules* means an invertible graded $\mathbf{k}$-linear map whose inverse is also graded.[19]

[16] Indeed, $J \otimes C + C \otimes J$ is contained in the kernel of the canonical map $C \otimes C \to (C/J) \otimes (C/J)$; therefore, the condition $\Delta(J) \subset J \otimes C + C \otimes J$ shows that the map $C \xrightarrow{\Delta} C \otimes C \twoheadrightarrow (C/J) \otimes (C/J)$ factors through a map $\overline{\Delta} : C/J \to (C/J) \otimes (C/J)$. Likewise, $\epsilon(J) = 0$ shows that the map $\epsilon : C \to \mathbf{k}$ factors through a map $\overline{\epsilon} : C/J \to \mathbf{k}$. Equipping C/J with these maps $\overline{\Delta}$ and $\overline{\epsilon}$, we obtain a coalgebra (as the commutativity of the required diagrams follows from the corresponding property of C).

[17] also known as an "$\mathbb{N}$-*graded* $\mathbf{k}$*-module*"

[18] This notation should not be taken too literally, as it would absurdly imply that $\deg(0)$ "equals" every $n \in \mathbb{N}$ at the same time, since $0 \in V_n$ for all n.

[19] We shall see in Exercise 1.3.18 that the "whose inverse is also graded" requirement is actually superfluous (i.e., it is automatically satisfied for an invertible graded $\mathbf{k}$-linear map); we are imposing it only in order to stick to our tradition of defining "isomorphisms" as invertible morphisms whose inverses are morphisms as well.

(d) Say that a $\mathbf{k}$-algebra (or coalgebra, or bialgebra) is *graded* if it is a graded $\mathbf{k}$-module and all of the relevant structure maps (u, ϵ, m, Δ) are graded.
(e) Say that a graded $\mathbf{k}$-module V is *connected* if $V_0 \cong \mathbf{k}$.
(f) Let V be a graded $\mathbf{k}$-module. Then, a *graded* $\mathbf{k}$*-submodule of* V (sometimes also called a *homogeneous* $\mathbf{k}$*-submodule of* V) means a graded $\mathbf{k}$-module W such that $W \subset V$ as sets, and such that the inclusion map $W \hookrightarrow V$ is a graded $\mathbf{k}$-linear map.
Note that if W is a graded $\mathbf{k}$-submodule of V, then the grading of W is uniquely determined by the underlying set of W and the grading of V – namely, the n-th graded component W_n of W is $W_n = W \cap V_n$ for each $n \in \mathbb{N}$. Thus, we can specify a graded $\mathbf{k}$-submodule of V without explicitly specifying its grading. From this point of view, a graded $\mathbf{k}$-submodule of V can also be defined as a $\mathbf{k}$-submodule W of V satisfying $W = \sum_{n \in \mathbb{N}} (W \cap V_n)$. (This sum is automatically a direct sum, and thus defines a grading on W.)

Example 1.3.16. Let $\mathbf{k}$ be a field. A path-connected space X has its homology and cohomology

$$H_*(X; \mathbf{k}) = \bigoplus_{i \geq 0} H_i(X; \mathbf{k}),$$
$$H^*(X; \mathbf{k}) = \bigoplus_{i \geq 0} H^i(X; \mathbf{k})$$

carrying the structure of connected graded coalgebras and algebras, respectively. If in addition, X is a topological group, or even less strongly, a *homotopy-associative* H*-space* (e.g. the *loop space* ΩY on some other space Y), the continuous multiplication map $X \times X \to X$ induces an algebra structure on $H_*(X; \mathbf{k})$ and a coalgebra structure on $H^*(X; \mathbf{k})$, so that each become bialgebras in the topologist's sense (i.e., with the twist as in (1.3.3)), and these bialgebras are dual to each other in a sense soon to be discussed. This was Hopf's motivation: the (co-)homology of a compact Lie group carries bialgebra structure that explains why it takes a certain form; see Cartier [35, §2].

Example 1.3.17. Let V be a graded $\mathbf{k}$-module. Then, its tensor algebra $T(V)$ and its symmetric algebra $\operatorname{Sym}(V)$ are graded Hopf algebras. The grading is given as follows: If $v_1, v_2, \ldots, v_k$ are homogeneous elements of V having degrees $i_1, i_2, \ldots, i_k$, respectively, then the elements $v_1 v_2 \cdots v_k$ of $T(V)$ and $\operatorname{Sym}(V)$ are homogeneous of degree $i_1 + i_2 + \cdots + i_k$. That is, we have

$$\deg(v_1 v_2 \cdots v_k) = \deg(v_1) + \deg(v_2) + \cdots + \deg(v_k)$$

for any homogeneous elements $v_1, v_2, \ldots, v_k$ of V.

Assuming that $V_0 = 0$, the graded algebras $T(V)$ and $\operatorname{Sym}(V)$ are connected. This is a fairly common situation in combinatorics. For example, we will often turn a (non-graded) $\mathbf{k}$-module V into a graded $\mathbf{k}$-module by declaring that all elements of V are homogeneous of degree 1, but at other times, it will make sense to have V live in different (positive) degrees.

Exercise 1.3.18. Let V and W be two graded $\mathbf{k}$-modules. Prove that if $f : V \to W$ is an invertible graded $\mathbf{k}$-linear map, then its inverse $f^{-1} : W \to V$ is also graded.

Exercise 1.3.19. Let $A = \bigoplus_{n \geq 0} A_n$ be a graded $\mathbf{k}$-bialgebra. We denote by $\mathfrak{p}$ the set of all primitive elements of A.

(a) Show that $\mathfrak{p}$ is a graded $\mathbf{k}$-submodule of A (that is, we have $\mathfrak{p} = \bigoplus_{n \geq 0} (\mathfrak{p} \cap A_n)$).
(b) Show that $\mathfrak{p}$ is a two-sided coideal of A.

Exercise 1.3.20. Let A be a connected graded $\mathbf{k}$-bialgebra. Show the following:

(a) The $\mathbf{k}$-submodule $\mathbf{k} = \mathbf{k} \cdot 1_A$ of A lies in A_0.
(b) The map u is an isomorphism $\mathbf{k} \overset{u}{\to} A_0$.
(c) We have $A_0 = \mathbf{k} \cdot 1_A$.
(d) The two-sided ideal $\ker \epsilon$ is the $\mathbf{k}$-module of positive degree elements $I = \bigoplus_{n > 0} A_n$.
(e) The map ϵ restricted to A_0 is the inverse isomorphism $A_0 \overset{\epsilon}{\to} \mathbf{k}$ to u.
(f) For every $x \in A$, we have

$$\Delta(x) \in x \otimes 1 + A \otimes I.$$

(g) Every x in I satisfies

$$\Delta(x) = 1 \otimes x + x \otimes 1 + \Delta_+(x), \qquad \text{where } \Delta_+(x) \text{ lies in } I \otimes I.$$

(h) Every $n > 0$ and every $x \in A_n$ satisfy

$$\Delta(x) = 1 \otimes x + x \otimes 1 + \Delta_+(x), \qquad \text{where } \Delta_+(x) \text{ lies in } \sum_{k=1}^{n-1} A_k \otimes A_{n-k}.$$

(Use only the gradedness of the unit u and counit ϵ maps, along with commutativity of diagrams (1.2.2), and (1.3.4) and the connectedness of A.)

Having discussed graded $\mathbf{k}$-modules, let us also define the concept of a *graded basis*, which is the analogue of the notion of a basis in the graded context. Roughly speaking, a graded basis of a graded $\mathbf{k}$-module is a basis that comprises bases of all its homogeneous components. More formally:

Definition 1.3.21. Let $V = \bigoplus_{n \geq 0} V_n$ be a graded $\mathbf{k}$-module. A *graded basis* of the graded $\mathbf{k}$-module V means a basis $\{v_i\}_{i \in I}$ of the $\mathbf{k}$-module V whose indexing set I is partitioned into subsets $I_0, I_1, I_2, \ldots$ (which are allowed to be empty) with the property that, for every $n \in \mathbb{N}$, the subfamily $\{v_i\}_{i \in I_n}$ is a basis of the $\mathbf{k}$-module V_n.

Example 1.3.22. Consider the polynomial ring $\mathbf{k}[x]$ in one variable x over $\mathbf{k}$. This is a graded $\mathbf{k}$-module (graded by the degree of a polynomial; thus, each x^n is homogeneous of degree n). Then, the family $(x^n)_{n \in \mathbb{N}} = (x^0, x^1, x^2, \ldots)$ is a graded basis of $\mathbf{k}[x]$ (presuming that its indexing set $\mathbb{N}$ is partitioned into the one-element subsets $\{0\}, \{1\}, \{2\}, \ldots$). The family $((-x)^n)_{n \in \mathbb{N}} = (x^0, -x^1, x^2, -x^3, \ldots)$ is a graded basis of $\mathbf{k}[x]$ as well. But the family $((1+x)^n)_{n \in \mathbb{N}}$ is not, since it contains non-homogeneous elements.

We end this section by discussing morphisms between bialgebras. They are defined as one would expect:

Definition 1.3.23. A *morphism of bialgebras* (also known as a $\mathbf{k}$-*bialgebra homomorphism*) is a $\mathbf{k}$-linear map $A \overset{\varphi}{\to} B$ between two $\mathbf{k}$-bialgebras A and B that is simultaneously a $\mathbf{k}$-algebra homomorphism and a $\mathbf{k}$-coalgebra homomorphism.

For example, any $\mathbf{k}$-linear map $f : V \to W$ between two $\mathbf{k}$-modules V and W induces a $\mathbf{k}$-linear map $T(f) : T(V) \to T(W)$ between their tensor algebras (which sends each $v_1 v_2 \cdots v_k \in T(V)$ to $f(v_1) f(v_2) \cdots f(v_k) \in T(W)$) as well as a $\mathbf{k}$-linear map $\operatorname{Sym}(f) : \operatorname{Sym}(V) \to \operatorname{Sym}(W)$ between their symmetric algebras; both of these maps $T(f)$ and $\operatorname{Sym}(f)$ are morphisms of bialgebras.

Graded bialgebras come with a special family of endomorphisms, as the following exercise shows:

Exercise 1.3.24. Fix $q \in \mathbf{k}$. Let $A = \bigoplus_{n \in \mathbb{N}} A_n$ be a graded $\mathbf{k}$-bialgebra (where the A_n are the homogeneous components of A). Let $D_q : A \to A$ be the $\mathbf{k}$-module endomorphism of A defined by setting

$$D_q(a) = q^n a \qquad \text{for each } n \in \mathbb{N} \text{ and each } a \in A_n.$$

(It is easy to see that this is well-defined; equivalently, D_q could be defined as the direct sum $\bigoplus_{n \in \mathbb{N}} (q^n \cdot \mathrm{id}_{A_n}) : \bigoplus_{n \in \mathbb{N}} A_n \to \bigoplus_{n \in \mathbb{N}} A_n$ of the maps $q^n \cdot \mathrm{id}_{A_n} : A_n \to A_n$.)

Prove that D_q is a $\mathbf{k}$-bialgebra homomorphism.

The tensor product of two bialgebras is canonically a bialgebra, as the following proposition shows:

Proposition 1.3.25. *Let A and B be two $\mathbf{k}$-bialgebras. Then, $A \otimes B$ is both a $\mathbf{k}$-algebra and a $\mathbf{k}$-coalgebra (by Definition 1.3.3). These two structures, combined, turn $A \otimes B$ into a $\mathbf{k}$-bialgebra.*

Exercise 1.3.26. (a) Prove Proposition 1.3.25.

(b) Let G and H be two groups. Show that the $\mathbf{k}$-bialgebra $\mathbf{k}G \otimes \mathbf{k}H$ (defined as in Proposition 1.3.25) is isomorphic to the $\mathbf{k}$-bialgebra $\mathbf{k}[G \times H]$. (The notation $\mathbf{k}[S]$ is a synonym for $\mathbf{k}S$.)

1.4. **Antipodes and Hopf algebras.** There is one more piece of structure needed to make a bialgebra a Hopf algebra, although it will come for free in the connected graded case.

Definition 1.4.1. For any coalgebra C and algebra A, one can endow the $\mathbf{k}$-module $\operatorname{Hom}(C, A)$ (which consists of all $\mathbf{k}$-linear maps from C to A) with an associative algebra structure called the *convolution* *algebra*: Define the product $f \star g$ of two maps f, g in $\mathrm{Hom}(C, A)$ by $(f \star g)(c) = \sum f(c_1)g(c_2)$, using the Sweedler notation[20] $\Delta(c) = \sum c_1 \otimes c_2$. Equivalently, $f \star g$ is the composite

$$\begin{tikzcd} C \arrow[r, "\Delta"] & C \otimes C \arrow[r, "f \otimes g"] & A \otimes A \arrow[r, "m"] & A . \end{tikzcd}$$

The associativity of this multiplication $\star$ is easy to check (see Exercise 1.4.2 below).

The map $u \circ \epsilon$ is a two-sided identity element for $\star$, meaning that every $f \in \mathrm{Hom}(C, A)$ satisfies

$$\sum f(c_1)\epsilon(c_2) = f(c) = \sum \epsilon(c_1)f(c_2)$$

for all $c \in C$. One sees this by adding a top row to (1.2.2):

(1.4.1)
$$\begin{tikzcd} A \otimes \mathbf{k} \arrow[r] & A & \mathbf{k} \otimes A \arrow[l] \\ C \otimes \mathbf{k} \arrow[r] \arrow[u, "f \otimes \mathrm{id}"] & C \arrow[u, "f"] & \mathbf{k} \otimes C \arrow[l] \arrow[u, "\mathrm{id} \otimes f"'] \\ C \otimes C \arrow[u, "\mathrm{id} \otimes \epsilon"] & C \arrow[l, "\Delta"] \arrow[r, "\Delta"'] \arrow[u, "\mathrm{id}"] & C \otimes C \arrow[u, "\epsilon \otimes \mathrm{id}"'] \end{tikzcd}$$

In particular, when one has a bialgebra A, the convolution product $\star$ gives an associative algebra structure on $\mathrm{End}(A) := \mathrm{Hom}(A, A)$.

Exercise 1.4.2. Let C be a $\mathbf{k}$-coalgebra and A be a $\mathbf{k}$-algebra. Show that the binary operation $\star$ on $\mathrm{Hom}(C, A)$ is associative.

The product $f \star g$ of two elements f and g in a convolution algebra $\mathrm{Hom}(C, A)$ is often called their *convolution*.

The following simple (but useful) property of convolution algebras says essentially that the $\mathbf{k}$-algebra $(\mathrm{Hom}(C, A), \star)$ is a covariant functor in A and a contravariant functor in C, acting on morphisms by pre- and post-composition:

Proposition 1.4.3. *Let C and C' be two $\mathbf{k}$-coalgebras, and let A and A' be two $\mathbf{k}$-algebras. Let $\gamma : C \to C'$ be a $\mathbf{k}$-coalgebra morphism. Let $\alpha : A \to A'$ be a $\mathbf{k}$-algebra morphism.*

The map

$$\mathrm{Hom}(C', A) \to \mathrm{Hom}(C, A'), \qquad f \mapsto \alpha \circ f \circ \gamma$$

is a $\mathbf{k}$-algebra homomorphism from the convolution algebra $(\mathrm{Hom}(C', A), \star)$ to the convolution algebra $(\mathrm{Hom}(C, A'), \star)$.

Proof of Proposition 1.4.3. Denote this map by φ. We must show that φ is a $\mathbf{k}$-algebra homomorphism.

Recall that α is an algebra morphism; thus, $\alpha \circ m_A = m_{A'} \circ (\alpha \otimes \alpha)$ and $\alpha \circ u_A = u_{A'}$. Also, γ is a coalgebra morphism; thus, $\Delta_{C'} \circ \gamma = (\gamma \otimes \gamma) \circ \Delta_C$ and $\epsilon_{C'} \circ \gamma = \epsilon_C$.

Now, the definition of φ yields $\varphi(u_A \circ \epsilon_{C'}) = \underbrace{\alpha \circ u_A}_{=u_{A'}} \circ \underbrace{\epsilon_{C'} \circ \gamma}_{=\epsilon_C} = u_{A'} \circ \epsilon_C$; in other words, φ sends the unity of the algebra $(\mathrm{Hom}(C', A), \star)$ to the unity of the algebra $(\mathrm{Hom}(C, A'), \star)$.

[20] See the paragraph around (1.2.3) for the meaning of this notation.

Furthermore, every $f \in \operatorname{Hom}(C', A)$ and $g \in \operatorname{Hom}(C', A)$ satisfy

$$\begin{aligned}
\varphi(f \star g) &= \alpha \circ \underbrace{(f \star g)}_{= m_A \circ (f \otimes g) \circ \Delta_{C'}} \circ \gamma \\
&= \underbrace{\alpha \circ m_A}_{= m_{A'} \circ (\alpha \otimes \alpha)} \circ (f \otimes g) \circ \underbrace{\Delta_{C'} \circ \gamma}_{= (\gamma \otimes \gamma) \circ \Delta_C} \\
&= m_{A'} \circ \underbrace{(\alpha \otimes \alpha) \circ (f \otimes g) \circ (\gamma \otimes \gamma)}_{= (\alpha \circ f \circ \gamma) \otimes (\alpha \circ g \circ \gamma)} \circ \Delta_C \\
&= m_{A'} \circ ((\alpha \circ f \circ \gamma) \otimes (\alpha \circ g \circ \gamma)) \circ \Delta_C \\
&= \underbrace{(\alpha \circ f \circ \gamma)}_{= \varphi(f)} \star \underbrace{(\alpha \circ g \circ \gamma)}_{= \varphi(g)} = \varphi(f) \star \varphi(g). \qquad (1.4.2)
\end{aligned}$$

Thus, φ is a $\mathbf{k}$-algebra homomorphism (since φ is a $\mathbf{k}$-linear map and sends the unity of the algebra $(\operatorname{Hom}(C', A), \star)$ to the unity of the algebra $(\operatorname{Hom}(C, A'), \star)$). □

Exercise 1.4.4. Let C and D be two $\mathbf{k}$-coalgebras, and let A and B be two $\mathbf{k}$-algebras. Prove that:

(a) If $f : C \to A$, $f' : C \to A$, $g : D \to B$ and $g' : D \to B$ are four $\mathbf{k}$-linear maps, then

$$(f \otimes g) \star (f' \otimes g') = (f \star f') \otimes (g \star g')$$

in the convolution algebra $\operatorname{Hom}(C \otimes D, A \otimes B)$.

(b) Let R be the $\mathbf{k}$-linear map $(\operatorname{Hom}(C, A), \star) \otimes (\operatorname{Hom}(D, B), \star) \to (\operatorname{Hom}(C \otimes D, A \otimes B), \star)$ which sends every tensor $f \otimes g \in (\operatorname{Hom}(C, A), \star) \otimes (\operatorname{Hom}(D, B), \star)$ to the map $f \otimes g : C \otimes D \to A \otimes B$. (Notice that the tensor $f \otimes g$ and the map $f \otimes g$ are different things which happen to be written in the same way.) Then, R is a $\mathbf{k}$-algebra homomorphism.

Exercise 1.4.5. Let C and D be two $\mathbf{k}$-coalgebras. Let A be a $\mathbf{k}$-algebra. Let Φ be the canonical $\mathbf{k}$-module isomorphism $\operatorname{Hom}(C \otimes D, A) \to \operatorname{Hom}(C, \operatorname{Hom}(D, A))$ (defined by $((\Phi(f))(c))(d) = f(c \otimes d)$ for all $f \in \operatorname{Hom}(C \otimes D, A)$, $c \in C$ and $d \in D$). Prove that Φ is a $\mathbf{k}$-algebra isomorphism

$$(\operatorname{Hom}(C \otimes D, A), \star) \to (\operatorname{Hom}(C, (\operatorname{Hom}(D, A), \star)), \star).$$

Definition 1.4.6. A bialgebra A is called a *Hopf algebra* if there is an element S (called an *antipode* for A) in $\operatorname{End}(A)$ which is a 2-sided inverse under $\star$ for the identity map id_A. In other words, this diagram commutes:

(1.4.3)

$$\begin{tikzcd}
 & A \otimes A \arrow[rr, "S \otimes \mathrm{id}_A"] & & A \otimes A \arrow[dr, "m"] & \\
A \arrow[ur, "\Delta"] \arrow[rr, "\epsilon"] \arrow[dr, "\Delta"'] & & \mathbf{k} \arrow[rr, "u"] & & A \\
 & A \otimes A \arrow[rr, "\mathrm{id}_A \otimes S"'] & & A \otimes A \arrow[ur, "m"'] &
\end{tikzcd}$$

Or equivalently, if we follow the Sweedler notation in writing $\Delta(a) = \sum a_1 \otimes a_2$, then

$$\sum_{(a)} S(a_1) a_2 = u(\epsilon(a)) = \sum_{(a)} a_1 S(a_2). \qquad (1.4.4)$$

Example 1.4.7. For a group algebra $\mathbf{k}G$, one can define an antipode $\mathbf{k}$-linearly via $S(t_g) = t_{g^{-1}}$. The top pentagon in the above diagram commutes because

$$(S \star \mathrm{id})(t_g) = m((S \otimes \mathrm{id})(t_g \otimes t_g)) = S(t_g) t_g = t_{g^{-1}} t_g = t_e = (u \circ \epsilon)(t_g).$$

Note that when it exists, the antipode S is unique, as with all 2-sided inverses in associative algebras: if S, S' are both 2-sided $\star$-inverses to id_A then

$$S' = (u \circ \epsilon) \star S' = (S \star \mathrm{id}_A) \star S' = S \star (\mathrm{id}_A \star S') = S \star (u \circ \epsilon) = S.$$

Thus, we can speak of "*the antipode*" of a Hopf algebra.

Unlike the comultiplication Δ, the antipode S of a Hopf algebra is not always an algebra homomorphism. It is instead an algebra *anti-homomorphism*, a notion we shall now introduce:

Definition 1.4.8. (a) For any two $\mathbf{k}$-modules U and V, we let $T_{U,V} : U \otimes V \to V \otimes U$ be the $\mathbf{k}$-linear map $U \otimes V \to V \otimes U$ sending every $u \otimes v$ to $v \otimes u$. This map $T_{U,V}$ is called the *twist map* for U and V.

(b) A $\mathbf{k}$-*algebra anti-homomorphism* means a $\mathbf{k}$-linear map $f : A \to B$ between two $\mathbf{k}$-algebras A and B which satisfies $f \circ m_A = m_B \circ (f \otimes f) \circ T_{A,A}$ and $f \circ u_A = u_B$.

(c) A $\mathbf{k}$-*coalgebra anti-homomorphism* means a $\mathbf{k}$-linear map $f : C \to D$ between two $\mathbf{k}$-coalgebras C and D which satisfies $\Delta_D \circ f = T_{D,D} \circ (f \otimes f) \circ \Delta_C$ and $\epsilon_D \circ f = \epsilon_C$.

(d) A $\mathbf{k}$-*algebra anti-endomorphism* of a $\mathbf{k}$-algebra A means a $\mathbf{k}$-algebra anti-homomorphism from A to A.

(e) A $\mathbf{k}$-*coalgebra anti-endomorphism* of a $\mathbf{k}$-coalgebra C means a $\mathbf{k}$-coalgebra anti-homomorphism from C to C.

Parts (b) and (c) of Definition 1.4.8 can be restated in terms of elements:

- A $\mathbf{k}$-linear map $f : A \to B$ between two $\mathbf{k}$-algebras A and B is a $\mathbf{k}$-algebra anti-homomorphism if and only if it satisfies $f(ab) = f(b) f(a)$ for all $a, b \in A$ as well as $f(1) = 1$.
- A $\mathbf{k}$-linear map $f : C \to D$ between two $\mathbf{k}$-coalgebras C and D is a $\mathbf{k}$-coalgebra anti-homomorphism if and only if it satisfies $\sum_{(f(c))} (f(c))_1 \otimes (f(c))_2 = \sum_{(c)} f(c_2) \otimes f(c_1)$ and $\epsilon(f(c)) = \epsilon(c)$ for all $c \in C$.

Example 1.4.9. Let $n \in \mathbb{N}$, and consider the $\mathbf{k}$-algebra $\mathbf{k}^{n\times n}$ of $n \times n$-matrices over $\mathbf{k}$. The map $\mathbf{k}^{n\times n} \to \mathbf{k}^{n\times n}$ that sends each matrix A to its transpose A^T is a $\mathbf{k}$-algebra anti-endomorphism of $\mathbf{k}^{n\times n}$.

We warn the reader that the composition of two $\mathbf{k}$-algebra anti-homomorphisms is not generally a $\mathbf{k}$-algebra anti-homomorphism again, but rather a $\mathbf{k}$-algebra homomorphism. The same applies to coalgebra anti-homomorphisms. Other than that, however, anti-homomorphisms share many of the helpful properties of homomorphisms. In particular, two $\mathbf{k}$-algebra anti-homomorphisms are identical if they agree on a generating set of their domain. Thus, the next proposition is useful when one wants to check that a certain map *is* the antipode in a particular Hopf algebra, by checking it on an algebra generating set.

Proposition 1.4.10. *The antipode S in a Hopf algebra A is an algebra anti-endomorphism: $S(1) = 1$, and $S(ab) = S(b)S(a)$ for all a, b in A.*

Proof. This is surprisingly nontrivial; the following argument comes from [213, proof of Proposition 4.0.1].

Since Δ is an algebra morphism, one has $\Delta(1) = 1 \otimes 1$, and therefore $1 = u\epsilon(1) = S(1) \cdot 1 = S(1)$.

To show $S(ab) = S(b)S(a)$, consider $A \otimes A$ as a coalgebra and A as an algebra. Then $\operatorname{Hom}(A \otimes A, A)$ is an associative algebra with a convolution product $\circledast$ (to be distinguished from the convolution $\star$ on $\operatorname{End}(A)$), having two-sided identity element $u_A \epsilon_{A\otimes A}$. We define three elements f, g, h of $\operatorname{Hom}(A \otimes A, A)$ by

$$\begin{aligned} f(a \otimes b) &= ab, \\ g(a \otimes b) &= S(b)S(a), \\ h(a \otimes b) &= S(ab). \end{aligned}$$

We will show that these three elements have the property that

$$h \circledast f = u_A \epsilon_{A\otimes A} = f \circledast g, \tag{1.4.5}$$

which would then show the desired equality $h = g$ via associativity:

$$h = h \circledast (u_A \epsilon_{A\otimes A}) = h \circledast (f \circledast g) = (h \circledast f) \circledast g = (u_A \epsilon_{A\otimes A}) \circledast g = g.$$

So we evaluate the three elements in (1.4.5) on $a \otimes b$. To do so, we use Sweedler notation – i.e., we assume $\Delta(a) = \sum_{(a)} a_1 \otimes a_2$ and $\Delta(b) = \sum_{(b)} b_1 \otimes b_2$, and hence $\Delta(ab) = \sum_{(a),(b)} a_1 b_1 \otimes a_2 b_2$ (by (1.3.6)); then,

$$(u_A \epsilon_{A\otimes A})(a \otimes b) = u_A(\epsilon_A(a)\epsilon_A(b)) = u_A(\epsilon_A(ab)).$$

$$\begin{aligned}(h \circledast f)(a \otimes b) &= \sum_{(a),(b)} h(a_1 \otimes b_1) f(a_2 \otimes b_2) \\ &= \sum_{(a),(b)} S(a_1 b_1) a_2 b_2 \\ &= (S \star \mathrm{id}_A)(ab) = u_A(\epsilon_A(ab)).\end{aligned}$$

$$\begin{aligned}(f \circledast g)(a \otimes b) &= \sum_{(a),(b)} f(a_1 \otimes b_1) g(a_2 \otimes b_2) \\ &= \sum_{(a),(b)} a_1 b_1 S(b_2) S(a_2) \\ &= \sum_{(a)} a_1 \cdot (\mathrm{id}_A \star S)(b) \cdot S(a_2) \\ &= u_A(\epsilon_A(b)) \sum_{(a)} a_1 S(a_2) = u_A(\epsilon_A(b)) u_A(\epsilon_A(a)) = u_A(\epsilon_A(ab)).\end{aligned}$$

These results are equal, so that (1.4.5) holds, and we conclude that $h = g$ as explained above. □

Remark 1.4.11. Recall from Remark 1.3.9 that the comultiplication on a bialgebra A allows one to define an A-module structure on the tensor product $M \otimes N$ of two A-modules M, N. Similarly, the anti-endomorphism S in a Hopf algebra allows one to turn *left* A-modules into *right* A-modules, or vice-versa.[21] E.g., left A-modules M naturally have a right A-module structure on the dual $\mathbf{k}$-module $M^* := \mathrm{Hom}(M, \mathbf{k})$, defined via $(fa)(m) := f(am)$ for f in M^* and a in A. The antipode S can be used to turn this back into a left A-module M^*, via $(af)(m) = f(S(a)m)$.

For groups G and left $\mathbf{k}G$-modules (group representations) M, this is how one defines the *contragredient action* of G on M^*, namely t_g acts as $(t_g f)(m) = f(t_{g^{-1}} m)$.

More generally, if A is a Hopf algebra and M and N are two left A-modules, then $\mathrm{Hom}\,(M, N)$ (the Hom here means $\mathrm{Hom}_{\mathbf{k}}$, not Hom_A) canonically becomes a left A-module by setting

$$(af)(m) = \sum_{(a)} a_1 f(S(a_2) m) \qquad \text{for all } a \in A,\ f \in \mathrm{Hom}\,(M, N) \text{ and } m \in M.$$

[22] When A is the group algebra $\mathbf{k}G$ of a group G, this leads to

$$(t_g f)(m) = t_g f\left(t_{g^{-1}} m\right) \qquad \text{for all } g \in G,\ f \in \mathrm{Hom}\,(M, N) \text{ and } m \in M.$$

This is precisely how one commonly makes $\mathrm{Hom}\,(M, N)$ a representation of G for two representations M and N.

[21] Be warned that these two transformations are not mutually inverse! Turning a left A-module into a right one and then again into a left one using the antipode might lead to a non-isomorphic A-module, unless the antipode S satisfies $S^2 = \mathrm{id}$.

[22] In more abstract terms, this A-module structure is given by the composition

$$\begin{tikzcd} A \arrow[r, "\Delta"] & A \otimes A \arrow[r, "\mathrm{id}_A \otimes S"] & A \otimes A^{\mathrm{op}} \arrow[r] & \mathrm{End}\,(\mathrm{Hom}\,(M, N)) \end{tikzcd},$$

where the last arrow is the morphism

$$\begin{aligned} A \otimes A^{\mathrm{op}} &\longrightarrow \mathrm{End}\,(\mathrm{Hom}\,(M, N)), \\ a \otimes b &\longmapsto (f \mapsto (M \to N,\ m \mapsto af(bm))). \end{aligned}$$

Here, A^{op} denotes the *opposite algebra* of A, which is the $\mathbf{k}$-algebra differing from A only in the multiplication being twisted (the product of a and b in A^{op} is defined to be the product of b and a in A). As $\mathbf{k}$-modules, $A^{\mathrm{op}} = A$, but we prefer to use A^{op} instead of A here to ensure that all morphisms in the above composition are algebra morphisms.

Along the same lines, whenever A is a $\mathbf{k}$-bialgebra, we are supposed to think of the counit $A \xrightarrow{\epsilon} \mathbf{k}$ as giving a way to make $\mathbf{k}$ into a *trivial* A-module. This A-module $\mathbf{k}$ behaves as one would expect: the canonical isomorphisms $\mathbf{k} \otimes M \to M$, $M \otimes \mathbf{k} \to M$ and (if A is a Hopf algebra) $\operatorname{Hom}(M, \mathbf{k}) \to M^*$ are A-module isomorphisms for any A-module M.

Corollary 1.4.12. *Let A be a commutative Hopf algebra. Then, its antipode is an involution: $S^2 = \mathrm{id}_A$.*

Proof. One checks that $S^2 = S \circ S$ is a right $\star$-inverse to S, as follows:

$$\begin{aligned}(S \star S^2)(a) &= \sum_{(a)} S(a_1) S^2(a_2) \\ &= S\left(\sum_{(a)} S(a_2) a_1\right) \qquad \text{(by Proposition 1.4.10)} \\ &= S\left(\sum_{(a)} a_1 S(a_2)\right) \qquad \text{(by commutativity of } A\text{)} \\ &= S\left(u(\epsilon(a))\right) \\ &= u(\epsilon(a)) \qquad \left(\text{since } S(1) = 1 \text{ by Proposition 1.4.10}\right).\end{aligned}$$

Since S itself is the $\star$-inverse to id_A, this shows that $S^2 = \mathrm{id}_A$. □

Remark 1.4.13. We won't need it, but it is easy to adapt the above proof to show that $S^2 = \mathrm{id}_A$ also holds for *cocommutative* Hopf algebras (the dual notion to commutativity; see Definition 1.5.2 below for the precise definition); see [157, Corollary 1.5.12] or [213, Proposition 4.0.1 6)] or Exercise 1.5.13 below. For a general Hopf algebra which is not finite-dimensional over a field $\mathbf{k}$, the antipode S may not even have finite order, even in the connected graded setting. E.g., Aguiar and Sottile [7] show that the Malvenuto-Reutenauer Hopf algebra of permutations has antipode of infinite order. In general, antipodes need not even be invertible [214].

Proposition 1.4.14. *Let A and B be two Hopf algebras. Then, the $\mathbf{k}$-bialgebra $A \otimes B$ (defined as in Proposition 1.3.25) is a Hopf algebra. The antipode of this Hopf algebra $A \otimes B$ is the map $S_A \otimes S_B : A \otimes B \to A \otimes B$, where S_A and S_B are the antipodes of the Hopf algebras A and B.*

Exercise 1.4.15. Prove Proposition 1.4.14.

In our frequent setting of connected graded bialgebras, antipodes come for free.

Proposition 1.4.16. *A connected graded bialgebra A has a unique antipode S, which is a graded map $A \xrightarrow{S} A$, endowing it with a Hopf structure.*

Proof. Let us try to define a ($\mathbf{k}$-linear) left $\star$-inverse S to id_A on each homogeneous component A_n, via induction on n.

In the base case $n = 0$, Proposition 1.4.10 and its proof show that one must define $S(1) = 1$ so S is the identity on $A_0 = \mathbf{k}$.

In the inductive step, recall from Exercise 1.3.20(h) that a homogeneous element a of degree $n > 0$ has $\Delta(a) = a \otimes 1 + \sum a'_1 \otimes a'_2$, with each $\deg(a'_1) < n$. (Here $\sum a'_1 \otimes a'_2$ stands for a sum of tensors $a'_{1,k} \otimes a'_{2,k}$, with each $a'_{1,k}$ being homogeneous of degree $\deg(a'_{1,k}) < n$. This is a slight variation on Sweedler notation.) Hence in order to have $S \star \mathrm{id}_A = u\epsilon$, one must define $S(a)$ in such a way that $S(a) \cdot 1 + \sum S(a'_1) a'_2 = u\epsilon(a) = 0$ and hence $S(a) := -\sum S(a'_1) a'_2$, where $S(a'_1)$ have already been uniquely defined by induction (since $\deg(a'_{1,k}) < n$). This does indeed define such a left $\star$-inverse S to id_A, by induction. It is also a graded map by induction.

The same argument shows how to define a right $\star$-inverse S' to id_A. Then $S = S'$ is a two-sided $\star$-inverse to id_A by the associativity of $\star$. □

Here is another consequence of the fact that $S(1) = 1$.

Proposition 1.4.17. *In bialgebras, primitive elements x have $\epsilon(x) = 0$, and in Hopf algebras, they have $S(x) = -x$.*

Proof. In a bialgebra, $\epsilon(1) = 1$. Hence $\Delta(x) = 1 \otimes x + x \otimes 1$ implies via (1.2.2) that $1 \cdot \epsilon(x) + \epsilon(1)x = x$, so $\epsilon(x) = 0$. It also implies via (1.4.3) that $S(x)1 + S(1)x = u\epsilon(x) = u(0) = 0$, so $S(x) = -x$. □

Thus, whenever A is a Hopf algebra generated as an algebra by its primitive elements, S is its unique $\mathbf{k}$-algebra anti-endomorphism that negates all primitive elements.

Example 1.4.18. The tensor and symmetric algebras $T(V)$ and $\operatorname{Sym}(V)$ are each generated by V, and each element of V is primitive when regarded as an element of either of them. Hence one has in $T(V)$ that

$$S(x_{i_1}x_{i_2}\cdots x_{i_k}) = (-x_{i_k})\cdots(-x_{i_2})(-x_{i_1}) = (-1)^k x_{i_k}\cdots x_{i_2}x_{i_1} \tag{1.4.6}$$

for each word $(i_1, \ldots, i_k)$ in the alphabet I if V is a free $\mathbf{k}$-module with basis $\{x_i\}_{i\in I}$. The same holds in $\operatorname{Sym}(V)$ for each multiset $\{i_1, \ldots, i_k\}_{\text{multiset}}$, recalling that the monomials are now commutative. In other words, for a commutative polynomial $f(x_1, x_2, \ldots, x_n)$ in $\operatorname{Sym}(V)$, the antipode S sends $f(x_1, x_2, \ldots, x_n)$ to $f(-x_1, -x_2, \ldots, -x_n)$, negating all the variables.

The antipode for a connected graded Hopf algebra has an interesting formula due to Takeuchi [214], reminiscent of P. Hall's formula for the Möbius function of a poset[23]. For the sake of stating this, consider (for every $k \in \mathbb{N}$) the k-fold *tensor power* $A^{\otimes k} = A \otimes \cdots \otimes A$ (defined in Example 1.1.2) and define *iterated multiplication and comultiplication* maps

$$A^{\otimes k} \xrightarrow{m^{(k-1)}} A \qquad \text{and} \qquad A \xrightarrow{\Delta^{(k-1)}} A^{\otimes k}$$

by induction over k, setting $m^{(-1)} = u$, $\Delta^{(-1)} = \epsilon$, $m^{(0)} = \Delta^{(0)} = \mathrm{id}_A$, and

$$\begin{aligned} m^{(k)} &= m \circ (\mathrm{id}_A \otimes m^{(k-1)}) && \text{for every } k \geq 1; \\ \Delta^{(k)} &= (\mathrm{id}_A \otimes \Delta^{(k-1)}) \circ \Delta && \text{for every } k \geq 1. \end{aligned}$$

Using associativity and coassociativity, one can see that for $k \geq 1$ these maps also satisfy

$$\begin{aligned} m^{(k)} &= m \circ (m^{(k-1)} \otimes \mathrm{id}_A) && \text{for every } k \geq 1; \\ \Delta^{(k)} &= (\Delta^{(k-1)} \otimes \mathrm{id}_A) \circ \Delta && \text{for every } k \geq 1 \end{aligned}$$

(so we could just as well have used $\mathrm{id}_A \otimes m^{(k-1)}$ instead of $m^{(k-1)} \otimes \mathrm{id}_A$ in defining them) and further symmetry properties (see Exercise 1.4.19 and Exercise 1.4.20). They are how one gives meaning to the right sides of these equations:

$$m^{(k)}(a^{(1)} \otimes \cdots \otimes a^{(k+1)}) = a^{(1)} \cdots a^{(k+1)};$$

$$\Delta^{(k)}(b) = \sum b_1 \otimes \cdots \otimes b_{k+1} \text{ in Sweedler notation.}$$

Exercise 1.4.19. Let A be a $\mathbf{k}$-algebra. Let us define, for every $k \in \mathbb{N}$, a $\mathbf{k}$-linear map $m^{(k)} : A^{\otimes(k+1)} \to A$. Namely, we define these maps by induction over k, with the induction base $m^{(0)} = \mathrm{id}_A$, and with the induction step $m^{(k)} = m \circ \left(\mathrm{id}_A \otimes m^{(k-1)}\right)$ for every $k \geq 1$. (This generalizes our definition of $m^{(k)}$ for Hopf algebras A given above, except for $m^{(-1)}$ which we have omitted.)

(a) Show that $m^{(k)} = m \circ \left(m^{(i)} \otimes m^{(k-1-i)}\right)$ for every $k \geq 0$ and $0 \leq i \leq k-1$.
(b) Show that $m^{(k)} = m \circ \left(m^{(k-1)} \otimes \mathrm{id}_A\right)$ for every $k \geq 1$.
(c) Show that $m^{(k)} = m^{(k-1)} \circ (\mathrm{id}_{A^{\otimes i}} \otimes m \otimes \mathrm{id}_{A^{\otimes(k-1-i)}})$ for every $k \geq 0$ and $0 \leq i \leq k-1$.
(d) Show that $m^{(k)} = m^{(k-1)} \circ (\mathrm{id}_{A^{\otimes(k-1)}} \otimes m) = m^{(k-1)} \circ (m \otimes \mathrm{id}_{A^{\otimes(k-1)}})$ for every $k \geq 1$.

Exercise 1.4.20. Let C be a $\mathbf{k}$-coalgebra. Let us define, for every $k \in \mathbb{N}$, a $\mathbf{k}$-linear map $\Delta^{(k)} : C \to C^{\otimes(k+1)}$. Namely, we define these maps by induction over k, with the induction base $\Delta^{(0)} = \mathrm{id}_C$, and with the induction step $\Delta^{(k)} = \left(\mathrm{id}_C \otimes \Delta^{(k-1)}\right) \circ \Delta$ for every $k \geq 1$. (This generalizes our definition of $\Delta^{(k)}$ for Hopf algebras A given above, except for $\Delta^{(-1)}$ which we have omitted.)

(a) Show that $\Delta^{(k)} = \left(\Delta^{(i)} \otimes \Delta^{(k-1-i)}\right) \circ \Delta$ for every $k \geq 0$ and $0 \leq i \leq k-1$.
(b) Show that $\Delta^{(k)} = \left(\Delta^{(k-1)} \otimes \mathrm{id}_C\right) \circ \Delta$ for every $k \geq 1$.
(c) Show that $\Delta^{(k)} = (\mathrm{id}_{C^{\otimes i}} \otimes \Delta \otimes \mathrm{id}_{C^{\otimes(k-1-i)}}) \circ \Delta^{(k-1)}$ for every $k \geq 0$ and $0 \leq i \leq k-1$.
(d) Show that $\Delta^{(k)} = (\mathrm{id}_{C^{\otimes(k-1)}} \otimes \Delta) \circ \Delta^{(k-1)} = (\Delta \otimes \mathrm{id}_{C^{\otimes(k-1)}}) \circ \Delta^{(k-1)}$ for every $k \geq 1$.

[23] In fact, for incidence Hopf algebras, Takeuchi's formula generalizes Hall's formula– see Corollary 7.2.3.

Remark 1.4.21. Exercise 1.4.19 holds more generally for nonunital associative algebras A (that is, $\mathbf{k}$-modules A equipped with a $\mathbf{k}$-linear map $m : A \otimes A \to A$ such that the diagram (1.1.1) is commutative, but not necessarily admitting a unit map u). Similarly, Exercise 1.4.20 holds for non-counital coassociative coalgebras C. The existence of a unit in A, respectively a counit in C, allows slightly extending these two exercises by additionally introducing maps $m^{(-1)} = u : \mathbf{k} \to A$ and $\Delta^{(-1)} = \epsilon : C \to \mathbf{k}$; however, not much is gained from this extension.[24]

Exercise 1.4.22. For every $k \in \mathbb{N}$ and every $\mathbf{k}$-bialgebra H, consider the map $\Delta_H^{(k)} : H \to H^{\otimes(k+1)}$ (this is the map $\Delta^{(k)}$ defined as in Exercise 1.4.20 for $C = H$), and the map $m_H^{(k)} : H^{\otimes(k+1)} \to H$ (this is the map $m^{(k)}$ defined as in Exercise 1.4.19 for $A = H$).

Let H be a $\mathbf{k}$-bialgebra. Let $k \in \mathbb{N}$. Show that:[25]

(a) The map $m_H^{(k)} : H^{\otimes(k+1)} \to H$ is a $\mathbf{k}$-coalgebra homomorphism.

(b) The map $\Delta_H^{(k)} : H \to H^{\otimes(k+1)}$ is a $\mathbf{k}$-algebra homomorphism.

(c) We have $m_{H^{\otimes(k+1)}}^{(\ell)} \circ \left(\Delta_H^{(k)}\right)^{\otimes(\ell+1)} = \Delta_H^{(k)} \circ m_H^{(\ell)}$ for every $\ell \in \mathbb{N}$.

(d) We have $\left(m_H^{(\ell)}\right)^{\otimes(k+1)} \circ \Delta_{H^{\otimes(\ell+1)}}^{(k)} = \Delta_H^{(k)} \circ m_H^{(\ell)}$ for every $\ell \in \mathbb{N}$.

The iterated multiplication and comultiplication maps allow explicitly computing the convolution of multiple maps; the following formula will often be used without explicit mention:

Exercise 1.4.23. Let C be a $\mathbf{k}$-coalgebra, and A be a $\mathbf{k}$-algebra. Let $k \in \mathbb{N}$. Let $f_1, f_2, \ldots, f_k$ be k elements of $\operatorname{Hom}(C, A)$. Show that

$$f_1 \star f_2 \star \cdots \star f_k = m_A^{(k-1)} \circ (f_1 \otimes f_2 \otimes \cdots \otimes f_k) \circ \Delta_C^{(k-1)}.$$

We are now ready to state Takeuchi's formula for the antipode:

Proposition 1.4.24. *In a connected graded Hopf algebra A, the antipode has formula*

$$\begin{aligned} S &= \sum_{k \geq 0} (-1)^k m^{(k-1)} f^{\otimes k} \Delta^{(k-1)} \\ &= u\epsilon - f + m \circ f^{\otimes 2} \circ \Delta - m^{(2)} \circ f^{\otimes 3} \circ \Delta^{(2)} + \cdots \end{aligned} \tag{1.4.7}$$

where $f := \operatorname{id}_A - u\epsilon$ in $\operatorname{End}(A)$.

Proof. We argue as in [214, proof of Lemma 14] or [7, §5]. For any f in $\operatorname{End}(A)$, the following explicit formula expresses its k-fold convolution power $f^{\star k} := f \star \cdots \star f$ in terms of its tensor powers $f^{\otimes k} := f \otimes \cdots \otimes f$ (according to Exercise 1.4.23):

$$f^{\star k} = m^{(k-1)} \circ f^{\otimes k} \circ \Delta^{(k-1)}.$$

Therefore any f annihilating A_0 will be *locally $\star$-nilpotent* on A, meaning that for each n one has that A_n is annihilated by $f^{\star m}$ for every $m > n$: homogeneity forces that for a in A_n, every summand of $\Delta^{(m-1)}(a)$ must contain among its m tensor factors at least one factor lying in A_0, so each summand is annihilated by $f^{\otimes m}$, and $f^{\star m}(a) = 0$.

In particular such f have the property that $u\epsilon + f$ has as two-sided $\star$-inverse

$$\begin{aligned} (u\epsilon + f)^{\star(-1)} &= u\epsilon - f + f \star f - f \star f \star f + \cdots \\ &= \sum_{k \geq 0} (-1)^k f^{\star k} = \sum_{k \geq 0} (-1)^k m^{(k-1)} \circ f^{\otimes k} \circ \Delta^{(k-1)}. \end{aligned}$$

The proposition follows upon taking $f := \operatorname{id}_A - u\epsilon$, which annihilates A_0. □

Remark 1.4.25. In fact, one can see that Takeuchi's formula applies more generally to define an antipode $A \xrightarrow{S} A$ in any (not necessarily graded) bialgebra A where the map $\operatorname{id}_A - u\epsilon$ is locally $\star$-nilpotent.

It is also worth noting that the proof of Proposition 1.4.24 gives an alternate proof of Proposition 1.4.16.

[24] The identity $m^{(k)} = m \circ \left(\operatorname{id}_A \otimes m^{(k-1)}\right)$ for a $\mathbf{k}$-algebra A still holds when $k = 0$ if it is interpreted in the right way (viz., if A is identified with $A \otimes \mathbf{k}$ using the canonical homomorphism).

[25] The following statements are taken from [167]; specifically, part (c) is [167, Lem. 1.8].

To finish our discussion of antipodes, we mention some properties (taken from [213, Lemma 4.0.3]) relating antipodes to convolutional inverses.

Proposition 1.4.26. *Let H be a Hopf algebra with antipode S.*

(a) *For any algebra A and algebra morphism $H \xrightarrow{\alpha} A$, one has $\alpha \circ S = \alpha^{\star -1}$, the convolutional inverse to α in $\operatorname{Hom}(H, A)$.*

(b) *For any coalgebra C and coalgebra morphism $C \xrightarrow{\gamma} H$, one has $S \circ \gamma = \gamma^{\star -1}$, the convolutional inverse to γ in $\operatorname{Hom}(C, H)$.*

Proof. We prove (a); the proof of (b) is similar.

For assertion (a), note that Proposition 1.4.3 (applied to H, H, H, A, id_H and α instead of C, C', A, A', γ and α) shows that the map

$$\operatorname{Hom}(H,H) \to \operatorname{Hom}(H,A), \qquad f \mapsto \alpha \circ f$$

is a $\mathbf{k}$-algebra homomorphism from the convolution algebra $(\operatorname{Hom}(H,H), \star)$ to the convolution algebra $(\operatorname{Hom}(H,A), \star)$. Denoting this homomorphism by φ, we thus have $\varphi\left((\mathrm{id}_H)^{\star -1}\right) = (\varphi(\mathrm{id}_H))^{\star -1}$ (since $\mathbf{k}$-algebra homomorphisms preserve inverses). Now,

$$\alpha \circ S = \varphi(S) = \varphi\left((\mathrm{id}_H)^{\star -1}\right) = (\varphi(\mathrm{id}_H))^{\star -1} = (\alpha \circ \mathrm{id}_H)^{\star -1} = \alpha^{\star -1}.$$

□

A rather useful consequence of Proposition 1.4.26 is the fact ([213, Lemma 4.0.4]) that a bialgebra morphism between Hopf algebras automatically respects the antipodes:

Corollary 1.4.27. *Let H_1 and H_2 be Hopf algebras with antipodes S_1 and S_2, respectively. Then, any bialgebra morphism $H_1 \xrightarrow{\beta} H_2$ is a Hopf morphism[26], that is, it commutes with the antipodes (i.e., we have $\beta \circ S_1 = S_2 \circ \beta$).*

Proof. Proposition 1.4.26(a) (applied to $H = H_1$, $S = S_1$, $A = H_2$ and $\alpha = \beta$) yields $\beta \circ S_1 = \beta^{\star -1}$. Proposition 1.4.26(b) (applied to $H = H_2$, $S = S_2$, $C = H_1$ and $\gamma = \beta$) yields $S_2 \circ \beta = \beta^{\star -1}$. Comparing these equalities shows that $\beta \circ S_1 = S_2 \circ \beta$, qed. □

Exercise 1.4.28. Prove that the antipode S of a Hopf algebra A is a coalgebra anti-endomorphism, i.e., that it satisfies $\epsilon \circ S = \epsilon$ and $\Delta \circ S = T \circ (S \otimes S) \circ \Delta$, where $T : A \otimes A \to A \otimes A$ is the twist map sending every $a \otimes b$ to $b \otimes a$.

Exercise 1.4.29. If C is a $\mathbf{k}$-coalgebra and if A is a $\mathbf{k}$-algebra, then a $\mathbf{k}$-linear map $f : C \to A$ is said to be *$\star$-invertible* if it is invertible as an element of the $\mathbf{k}$-algebra $(\operatorname{Hom}(C,A), \star)$. In this case, the multiplicative inverse $f^{\star(-1)}$ of f in $(\operatorname{Hom}(C,A), \star)$ is called the *$\star$-inverse* of f.

Recall the concepts introduced in Definition 1.4.8.

(a) If C is a $\mathbf{k}$-bialgebra, if A is a $\mathbf{k}$-algebra, and if $r : C \to A$ is a $\star$-invertible $\mathbf{k}$-algebra homomorphism, then prove that the $\star$-inverse $r^{\star(-1)}$ of r is a $\mathbf{k}$-algebra anti-homomorphism.

(b) If C is a $\mathbf{k}$-bialgebra, if A is a $\mathbf{k}$-coalgebra, and if $r : A \to C$ is a $\star$-invertible $\mathbf{k}$-coalgebra homomorphism, then prove that the $\star$-inverse $r^{\star(-1)}$ of r is a $\mathbf{k}$-coalgebra anti-homomorphism.

(c) Derive Proposition 1.4.10 from Exercise 1.4.29(a), and derive Exercise 1.4.28 from Exercise 1.4.29(b).

(d) Prove Corollary 1.4.12 again using Proposition 1.4.26.

(e) If C is a graded $\mathbf{k}$-coalgebra, if A is a graded $\mathbf{k}$-algebra, and if $r : C \to A$ is a $\star$-invertible $\mathbf{k}$-linear map that is graded, then prove that the $\star$-inverse $r^{\star(-1)}$ of r is also graded.

Exercise 1.4.30. (a) Let A be a Hopf algebra. If $P : A \to A$ is a $\mathbf{k}$-linear map such that every $a \in A$ satisfies

$$\sum_{(a)} P(a_2) \cdot a_1 = u(\epsilon(a)),$$

then prove that the antipode S of A is invertible and its inverse is P.

[26] A *Hopf morphism* (or, more officially, a *Hopf algebra morphism*, or *homomorphism of Hopf algebras*) between two Hopf algebras A and B is defined to be a bialgebra morphism $f : A \to B$ that satisfies $f \circ S_A = S_B \circ f$.

(b) Let A be a Hopf algebra. If $P : A \to A$ is a $\mathbf{k}$-linear map such that every $a \in A$ satisfies
$$\sum_{(a)} a_2 \cdot P(a_1) = u(\epsilon(a)),$$
then prove that the antipode S of A is invertible and its inverse is P.

(c) Show that the antipode of a connected graded Hopf algebra is invertible.

(Compare this exercise to [157, Lemma 1.5.11].)

Definition 1.4.31. Let C be a $\mathbf{k}$-coalgebra. A *subcoalgebra* of C means a $\mathbf{k}$-coalgebra D such that $D \subset C$ and such that the canonical inclusion map $D \to C$ is a $\mathbf{k}$-coalgebra homomorphism[27]. When $\mathbf{k}$ is a field, we can equivalently define a subcoalgebra of C as a $\mathbf{k}$-submodule D of C such that $\Delta_C(D)$ is a subset of the $\mathbf{k}$-submodule $D \otimes D$ of $C \otimes C$; however, this might no longer be equivalent when $\mathbf{k}$ is not a field[28].

Similarly, a *subbialgebra* of a bialgebra C is a $\mathbf{k}$-bialgebra D such that $D \subset C$ and such that the canonical inclusion map $D \to C$ is a $\mathbf{k}$-bialgebra homomorphism. Also, a *Hopf subalgebra* of a Hopf algebra C is a $\mathbf{k}$-Hopf algebra D such that $D \subset C$ and such that the canonical inclusion map $D \to C$ is a $\mathbf{k}$-Hopf algebra homomorphism.[29]

Exercise 1.4.32. Let C be a $\mathbf{k}$-coalgebra. Let D be a $\mathbf{k}$-submodule of C such that D is a direct summand of C as a $\mathbf{k}$-module (i.e., there exists a $\mathbf{k}$-submodule E of C such that $C = D \oplus E$). (This is automatically satisfied if $\mathbf{k}$ is a field.) Assume that $\Delta(D) \subset C \otimes D$ and $\Delta(D) \subset D \otimes C$. (Here, we are abusing the notation $C \otimes D$ to denote the $\mathbf{k}$-submodule of $C \otimes C$ spanned by tensors of the form $c \otimes d$ with $c \in C$ and $d \in D$; similarly, $D \otimes C$ should be understood.) Show that there is a canonically defined $\mathbf{k}$-coalgebra structure on D which makes D a subcoalgebra of C.

The next exercise is implicit in [4, §5]:

Exercise 1.4.33. Let $\mathbf{k}$ be a field. Let C be a $\mathbf{k}$-coalgebra, and let U be any $\mathbf{k}$-module. Let $f : C \to U$ be a $\mathbf{k}$-linear map. Recall the map $\Delta^{(2)} : C \to C^{\otimes 3}$ from Exercise 1.4.20. Let $K = \ker\left((\mathrm{id}_C \otimes f \otimes \mathrm{id}_C) \circ \Delta^{(2)}\right)$.

(a) Show that K is a $\mathbf{k}$-subcoalgebra of C.

(b) Show that every $\mathbf{k}$-subcoalgebra of C which is a subset of $\ker f$ must be a subset of K.

Exercise 1.4.34. (a) Let $C = \bigoplus_{n \geq 0} C_n$ be a graded $\mathbf{k}$-coalgebra, and A be any $\mathbf{k}$-algebra. Notice that C_0 itself is a $\mathbf{k}$-subcoalgebra of C. Let $h : C \to A$ be a $\mathbf{k}$-linear map such that the restriction $h \mid_{C_0}$ is a $\star$-invertible map in $\operatorname{Hom}(C_0, A)$. Prove that h is a $\star$-invertible map in $\operatorname{Hom}(C, A)$. (This is a weaker version of Takeuchi's [214, Lemma 14].)

(b) Let $A = \bigoplus_{n \geq 0} A_n$ be a graded $\mathbf{k}$-bialgebra. Notice that A_0 is a subbialgebra of A. Assume that A_0 is a Hopf algebra. Show that A is a Hopf algebra.

(c) Obtain yet another proof of Proposition 1.4.16.

Exercise 1.4.35. Let $A = \bigoplus_{n \geq 0} A_n$ be a connected graded $\mathbf{k}$-bialgebra. Let $\mathfrak{p}$ be the $\mathbf{k}$-submodule of A consisting of the primitive elements of A.

(a) If I is a two-sided coideal of A such that $I \cap \mathfrak{p} = 0$ and such that $I = \bigoplus_{n \geq 0} (I \cap A_n)$, then prove that $I = 0$.

(b) Let $f : A \to C$ be a graded surjective coalgebra homomorphism from A to a graded $\mathbf{k}$-coalgebra C. If $f \mid_{\mathfrak{p}}$ is injective, then prove that f is injective.

(c) Assume that $\mathbf{k}$ is a field. Show that the claim of Exercise 1.4.35(b) is valid even without requiring f to be surjective.

Remark 1.4.36. Exercise 1.4.35 (b) and (c) are often used in order to prove that certain coalgebra homomorphisms are injective.

The word “bialgebra” can be replaced by “coalgebra” in Exercise 1.4.35, provided that the notion of a connected graded coalgebra is defined correctly (namely, as a graded coalgebra such that the restriction of

[27] In this definition, we follow [162, p. 55] and [225, §6.7]; other authors may use other definitions.

[28] This is because the $\mathbf{k}$-submodule $D \otimes D$ of $C \otimes C$ is generally not isomorphic to the $\mathbf{k}$-module $D \otimes D$. See [162, p. 56] for specific counterexamples for the non-equivalence of the two notions of a subcoalgebra. Notice that the equivalence is salvaged if D is a direct summand of C as a $\mathbf{k}$-module (see Exercise 1.4.32 for this).

[29] By Corollary 1.4.27, we can also define it as a subbialgebra of C that happens to be a Hopf algebra.

ϵ to the 0-th graded component is an isomorphism), and the notion of the element 1 of a connected graded coalgebra is defined accordingly (namely, as the preimage of $1 \in \mathbf{k}$ under the restriction of ϵ to the 0-th graded component).

1.5. **Commutativity, cocommutativity.** Recall that a $\mathbf{k}$-algebra A is *commutative* if and only if all $a, b \in A$ satisfy $ab = ba$. Here is a way to restate this classical definition using tensors instead of pairs of elements:

Definition 1.5.1. A $\mathbf{k}$-algebra A is said to be *commutative* if the following diagram commutes:

(1.5.1)
$$\begin{tikzcd} A \otimes A \arrow[rr, "T"] \arrow[dr, "m"'] & & A \otimes A \arrow[dl, "m"] \\ & A & \end{tikzcd}$$

where T is the twist map $T_{A,A}$ (see Definition 1.4.8(a) for its definition).

Having thus redefined commutative algebras in terms of tensors and linear maps, we can dualize this definition (reversing all arrows) and obtain the notion of *cocommutative coalgebras*:

Definition 1.5.2. A $\mathbf{k}$-coalgebra C is said to be *cocommutative* if the following diagram commutes:

(1.5.2)
$$\begin{tikzcd} C \otimes C \arrow[rr, "T"] & & C \otimes C \\ & C \arrow[ul, "\Delta"] \arrow[ur, "\Delta"'] & \end{tikzcd}$$

where T is the twist map $T_{C,C}$ (see Definition 1.4.8(a) for its definition).

Example 1.5.3. Group algebras $\mathbf{k}G$ are always cocommutative. They are commutative if and only if G is abelian or $\mathbf{k} = 0$.

Tensor algebras $T(V)$ are always cocommutative, but not generally commutative[30].

Symmetric algebras $\operatorname{Sym}(V)$ are always cocommutative and commutative.

Homology and cohomology of H-spaces are always cocommutative and commutative *in the topologist's sense* where one reinterprets that twist map $A \otimes A \xrightarrow{T} A \otimes A$ to have the extra sign as in (1.3.3).

Note how the cocommutative Hopf algebras $T(V), \operatorname{Sym}(V)$ have much of their structure controlled by their $\mathbf{k}$-submodules V, which consist of primitive elements only (although, in general, not of all their primitive elements). This is not far from the truth in general, and closely related to Lie algebras.

Exercise 1.5.4. Recall that a *Lie algebra* over $\mathbf{k}$ is a $\mathbf{k}$-module $\mathfrak{g}$ with a $\mathbf{k}$-bilinear map $[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ that satisfies $[x, x] = 0$ for x in $\mathfrak{g}$, and the *Jacobi identity*

$$[x, [y, z]] = [[x, y], z] + [y, [x, z]], \text{ or equivalently}$$
$$[x, [y, z]] + [z, [x, y]] + [y, [z, x]] = 0$$

for all $x, y, z \in \mathfrak{g}$. This $\mathbf{k}$-bilinear map $[\cdot, \cdot]$ is called the *Lie bracket* of $\mathfrak{g}$.

(a) Check that any associative algebra A gives rise to a Lie algebra by means of the commutator operation $[a, b] := ab - ba$.

(b) If A is also a bialgebra, show that the $\mathbf{k}$-submodule of primitive elements $\mathfrak{p} \subset A$ is closed under the Lie bracket, that is, $[\mathfrak{p}, \mathfrak{p}] \subset \mathfrak{p}$, and hence forms a Lie subalgebra.

Conversely, given a Lie algebra $\mathfrak{p}$, one constructs the *universal enveloping algebra* $\mathcal{U}(\mathfrak{p}) := T(\mathfrak{p})/J$ as the quotient of the tensor algebra $T(\mathfrak{p})$ by the two-sided ideal J generated by all elements $xy - yx - [x, y]$ for x, y in $\mathfrak{p}$.

(c) Show that J is also a two-sided coideal in $T(\mathfrak{p})$ for its usual coalgebra structure, and hence the quotient $\mathcal{U}(\mathfrak{p})$ inherits the structure of a cocommutative bialgebra.

(d) Show that the antipode S on $T(\mathfrak{p})$ preserves J, meaning that $S(J) \subset J$, and hence $\mathcal{U}(\mathfrak{p})$ inherits the structure of a (cocommutative) Hopf algebra.

[30] If $\mathbf{k}$ is a field, then $T(V)$ is commutative if and only if $\dim_{\mathbf{k}} V \leq 1$.

There are theorems, discussed in [35, §3.8], [157, Chap. 5], [60, §3.2] giving various mild hypotheses in addition to cocommutativity which imply that the inclusion of the $\mathbf{k}$-module $\mathfrak{p}$ of primitives in a Hopf algebra A extends to a Hopf isomorphism $\mathcal{U}(\mathfrak{p}) \cong A$.

Exercise 1.5.5. Let C be a cocommutative $\mathbf{k}$-coalgebra. Let A be a commutative $\mathbf{k}$-algebra. Show that the convolution algebra $(\operatorname{Hom}(C, A), \star)$ is commutative (i.e., every $f, g \in \operatorname{Hom}(C, A)$ satisfy $f \star g = g \star f$).

Exercise 1.5.6. (a) Let C be a $\mathbf{k}$-coalgebra. Show that C is cocommutative if and only if its comultiplication $\Delta_C : C \to C \otimes C$ is a $\mathbf{k}$-coalgebra homomorphism.

(b) Let A be a $\mathbf{k}$-algebra. Show that A is commutative if and only if its multiplication $m_A : A \otimes A \to A$ is a $\mathbf{k}$-algebra homomorphism.

Remark 1.5.7. If C is a $\mathbf{k}$-coalgebra, then $\epsilon_C : C \to \mathbf{k}$ is always a $\mathbf{k}$-coalgebra homomorphism. Similarly, $u_A : \mathbf{k} \to A$ is a $\mathbf{k}$-algebra homomorphism whenever A is a $\mathbf{k}$-algebra.

Exercise 1.5.8. (a) Let A and B be two $\mathbf{k}$-algebras, at least one of which is commutative. Prove that the $\mathbf{k}$-algebra anti-homomorphisms from A to B are the same as the $\mathbf{k}$-algebra homomorphisms from A to B.

(b) State and prove the dual of this result.

Exercise 1.5.9. Let A be a commutative $\mathbf{k}$-algebra, and let $k \in \mathbb{N}$. The symmetric group $\mathfrak{S}_k$ acts on the k-fold tensor power $A^{\otimes k}$ by permuting the tensor factors: $\sigma(v_1 \otimes v_2 \otimes \cdots \otimes v_k) = v_{\sigma^{-1}(1)} \otimes v_{\sigma^{-1}(2)} \otimes \cdots \otimes v_{\sigma^{-1}(k)}$ for all $v_1, v_2, \ldots, v_k \in A$ and $\sigma \in \mathfrak{S}_k$. For every $\pi \in \mathfrak{S}_k$, denote by $\rho(\pi)$ the action of π on $A^{\otimes k}$ (this is an endomorphism of $A^{\otimes k}$). Show that every $\pi \in \mathfrak{S}_k$ satisfies $m^{(k-1)} \circ (\rho(\pi)) = m^{(k-1)}$. (Recall that $m^{(k-1)} : A^{\otimes k} \to A$ is defined as in Exercise 1.4.19 for $k \geq 1$, and by $m^{(-1)} = u : \mathbf{k} \to A$ for $k = 0$.)

Exercise 1.5.10. State and solve the analogue of Exercise 1.5.9 for cocommutative $\mathbf{k}$-coalgebras.

Exercise 1.5.11. (a) If H is a $\mathbf{k}$-bialgebra and A is a commutative $\mathbf{k}$-algebra, and if f and g are two $\mathbf{k}$-algebra homomorphisms $H \to A$, then prove that $f \star g$ also is a $\mathbf{k}$-algebra homomorphism $H \to A$.

(b) If H is a $\mathbf{k}$-bialgebra and A is a commutative $\mathbf{k}$-algebra, and if $f_1, f_2, \ldots, f_k$ are several $\mathbf{k}$-algebra homomorphisms $H \to A$, then prove that $f_1 \star f_2 \star \cdots \star f_k$ also is a $\mathbf{k}$-algebra homomorphism $H \to A$.

(c) If H is a Hopf algebra and A is a commutative $\mathbf{k}$-algebra, and if $f : H \to A$ is a $\mathbf{k}$-algebra homomorphism, then prove that $f \circ S : H \to A$ (where S is the antipode of H) is again a $\mathbf{k}$-algebra homomorphism, and is a $\star$-inverse to f.

(d) If A is a commutative $\mathbf{k}$-algebra, then show that $m^{(k)}$ is a $\mathbf{k}$-algebra homomorphism for every $k \in \mathbb{N}$. (The map $m^{(k)} : A^{\otimes(k+1)} \to A$ is defined as in Exercise 1.4.19.)

(e) If C' and C are two $\mathbf{k}$-coalgebras, if $\gamma : C \to C'$ is a $\mathbf{k}$-coalgebra homomorphism, if A and A' are two $\mathbf{k}$-algebras, if $\alpha : A \to A'$ is a $\mathbf{k}$-algebra homomorphism, and if $f_1, f_2, \ldots, f_k$ are several $\mathbf{k}$-linear maps $C' \to A$, then prove that

$$\alpha \circ (f_1 \star f_2 \star \cdots \star f_k) \circ \gamma = (\alpha \circ f_1 \circ \gamma) \star (\alpha \circ f_2 \circ \gamma) \star \cdots \star (\alpha \circ f_k \circ \gamma).$$

(f) If H is a commutative $\mathbf{k}$-bialgebra, and k and ℓ are two nonnegative integers, then prove that $\operatorname{id}_H^{\star k} \circ \operatorname{id}_H^{\star \ell} = \operatorname{id}_H^{\star(k\ell)}$.

(g) If H is a commutative $\mathbf{k}$-Hopf algebra, and k and ℓ are two integers, then prove that $\operatorname{id}_H^{\star k} \circ \operatorname{id}_H^{\star \ell} = \operatorname{id}_H^{\star(k\ell)}$. (These powers $\operatorname{id}_H^{\star k}$, $\operatorname{id}_H^{\star \ell}$ and $\operatorname{id}_H^{\star(k\ell)}$ are well-defined since id_H is $\star$-invertible.)

(h) State and prove the duals of parts (a)–(g) of this exercise.

Remark 1.5.12. The maps $\operatorname{id}_H^{\star k}$ for $k \in \mathbb{N}$ are known as the *Adams operators* of the bialgebra H; they are studied, inter alia, in [5]. Particular cases (and variants) of Exercise 1.5.11(f) appear in [167, Corollaire II.9] and [78, Theorem 1]. Exercise 1.5.11(f) and its dual are [135, Prop. 1.6].

Exercise 1.5.13. Prove that the antipode S of a cocommutative Hopf algebra A satisfies $S^2 = \operatorname{id}_A$. (This was a statement made in Remark 1.4.13.)

Exercise 1.5.14. Let A be a cocommutative graded Hopf algebra with antipode S. Define a $\mathbf{k}$-linear map $E : A \to A$ by having $E(a) = (\deg a) \cdot a$ for every homogeneous element a of A.

(a) Prove that for every $a \in A$, the elements $(S \star E)(a)$ and $(E \star S)(a)$ (where $\star$ denotes convolution in $\operatorname{Hom}(A, A)$) are primitive.

(b) Prove that for every primitive $p \in A$, we have $(S \star E)(p) = (E \star S)(p) = E(p)$.

(c) Prove that for every $a \in A$ and every primitive $p \in A$, we have $(S \star E)(ap) = [(S \star E)(a), p] + \epsilon(a) E(p)$, where $[u, v]$ denotes the commutator $uv - vu$ of u and v.

(d) If A is connected and $\mathbb{Q}$ is a subring of $\mathbf{k}$, prove that the $\mathbf{k}$-algebra A is generated by the $\mathbf{k}$-submodule $\mathfrak{p}$ consisting of the primitive elements of A.

(e) Assume that A is the tensor algebra $T(V)$ of a $\mathbf{k}$-module V, and that the $\mathbf{k}$-submodule $V = V^{\otimes 1}$ of $T(V)$ is the degree-1 homogeneous component of A. Show that $(S \star E)(x_1 x_2 \ldots x_n) = [\ldots[[x_1, x_2], x_3], \ldots, x_n]$ for any $n \geq 1$ and any $x_1, x_2, \ldots, x_n \in V$.

Remark 1.5.15. Exercise 1.5.14 gives rise to a certain idempotent map $A \to A$ when $\mathbf{k}$ is a commutative $\mathbb{Q}$-algebra and A is a cocommutative connected graded $\mathbf{k}$-Hopf algebra. Namely, the $\mathbf{k}$-linear map $A \to A$ sending every homogeneous $a \in A$ to $\frac{1}{\deg a}(S \star E)(a)$ (or 0 if $\deg a = 0$) is idempotent and is a projection on the $\mathbf{k}$-module of primitive elements of A. It is called the *Dynkin idempotent*; see [168] for more of its properties.[31] Part (c) of the exercise is more or less Baker's identity.

1.6. **Duals.** Recall that for $\mathbf{k}$-modules V, taking the dual $\mathbf{k}$-module $V^* := \operatorname{Hom}(V, \mathbf{k})$ reverses $\mathbf{k}$-linear maps. That is, every $\mathbf{k}$-linear map $V \xrightarrow{\varphi} W$ induces an *adjoint map* $W^* \xrightarrow{\varphi^*} V^*$ defined uniquely by

$$(f, \varphi(v)) = (\varphi^*(f), v)$$

in which (f, v) is the bilinear pairing $V^* \times V \to \mathbf{k}$ sending $(f, v) \mapsto f(v)$. If V and W are finite free $\mathbf{k}$-modules[32], more can be said: When φ is expressed in terms of a basis $\{v_i\}_{i \in I}$ for V and a basis $\{w_j\}_{j \in J}$ for W by some matrix, the map φ^* is expressed by the transpose matrix in terms of the dual bases of these two bases[33].

The correspondence $\varphi \mapsto \varphi^*$ between $\mathbf{k}$-linear maps $V \xrightarrow{\varphi} W$ and $\mathbf{k}$-linear maps $W^* \xrightarrow{\varphi^*} V^*$ is one-to-one when W is finite free. However, this is not the case in many combinatorial situations (in which W is usually free but not finite free). Fortunately, many of the good properties of finite free modules carry over to a certain class of graded modules as long as the dual V^* is replaced by a smaller module V^o called the graded dual. Let us first introduce the latter:

When $V = \bigoplus_{n \geq 0} V_n$ is a graded $\mathbf{k}$-module, note that the dual $V^* = \prod_{n \geq 0} (V_n)^*$ can contain functionals f supported on infinitely many V_n. However, we can consider the $\mathbf{k}$-submodule $V^o := \bigoplus_{n \geq 0} (V_n)^* \subset \prod_{n \geq 0} (V_n)^* = V^*$, sometimes called the *graded dual*[34], consisting of the functions f that vanish on all but finitely many V_n. Notice that V^o is graded, whereas V^* (in general) is not. If $V \xrightarrow{\varphi} W$ is a graded $\mathbf{k}$-linear map, then the adjoint map $W^* \xrightarrow{\varphi^*} V^*$ restricts to a graded $\mathbf{k}$-linear map $W^o \to V^o$, which we (abusively) still denote by φ^*.

A graded $\mathbf{k}$-module $V = \bigoplus_{n \geq 0} V_n$ is said to be *of finite type* if each V_n is a finite free $\mathbf{k}$-module[35]. When the graded $\mathbf{k}$-module V is of finite type, the graded $\mathbf{k}$-module V^o is again of finite type[36] and satisfies $(V^o)^o \cong V$. Many other properties of finite free modules are salvaged in this situation; most importantly: The correspondence $\varphi \mapsto \varphi^*$ between graded $\mathbf{k}$-linear maps $V \to W$ and graded $\mathbf{k}$-linear maps $W^o \to V^o$ is one-to-one when W is of finite type[37].

Reversing the diagrams should then make it clear that, in the finite free or finite-type situation, duals of algebras are coalgebras, and vice-versa, and duals of bialgebras or Hopf algebras are bialgebras or Hopf

[31] We will see another such idempotent in Exercise 5.4.6.

[32] A $\mathbf{k}$-module is said to be *finite free* if it has a finite basis. If $\mathbf{k}$ is a field, then a finite free $\mathbf{k}$-module is the same as a finite-dimensional $\mathbf{k}$-vector space.

[33] If $\{v_i\}_{i \in I}$ is a basis of a finite free $\mathbf{k}$-module V, then the *dual basis* of this basis is defined as the basis $\{f_i\}_{i \in I}$ of V^* that satisfies $(f_i, v_j) = \delta_{i,j}$ for all i and j. (Recall that $\delta_{i,j}$ is the Kronecker delta: $\delta_{i,j} = 1$ if $i = j$ and 0 else.)

[34] Do not mistake this for the coalgebraic restricted dual A° of [213, §6.0].

[35] This meaning of "finite type" can differ from the standard one.

[36] More precisely: Let $V = \bigoplus_{n \geq 0} V_n$ be of finite type, and let $\{v_i\}_{i \in I}$ be a *graded basis* of V, that is, a basis of the $\mathbf{k}$-module V such that the indexing set I is partitioned into subsets $I_0, I_1, I_2, \ldots$ (which are allowed to be empty) with the property that, for every $n \in \mathbb{N}$, the subfamily $\{v_i\}_{i \in I_n}$ is a basis of the $\mathbf{k}$-module V_n. Then, we can define a family $\{f_i\}_{i \in I}$ of elements of V^o by setting $(f_i, v_j) = \delta_{i,j}$ for all $i, j \in I$. This family $\{f_i\}_{i \in I}$ is a graded basis of the graded $\mathbf{k}$-module V^o. (Actually, for every $n \in \mathbb{N}$, the subfamily $\{f_i\}_{i \in I_n}$ is a basis of the $\mathbf{k}$-submodule $(V_n)^*$ of V^o – indeed the dual basis to the basis $\{v_i\}_{i \in I_n}$ of V_n.) This basis $\{f_i\}_{i \in I}$ is said to be the *dual basis* to the basis $\{v_i\}_{i \in I}$ of V.

[37] Only W has to be of finite type here; V can be any graded $\mathbf{k}$-module.

algebras. For example, the product in a Hopf algebra A of finite type uniquely defines the coproduct of A^o via adjointness:

$$(\Delta_{A^o}(f), a \otimes b)_{A \otimes A} = (f, ab)_A.$$

Thus if A has a basis $\{a_i\}_{i \in I}$ with *product structure constants* $\{c^i_{j,k}\}$, meaning

$$a_j a_k = \sum_{i \in I} c^i_{j,k} a_i,$$

then the dual basis $\{f_i\}_{i \in I}$ has the same $\{c^i_{j,k}\}$ as its *coproduct structure constants*:

$$\Delta_{A^o}(f_i) = \sum_{(j,k) \in I \times I} c^i_{j,k} f_j \otimes f_k.$$

The assumption that A be of finite type was indispensable here; in general, the dual of a $\mathbf{k}$-algebra does not become a $\mathbf{k}$-coalgebra. However, the dual of a $\mathbf{k}$-coalgebra still becomes a $\mathbf{k}$-algebra, as shown in the following exercise:

Exercise 1.6.1. For any two $\mathbf{k}$-modules U and V, let $\rho_{U,V} : U^* \otimes V^* \to (U \otimes V)^*$ be the $\mathbf{k}$-linear map which sends every tensor $f \otimes g \in U^* \otimes V^*$ to the composition $U \otimes V \xrightarrow{f \otimes g} \mathbf{k} \otimes \mathbf{k} \xrightarrow{m_{\mathbf{k}}} \mathbf{k}$ of the map[38] $f \otimes g$ with the canonical isomorphism $\mathbf{k} \otimes \mathbf{k} \xrightarrow{m_{\mathbf{k}}} \mathbf{k}$. When $\mathbf{k}$ is a field and U is finite-dimensional, this map $\rho_{U,V}$ is a $\mathbf{k}$-vector space isomorphism (and usually regarded as the identity); more generally, it is injective whenever $\mathbf{k}$ is a field[39]. Also, let $s : \mathbf{k} \to \mathbf{k}^*$ be the canonical isomorphism. Prove that:

(a) If C is a $\mathbf{k}$-coalgebra, then C^* becomes a $\mathbf{k}$-algebra if we define its associative operation by $m_{C^*} = \Delta^*_C \circ \rho_{C,C} : C^* \otimes C^* \to C^*$ and its unit map to be $\epsilon^*_C \circ s : \mathbf{k} \to C^*$. [40]

(b) The $\mathbf{k}$-algebra structure defined on C^* in part (a) is precisely the one defined on $\operatorname{Hom}(C, \mathbf{k}) = C^*$ in Definition 1.4.1 applied to $A = \mathbf{k}$.

(c) If C is a graded $\mathbf{k}$-coalgebra, then C^o is a $\mathbf{k}$-subalgebra of the $\mathbf{k}$-algebra C^* defined in part (a).

(d) If $f : C \to D$ is a homomorphism of $\mathbf{k}$-coalgebras, then $f^* : D^* \to C^*$ is a homomorphism of $\mathbf{k}$-algebras.

(e) Let U be a graded $\mathbf{k}$-module (not necessarily of finite type), and let V be a graded $\mathbf{k}$-module of finite type. Then, there is a 1-to-1 correspondence between graded $\mathbf{k}$-linear maps $U \to V$ and graded $\mathbf{k}$-linear maps $V^o \to U^o$ given by $f \mapsto f^*$.

(f) Let C be a graded $\mathbf{k}$-coalgebra (not necessarily of finite type), and let D be a graded $\mathbf{k}$-coalgebra of finite type. Part (e) of this exercise shows that there is a 1-to-1 correspondence between graded $\mathbf{k}$-linear maps $C \to D$ and graded $\mathbf{k}$-linear maps $D^o \to C^o$ given by $f \mapsto f^*$. This correspondence has the property that a given graded $\mathbf{k}$-linear map $f : C \to D$ is a $\mathbf{k}$-coalgebra morphism if and only if $f^* : D^o \to C^o$ is a $\mathbf{k}$-algebra morphism.

Another example of a Hopf algebra is provided by the so-called shuffle algebra. Before we introduce it, let us define the *shuffles* of two words:

Definition 1.6.2. Given two words $a = (a_1, a_2, \ldots, a_n)$ and $b = (b_1, b_2, \ldots, b_m)$, the *multiset of shuffles of a and b* is defined as the multiset

$$\left\{ \left(c_{w(1)}, c_{w(2)}, \ldots, c_{w(n+m)} \right) \ : \ w \in \operatorname{Sh}_{n,m} \right\}_{\text{multiset}},$$

where $(c_1, c_2, \ldots, c_{n+m})$ is the concatenation $a \cdot b = (a_1, a_2, \ldots, a_n, b_1, b_2, \ldots, b_m)$, and where $\operatorname{Sh}_{n,m}$ is the subset[41]

$$\left\{ w \in \mathfrak{S}_{n+m} \ : \ w^{-1}(1) < w^{-1}(2) < \cdots < w^{-1}(n) ; \ w^{-1}(n+1) < w^{-1}(n+2) < \cdots < w^{-1}(n+m) \right\}$$

of the symmetric group $\mathfrak{S}_{n+m}$. Informally speaking, the shuffles of the two words a and b are the words obtained by overlaying the words a and b, after first moving their letters apart so that no letters get

[38] Keep in mind that the *tensor* $f \otimes g \in U^* \otimes V^*$ is not the same as the *map* $U \otimes V \xrightarrow{f \otimes g} \mathbf{k} \otimes \mathbf{k}$.

[39] Over arbitrary rings it does not have to be even that!

[40] If C is a finite free $\mathbf{k}$-module, then this $\mathbf{k}$-algebra structure is the same as the one defined above by adjointness. But the advantage of the new definition is that it works even if C is not a finite free $\mathbf{k}$-module.

[41] **Warning:** This definition of $\operatorname{Sh}_{n,m}$ is highly nonstandard, and many authors define $\operatorname{Sh}_{n,m}$ to be the set of the inverses of the permutations belonging to what we call $\operatorname{Sh}_{n,m}$.

superimposed when the words are overlayed[42]. In particular, any shuffle of a and b contains a and b as subsequences. The multiset of shuffles of a and b has $\binom{m+n}{n}$ elements (counted with multiplicity) and is denoted by $a \sqcup\!\sqcup\, b$. For instance, the shuffles of $(1,2,1)$ and $(3,2)$ are

$$\begin{aligned}&(\underline{1},\underline{2},\underline{1},3,2)\,,(\underline{1},\underline{2},3,\underline{1},2)\,,(\underline{1},\underline{2},3,2,\underline{1})\,,(\underline{1},3,\underline{2},\underline{1},2)\,,(\underline{1},3,\underline{2},2,\underline{1})\,,\\&(\underline{1},3,2,\underline{2},\underline{1})\,,(3,\underline{1},\underline{2},\underline{1},2)\,,(3,\underline{1},\underline{2},2,\underline{1})\,,(3,\underline{1},2,\underline{2},\underline{1})\,,(3,2,\underline{1},\underline{2},\underline{1})\,,\end{aligned}$$

listed here as often as they appear in the multiset $(1,2,1) \sqcup\!\sqcup\, (3,2)$. Here we have underlined the letters taken from a – that is, the letters at positions $w^{-1}(1)$, $w^{-1}(2)$, $\ldots$, $w^{-1}(n)$.

Example 1.6.3. When $A = T(V)$ is the tensor algebra for a finite free $\mathbf{k}$-module V, having $\mathbf{k}$-basis $\{x_i\}_{i \in I}$, its graded dual A^o is another Hopf algebra whose basis $\left\{y_{(i_1,\ldots,i_\ell)}\right\}$ (the dual basis of the basis $\{x_{i_1} \cdots x_{i_\ell}\}$ of $A = T(V)$) is indexed by words in the alphabet I. This Hopf algebra A^o could be called the *shuffle algebra* of V^*. (To be more precise, it is isomorphic to the shuffle algebra of V^* introduced in Proposition 1.6.7 further below; we prefer not to call A^o itself the shuffle algebra of V^*, since A^o has several disadvantages[43].) Duality shows that the *cut* coproduct in A^o is defined by

$$\Delta y_{(i_1,\ldots,i_\ell)} = \sum_{j=0}^{\ell} y_{(i_1,\ldots,i_j)} \otimes y_{(i_{j+1},i_{j+2},\ldots,i_\ell)}. \tag{1.6.1}$$

For example,

$$\Delta y_{abcb} = y_\varnothing \otimes y_{abcb} + y_a \otimes y_{bcb} + y_{ab} \otimes y_{cb} + y_{abc} \otimes y_b + y_{abcb} \otimes y_\varnothing .$$

Duality also shows that the *shuffle* product in A^o will be given by

$$y_{(i_1,\ldots,i_\ell)} y_{(j_1,\ldots,j_m)} = \sum_{\mathbf{k}=(k_1,\ldots,k_{\ell+m}) \in \mathbf{i} \sqcup\!\sqcup\, \mathbf{j}} y_{(k_1,\ldots,k_{\ell+m})} \tag{1.6.2}$$

where $\mathbf{i} \sqcup\!\sqcup\, \mathbf{j}$ (as in Definition 1.6.2) denotes the multiset of the $\binom{\ell+m}{\ell}$ words obtained as *shuffles* of the two words $\mathbf{i} = (i_1, \ldots, i_\ell)$ and $\mathbf{j} = (j_1, \ldots, j_m)$. For example,

$$\begin{aligned} y_{ab} y_{cb} &= y_{abcb} + y_{acbb} + y_{cabb} + y_{cabb} + y_{acbb} + y_{cbab} \\ &= y_{abcb} + 2y_{acbb} + 2y_{cabb} + y_{cbab}. \end{aligned}$$

Equivalently, one has

$$y_{(i_1,i_2,\ldots,i_\ell)} y_{(i_{\ell+1},i_{\ell+2},\ldots,i_{\ell+m})} = \sum_{\substack{w \in \mathfrak{S}_{\ell+m}: \\ w(1)<\cdots<w(\ell), \\ w(\ell+1)<\cdots<w(\ell+m)}} y_{\left(i_{w^{-1}(1)},i_{w^{-1}(2)},\ldots,i_{w^{-1}(\ell+m)}\right)} \tag{1.6.3}$$

$$= \sum_{\sigma \in \mathrm{Sh}_{\ell,m}} y_{\left(i_{\sigma(1)},i_{\sigma(2)},\ldots,i_{\sigma(\ell+m)}\right)} \tag{1.6.4}$$

(using the notations of Definition 1.6.2 again). Lastly, the antipode S of A^o is the adjoint of the antipode of $A = T(V)$ described in (1.4.6):

$$S y_{(i_1,i_2,\ldots,i_\ell)} = (-1)^\ell y_{(i_\ell,\ldots,i_2,i_1)}.$$

Since the coalgebra $T(V)$ is cocommutative, its graded dual $T(V)^o$ is commutative.

Exercise 1.6.4. Let V be a 1-dimensional free $\mathbf{k}$-module with basis element x, so $\mathrm{Sym}(V) \cong \mathbf{k}[x]$, with $\mathbf{k}$-basis $\{1 = x^0, x^1, x^2, \ldots\}$.

[42] For instance, if $a = (1,3,2,1)$ and $b = (2,4)$, then the shuffle $(1,2,3,2,4,1)$ of a and b can be obtained by moving the letters of a and b apart as follows:

$$\begin{array}{rcccccc} a = & 1 & & 3 & 2 & & 1 \\ b = & & 2 & & & 4 & \end{array}$$

and then overlaying them to obtain $\begin{array}{cccccc} 1 & 2 & 3 & 2 & 4 & 1 \end{array}$. Other ways of moving letters apart lead to further shuffles (not always distinct).

[43] Specifically, A^o has the disadvantages of being defined only when V^* is the dual of a finite free $\mathbf{k}$-module V, and depending on a choice of basis, whereas Proposition 1.6.7 will define shuffle algebras in full generality and canonically.

(a) Check that the powers x^i satisfy

$$\begin{aligned} x^i \cdot x^j &= x^{i+j}, \\ \Delta(x^n) &= \sum_{i+j=n} \binom{n}{i} x^i \otimes x^j, \\ S(x^n) &= (-1)^n x^n. \end{aligned}$$

(b) Check that the dual basis elements $\{f^{(0)}, f^{(1)}, f^{(2)}, \ldots\}$ for $\operatorname{Sym}(V)^o$, defined by $f^{(i)}(x^j) = \delta_{i,j}$, satisfy

$$\begin{aligned} f^{(i)} f^{(j)} &= \binom{i+j}{i} f^{(i+j)}, \\ \Delta(f^{(n)}) &= \sum_{i+j=n} f^{(i)} \otimes f^{(j)}, \\ S(f^{(n)}) &= (-1)^n f^{(n)}. \end{aligned}$$

(c) Show that if $\mathbb{Q}$ is a subring of $\mathbf{k}$, then the $\mathbf{k}$-linear map $\operatorname{Sym}(V)^o \to \operatorname{Sym}(V)$ sending $f^{(n)} \mapsto \frac{x^n}{n!}$ is a graded Hopf isomorphism.

For this reason, the Hopf structure on $\operatorname{Sym}(V)^o$ is called a *divided power algebra*.

(d) Show that when $\mathbf{k}$ is a field of characteristic $p > 0$, one has $(f^{(1)})^p = 0$, and hence why there can be no Hopf isomorphism $\operatorname{Sym}(V)^o \to \operatorname{Sym}(V)$.

Exercise 1.6.5. Let V have $\mathbf{k}$-basis $\{x_1, \ldots, x_n\}$, and let $V \oplus V$ have $\mathbf{k}$-basis $\{x_1, \ldots, x_n, y_1, \ldots, y_n\}$, so that one has isomorphisms

$$\operatorname{Sym}(V \oplus V) \cong \mathbf{k}[\mathbf{x}, \mathbf{y}] \cong \mathbf{k}[\mathbf{x}] \otimes \mathbf{k}[\mathbf{y}] \cong \operatorname{Sym}(V) \otimes \operatorname{Sym}(V).$$

Here we are using the abbreviations $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ and $\mathbf{y} = (y_1, y_2, \ldots, y_n)$.

(a) Show that our usual coproduct on $\operatorname{Sym}(V)$ can be re-expressed as follows:

$$\begin{tikzcd} \operatorname{Sym}(V) \arrow[d, equal] & \operatorname{Sym}(V) \otimes \operatorname{Sym}(V) \arrow[d, equal] \\ \mathbf{k}[\mathbf{x}] \arrow[r, "\Delta"] & \mathbf{k}[\mathbf{x}, \mathbf{y}], \\ f(x_1, \ldots, x_n) \arrow[r, mapsto] & f(x_1 + y_1, \ldots, x_n + y_n). \end{tikzcd}$$

In other words, it is induced from the diagonal map

$$\begin{array}{ccc} V & \longrightarrow & V \oplus V, \\ x_i & \longmapsto & x_i + y_i. \end{array} \tag{1.6.5}$$

(b) One can similarly define a coproduct on the *exterior algebra* $\wedge V$, which is the quotient $T(V)/J$ where J is the two-sided ideal generated by the elements $\{x^2 (= x \otimes x)\}_{x \in V}$ in $T^2(V)$. The ideal J is a graded $\mathbf{k}$-submodule of $T(V)$ (this is not obvious!), and the quotient $T(V)/J$ becomes a graded commutative algebra

$$\wedge V = \bigoplus_{d=0}^{n} \wedge^d V \left(= \bigoplus_{d=0}^{\infty} \wedge^d V \right),$$

if one views the elements of $V = \wedge^1 V$ as having *odd* degree, and uses the topologist's sign convention (as in (1.3.3)). One again has $\wedge(V \oplus V) = \wedge V \otimes \wedge V$ as graded algebras. Show that one can again let the diagonal map (1.6.5) induce a map

$$\begin{tikzcd} \wedge(V) \arrow[r, "\Delta"] & \wedge V \otimes \wedge V, \\ f(x_1, \ldots, x_n) \arrow[r, mapsto] \arrow[d, equal] & f(x_1 + y_1, \ldots, x_n + y_n) \arrow[d, equal] \\ \sum c_{i_1, \ldots, i_d} \cdot x_{i_1} \wedge \cdots \wedge x_{i_d} & \sum c_{i_1, \ldots, i_d} \cdot (x_{i_1} + y_{i_1}) \wedge \cdots \wedge (x_{i_d} + y_{i_d}), \end{tikzcd} \tag{1.6.6}$$

which makes $\wedge V$ into a connected graded Hopf algebra.

(c) Show that in the tensor algebra $T(V)$, if one views the elements of $V = V^{\otimes 1}$ as having odd degree, and uses the topologist's sign convention (1.3.3) in the twist map when defining $T(V)$, then for any x in V one has $\Delta(x^2) = 1 \otimes x^2 + x^2 \otimes 1$.

(d) Let us use the convention (1.3.3) as in part (c). Show that the two-sided ideal $J \subset T(V)$ generated by $\{x^2\}_{x \in V}$ is also a two-sided coideal and a graded $\mathbf{k}$-submodule of $T(V)$, and hence the quotient $\wedge V = T(V)/J$ inherits the structure of a graded bialgebra. Check that the coproduct on $\wedge V$ inherited from $T(V)$ is the same as the one defined in part (b).

[**Hint:** The ideal J in part (b) is a graded $\mathbf{k}$-submodule of $T(V)$, but this is not completely obvious (not all elements of V have to be homogeneous!).]

Exercise 1.6.6. Let C be a $\mathbf{k}$-coalgebra. As we know from Exercise 1.6.1(a), this makes C^* into a $\mathbf{k}$-algebra.
Let A be a $\mathbf{k}$-algebra which is finite free as $\mathbf{k}$-module. This makes A^* into a $\mathbf{k}$-coalgebra.
Let $f : C \to A$ and $g : C \to A$ be two $\mathbf{k}$-linear maps. Show that $f^* \star g^* = (f \star g)^*$.

The above arguments might have created the impression that duals of bialgebras have good properties only under certain restrictive conditions (e.g., the dual of a bialgebra H does not generally become a bialgebra unless H is of finite type), and so they cannot be used in proofs and constructions unless one is willing to sacrifice some generality (e.g., we had to require V to be finite free in Example 1.6.3). While the first part of this impression is true, the second is not always; often there is a way to gain back the generality lost from using duals. As an example of this, let us define the shuffle algebra of an arbitrary $\mathbf{k}$-module (not just of a dual of a finite free $\mathbf{k}$-module as in Example 1.6.3):

Proposition 1.6.7. *Let V be a $\mathbf{k}$-module. Define a $\mathbf{k}$-linear map $\Delta_{\sqcup\!\sqcup} : T(V) \to T(V) \otimes T(V)$ by setting*

$$\Delta_{\sqcup\!\sqcup}(v_1 v_2 \cdots v_n) = \sum_{k=0}^{n} (v_1 v_2 \cdots v_k) \otimes (v_{k+1} v_{k+2} \cdots v_n) \qquad \text{for all } n \in \mathbb{N} \text{ and } v_1, v_2, \ldots, v_n \in V.$$

[44] *Define a $\mathbf{k}$-bilinear map $\underline{\sqcup\!\sqcup} : T(V) \times T(V) \to T(V)$, which will be written in infix notation (that is, we will write $a \,\underline{\sqcup\!\sqcup}\, b$ instead of $\underline{\sqcup\!\sqcup}(a, b)$), by setting*[45]

$$(v_1 v_2 \cdots v_\ell) \,\underline{\sqcup\!\sqcup}\, (v_{\ell+1} v_{\ell+2} \cdots v_{\ell+m}) = \sum_{\sigma \in \mathrm{Sh}_{\ell,m}} v_{\sigma(1)} v_{\sigma(2)} \cdots v_{\sigma(\ell+m)}$$
$$\text{for all } \ell, m \in \mathbb{N} \text{ and } v_1, v_2, \ldots, v_{\ell+m} \in V.$$

[46] *Consider also the comultiplication ϵ of the Hopf algebra $T(V)$.*

Then, the $\mathbf{k}$-module $T(V)$, endowed with the multiplication $\underline{\sqcup\!\sqcup}$, the unit $1_{T(V)} \in V^{\otimes 0} \subset T(V)$, the comultiplication $\Delta_{\sqcup\!\sqcup}$ and the counit ϵ, becomes a commutative Hopf algebra. This Hopf algebra is called the shuffle algebra *of V, and denoted by* $\mathrm{Sh}(V)$. *The antipode of the Hopf algebra* $\mathrm{Sh}(V)$ *is precisely the antipode S of $T(V)$.*

Exercise 1.6.8. Prove Proposition 1.6.7.

[**Hint:** When V is a finite free $\mathbf{k}$-module, Proposition 1.6.7 follows from Example 1.6.3. The trick is to derive the general case from this specific one. Every $\mathbf{k}$-linear map $f : W \to V$ between two $\mathbf{k}$-modules W and V induces a map $T(f) : T(W) \to T(V)$ which preserves $\Delta_{\sqcup\!\sqcup}$, $\underline{\sqcup\!\sqcup}$, $1_{T(W)}$, ϵ and S (in the appropriate meanings – e.g., preserving $\Delta_{\sqcup\!\sqcup}$ means $\Delta_{\sqcup\!\sqcup} \circ T(f) = (T(f) \otimes T(f)) \circ \Delta_{\sqcup\!\sqcup}$). Show that each of the equalities that need to be proven in order to verify Proposition 1.6.7 can be "transported" along such a map $T(f)$ from a $T(W)$ for a suitably chosen finite free $\mathbf{k}$-module W.]

It is also possible to prove Proposition 1.6.7 "by foot", as long as one is ready to make combinatorial arguments about cutting shuffles.

Remark 1.6.9. (a) Let V be a finite free $\mathbf{k}$-module. The Hopf algebra $T(V)^o$ (studied in Example 1.6.3) is naturally isomorphic to the shuffle algebra $\mathrm{Sh}(V^*)$ (defined as in Proposition 1.6.7 but for V^* instead of V) as Hopf algebras, by the obvious isomorphism (namely, the direct sum of the isomorphisms $(V^{\otimes n})^* \to (V^*)^{\otimes n}$ over all $n \in \mathbb{N}$). [47]

[44] This is well-defined, because the right hand side is n-multilinear in $v_1, v_2, \ldots, v_n$, and because any n-multilinear map $V^{\times n} \to M$ into a $\mathbf{k}$-module M gives rise to a unique $\mathbf{k}$-linear map $V^{\otimes n} \to M$.

[45] Many authors use the symbol $\sqcup\!\sqcup$ instead of $\underline{\sqcup\!\sqcup}$ here, but we prefer to reserve the former notation for the shuffle product of words.

[46] Again, this is well-defined by the $\ell + m$-multilinearity of the right hand side.

[47] This can be verified by comparing (1.6.1) with the definition of $\Delta_{\sqcup\!\sqcup}$, and comparing (1.6.4) with the definition of $\underline{\sqcup\!\sqcup}$.

(b) The same statement applies to the case when V is a graded $\mathbf{k}$-module of finite type satisfying $V_0 = 0$ rather than a finite free $\mathbf{k}$-module, provided that V^* and $\left(V^{\otimes n}\right)^*$ are replaced by V^o and $\left(V^{\otimes n}\right)^o$.

We shall return to shuffle algebras in Section 6.3, where we will show that under certain conditions ($\mathbb{Q}$ being a subring of $\mathbf{k}$, and V being a free $\mathbf{k}$-module) the algebra structure on a shuffle algebra $\operatorname{Sh}(V)$ is a polynomial algebra in an appropriately chosen set of generators[48].

1.7. **Infinite sums and Leray's theorem.** In this section (which can be skipped, as it will not be used except in a few exercises), we will see how a Hopf algebra structure on a $\mathbf{k}$-algebra reveals knowledge about the $\mathbf{k}$-algebra itself. Specifically, we will show that if $\mathbf{k}$ is a commutative $\mathbb{Q}$-algebra, and if A is any commutative connected graded $\mathbf{k}$-Hopf algebra, then A as a $\mathbf{k}$-algebra must be (isomorphic to) a symmetric algebra of a $\mathbf{k}$-module[49]. This is a specimen of a class of facts which are commonly called *Leray theorems*; for different specimens, see [156, Theorem 7.5] or [35, p. 17, "Hopf's theorem"] or [35, §2.5, A, B, C] or [35, Theorem 3.8.3].[50] In a sense, these facts foreshadow Zelevinsky's theory of positive self-dual Hopf algebras, which we shall encounter in Chapter 3; however, the latter theory works in a much less general setting (and makes much stronger claims).

We shall first explore the possibilities of applying a formal power series v to a linear map $f : C \to A$ from a coalgebra C to an algebra A. We have already seen an example of this in the proof of Proposition 1.4.7 above (where the power series $\sum_{k\geq 0} (-1)^k T^k \in \mathbf{k}[[T]]$ was applied to the locally $\star$-nilpotent map $\operatorname{id}_A - u_A \epsilon_A : A \to A$); we shall now take a more systematic approach and establish general criteria for when such applications are possible. First, we will have to make sense of infinite sums of maps from a coalgebra to an algebra. This is somewhat technical, but the effort will pay off.

Definition 1.7.1. Let A be an abelian group (written additively).

We say that a family $\left(a_q\right)_{q\in Q} \in A^Q$ of elements of A is *finitely supported* if all but finitely many $q \in Q$ satisfy $a_q = 0$. Clearly, if $\left(a_q\right)_{q\in Q} \in A^Q$ is a finitely supported family, then the sum $\sum_{q\in Q} a_q$ is well-defined (since all but finitely many of its addends are 0). Sums like this satisfy the usual rules for sums, even though their indexing set Q may be infinite. (For example, if $\left(a_q\right)_{q\in Q}$ and $\left(b_q\right)_{q\in Q}$ are two finitely supported families in A^Q, then the family $\left(a_q + b_q\right)_{q\in Q}$ is also finitely supported, and we have $\sum_{q\in Q} a_q + \sum_{q\in Q} b_q = \sum_{q\in Q} \left(a_q + b_q\right)$.)

Definition 1.7.2. Let C and A be two $\mathbf{k}$-modules.

We say that a family $\left(f_q\right)_{q\in Q} \in \left(\operatorname{Hom}(C, A)\right)^Q$ of maps $f_q \in \operatorname{Hom}(C, A)$ is *pointwise finitely supported* if for each $x \in C$, the family $\left(f_q(x)\right)_{q\in Q} \in A^Q$ of elements of A is finitely supported.[51] If $\left(f_q\right)_{q\in Q} \in \left(\operatorname{Hom}(C, A)\right)^Q$ is a pointwise finitely supported family, then the sum $\sum_{q\in Q} f_q$ is defined to be the map $C \to A$ sending each $x \in C$ to $\sum_{q\in Q} f_q(x)$. [52]

Note that the concept of a "pointwise finitely supported" family $\left(f_q\right)_{q\in Q} \in \left(\operatorname{Hom}(C, A)\right)^Q$ is precisely the concept of a "summable" family in [60, Definition 1].

Definition 1.7.3. For the rest of Section 1.7, we shall use the following conventions:

- Let C be a $\mathbf{k}$-coalgebra. Let A be a $\mathbf{k}$-algebra.

[48] This says nothing about the coalgebra structure on $\operatorname{Sh}(V)$ – which is much more complicated in these generators.

[49] If $\mathbf{k}$ is a field, then this simply means that A as a $\mathbf{k}$-algebra must be a polynomial ring over $\mathbf{k}$.

[50] Notice that many of these sources assume $\mathbf{k}$ to be a field; some of their proofs rely on this assumption.

[51] Here are some examples of pointwise finitely supported families:

- If Q is a finite set, then any family $\left(f_q\right)_{q\in Q} \in \left(\operatorname{Hom}(C, A)\right)^Q$ is pointwise finitely supported.
- More generally, any finitely supported family $\left(f_q\right)_{q\in Q} \in \left(\operatorname{Hom}(C, A)\right)^Q$ is pointwise finitely supported.
- If C is a graded $\mathbf{k}$-module, and if $\left(f_n\right)_{n\in\mathbb{N}} \in \left(\operatorname{Hom}(C, A)\right)^{\mathbb{N}}$ is a family of maps such that $f_n(C_m) = 0$ whenever $n \neq m$, then the family $\left(f_n\right)_{n\in\mathbb{N}}$ is pointwise finitely supported.
- If C is a graded $\mathbf{k}$-coalgebra and A is any $\mathbf{k}$-algebra, and if $f \in \operatorname{Hom}(C, A)$ satisfies $f(C_0) = 0$, then the family $\left(f^{\star n}\right)_{n\in\mathbb{N}} \in \left(\operatorname{Hom}(C, A)\right)^{\mathbb{N}}$ is pointwise finitely supported. (This will be proven in Proposition 1.7.11(h).)

[52] This definition of $\sum_{q\in Q} f_q$ generalizes the usual definition of $\sum_{q\in Q} f_q$ when Q is a finite set (because if Q is a finite set, then any family $\left(f_q\right)_{q\in Q} \in \left(\operatorname{Hom}(C, A)\right)^Q$ is pointwise finitely supported).

- We shall avoid our standard practice of denoting the unit map $u_A : \mathbf{k} \to A$ of a $\mathbf{k}$-algebra A by u; instead, we will use the letter u (without the subscript A) for other purposes.

Definition 1.7.2 allows us to work with infinite sums in $\operatorname{Hom}(C, A)$, provided that we are summing a pointwise finitely supported family. We shall next state some properties of such sums:[53]

Proposition 1.7.4. *Let $(f_q)_{q\in Q} \in (\operatorname{Hom}(C, A))^Q$ be a pointwise finitely supported family. Then, the map $\sum_{q\in Q} f_q$ belongs to $\operatorname{Hom}(C, A)$.*

Proposition 1.7.5. *Let $(f_q)_{q\in Q}$ and $(g_q)_{q\in Q}$ be two pointwise finitely supported families in $(\operatorname{Hom}(C, A))^Q$. Then, the family $(f_q + g_q)_{q\in Q} \in (\operatorname{Hom}(C, A))^Q$ is also pointwise finitely supported, and satisfies*

$$\sum_{q\in Q} f_q + \sum_{q\in Q} g_q = \sum_{q\in Q} (f_q + g_q) .$$

Proposition 1.7.6. *Let $(f_q)_{q\in Q} \in (\operatorname{Hom}(C, A))^Q$ and $(g_r)_{r\in R} \in (\operatorname{Hom}(C, A))^R$ be two pointwise finitely supported families. Then, the family $(f_q \star g_r)_{(q,r)\in Q\times R} \in (\operatorname{Hom}(C, A))^{Q\times R}$ is pointwise finitely supported, and satisfies*

$$\sum_{(q,r)\in Q\times R} (f_q \star g_r) = \left(\sum_{q\in Q} f_q\right) \star \left(\sum_{r\in R} g_r\right) .$$

Roughly speaking, the above three propositions say that sums of the form $\sum_{q\in Q} f_q$ (where $(f_q)_{q\in Q}$ is a pointwise finitely supported family) satisfy the usual rules for finite sums. Furthermore, the following properties of pointwise finitely supported families hold:

Proposition 1.7.7. *Let $(f_q)_{q\in Q} \in (\operatorname{Hom}(C, A))^Q$ be a pointwise finitely supported family. Let $(\lambda_q)_{q\in Q} \in \mathbf{k}^Q$ be any family of elements of $\mathbf{k}$. Then, the family $(\lambda_q f_q)_{q\in Q} \in (\operatorname{Hom}(C, A))^Q$ is pointwise finitely supported.*

Proposition 1.7.8. *Let $(f_q)_{q\in Q} \in (\operatorname{Hom}(C, A))^Q$ and $(g_q)_{q\in Q} \in (\operatorname{Hom}(C, A))^Q$ be two families such that $(f_q)_{q\in Q}$ is pointwise finitely supported. Then, the family $(f_q \star g_q)_{q\in Q} \in (\operatorname{Hom}(C, A))^Q$ is also pointwise finitely supported.*

Exercise 1.7.9. Prove Propositions 1.7.4, 1.7.5, 1.7.6, 1.7.7 and 1.7.8.

We can now define the notion of a "pointwise $\star$-nilpotent" map. Roughly speaking, this will mean an element of $(\operatorname{Hom}(C, A), \star)$ that can be substituted into any power series because its powers (with respect to the convolution $\star$) form a pointwise finitely supported family. Here is the definition:

Definition 1.7.10.
(a) A map $f \in \operatorname{Hom}(C, A)$ is said to be *pointwise $\star$-nilpotent* if and only if the family $(f^{\star n})_{n\in\mathbb{N}} \in (\operatorname{Hom}(C, A))^{\mathbb{N}}$ is pointwise finitely supported. Equivalently, a map $f \in \operatorname{Hom}(C, A)$ is pointwise $\star$-nilpotent if and only if for each $x \in C$, the family $(f^{\star n}(x))_{n\in\mathbb{N}}$ of elements of A is finitely supported.
(b) If $f \in \operatorname{Hom}(C, A)$ is a pointwise $\star$-nilpotent map, and if $(\lambda_n)_{n\in\mathbb{N}} \in \mathbf{k}^{\mathbb{N}}$ is any family of scalars, then the family $(\lambda_n f^{\star n})_{n\in\mathbb{N}} \in (\operatorname{Hom}(C, A))^{\mathbb{N}}$ is pointwise finitely supported[54], and thus the infinite sum $\sum_{n\geq 0} \lambda_n f^{\star n} = \sum_{n\in\mathbb{N}} \lambda_n f^{\star n}$ is well-defined and belongs to $\operatorname{Hom}(C, A)$ (by Proposition 1.7.4).[55]

[53] See Exercise 1.7.9 below for the proofs of these properties.

[54] This follows easily from Proposition 1.7.7 above. (In fact, the map f is pointwise $\star$-nilpotent, and thus the family $(f^{\star n})_{n\in\mathbb{N}} \in (\operatorname{Hom}(C, A))^{\mathbb{N}}$ is pointwise finitely supported (by the definition of "pointwise $\star$-nilpotent"). Hence, Proposition 1.7.7 (applied to $Q = \mathbb{N}$ and $(f_q)_{q\in Q} = (f^{\star n})_{n\in\mathbb{N}}$ and $(\lambda_q)_{q\in Q} = (\lambda_n)_{n\in\mathbb{N}}$) shows that the family $(\lambda_n f^{\star n})_{n\in\mathbb{N}} \in (\operatorname{Hom}(C, A))^{\mathbb{N}}$ is pointwise finitely supported.)

[55] Notice that the concept of "local $\star$-nilpotence" we used in the proof of Proposition 1.4.24 serves the same function (viz., ensuring that the sum $\sum_{n\in\mathbb{N}} \lambda_n f^{\star n}$ is well-defined). But local $\star$-nilpotence is only defined when a grading is present, whereas pointwise $\star$-nilpotence is defined in the general case. Also, local $\star$-nilpotence is more restrictive (i.e., a locally $\star$-nilpotent map is always pointwise $\star$-nilpotent, but the converse does not always hold).

(c) We let $\mathfrak{n}(C,A)$ be the set of all pointwise $\star$-nilpotent maps $f \in \operatorname{Hom}(C,A)$. Note that this is not necessarily a $\mathbf{k}$-submodule of $\operatorname{Hom}(C,A)$.

(d) Consider the ring $\mathbf{k}[[T]]$ of formal power series in an indeterminate T over $\mathbf{k}$. For any power series $u \in \mathbf{k}[[T]]$ and any $f \in \mathfrak{n}(C,A)$, we define a map $u^\star(f) \in \operatorname{Hom}(C,A)$ by $u^\star(f) = \sum_{n\geq 0} u_n f^{\star n}$, where u is written in the form $u = \sum_{n\geq 0} u_n T^n$ with $(u_n)_{n\geq 0} \in \mathbf{k}^{\mathbb{N}}$. (This sum $\sum_{n\geq 0} u_n f^{\star n}$ is well-defined in $\operatorname{Hom}(C,A)$, since f is pointwise $\star$-nilpotent.)

The following proposition gathers some properties of pointwise $\star$-nilpotent maps[56]:

Proposition 1.7.11. (a) *For any $f \in \mathfrak{n}(C,A)$ and $k \in \mathbb{N}$, we have*

$$\left(T^k\right)^\star(f) = f^{\star k}. \tag{1.7.1}$$

(b) *For any $f \in \mathfrak{n}(C,A)$ and $u, v \in \mathbf{k}[[T]]$, we have*

$$(u+v)^\star(f) = u^\star(f) + v^\star(f) \qquad \text{and} \tag{1.7.2}$$

$$(uv)^\star(f) = u^\star(f) \star v^\star(f). \tag{1.7.3}$$

Also, for any $f \in \mathfrak{n}(C,A)$ and $u \in \mathbf{k}[[T]]$ and $\lambda \in \mathbf{k}$, we have

$$(\lambda u)^\star(f) = \lambda u^\star(f). \tag{1.7.4}$$

Also, for any $f \in \mathfrak{n}(C,A)$, we have

$$0^\star(f) = 0 \qquad \text{and} \tag{1.7.5}$$

$$1^\star(f) = u_A \epsilon_C. \tag{1.7.6}$$

(c) *If $f, g \in \mathfrak{n}(C,A)$ satisfy $f \star g = g \star f$, then $f + g \in \mathfrak{n}(C,A)$.*

(d) *For any $\lambda \in \mathbf{k}$ and $f \in \mathfrak{n}(C,A)$, we have $\lambda f \in \mathfrak{n}(C,A)$.*

(e) *If $f \in \mathfrak{n}(C,A)$ and $g \in \operatorname{Hom}(C,A)$ satisfy $f \star g = g \star f$, then $f \star g \in \mathfrak{n}(C,A)$.*

(f) *If $v \in \mathbf{k}[[T]]$ is a power series whose constant term is 0, then $v^\star(f) \in \mathfrak{n}(C,A)$ for each $f \in \mathfrak{n}(C,A)$.*

(g) *If $u, v \in \mathbf{k}[[T]]$ are two power series such that the constant term of v is 0, and if $f \in \mathfrak{n}(C,A)$ is arbitrary, then*

$$(u[v])^\star(f) = u^\star(v^\star(f)). \tag{1.7.7}$$

Here, $u[v]$ denotes the composition *of u with v; this is the power series obtained by substituting v for T in u. (This power series is well-defined, since v has constant term 0.) Furthermore, notice that the right hand side of (1.7.7) is well-defined, since Proposition 1.7.11(f) shows that $v^\star(f) \in \mathfrak{n}(C,A)$.*

(h) *If C is a graded $\mathbf{k}$-coalgebra, and if $f \in \operatorname{Hom}(C,A)$ satisfies $f(C_0) = 0$, then $f \in \mathfrak{n}(C,A)$.*

(i) *If B is any $\mathbf{k}$-algebra, and if $s : A \to B$ is any $\mathbf{k}$-algebra homomorphism, then every $u \in \mathbf{k}[[T]]$ and $f \in \mathfrak{n}(C,A)$ satisfy*

$$s \circ f \in \mathfrak{n}(C,B) \qquad \text{and} \qquad u^\star(s \circ f) = s \circ (u^\star(f)).$$

(j) *If C is a connected graded $\mathbf{k}$-bialgebra, and if $F : C \to A$ is a $\mathbf{k}$-algebra homomorphism, then $F - u_A \epsilon_C \in \mathfrak{n}(C,A)$.*

Example 1.7.12. Let C be a graded $\mathbf{k}$-coalgebra. Let $f \in \operatorname{Hom}(C,A)$ be such that $f(C_0) = 0$. Then, we claim that the map $u_A\epsilon_C + f : C \to A$ is $\star$-invertible. (This observation has already been made in the proof of Proposition 1.4.24, at least in the particular case when $C = A$.)

Let us see how this claim follows from Proposition 1.7.11. First, Proposition 1.7.11(h) shows that $f \in \mathfrak{n}(C,A)$. Now, define a power series $u \in \mathbf{k}[[T]]$ by $u = 1 + T$. Then, the power series u has constant term 1, and thus has a multiplicative inverse $v = u^{-1} \in \mathbf{k}[[T]]$. Consider this v. (Explicitly, $v = \sum_{n\geq 0} (-1)^n T^n$, but this does not matter for us.) Now, (1.7.3) yields $(uv)^\star(f) = u^\star(f) \star v^\star(f)$. Since $uv = 1$ (because $v = u^{-1}$), we have $(uv)^\star(f) = 1^\star(f) = u_A\epsilon_C$ (by (1.7.6)). Thus, $u^\star(f) \star v^\star(f) = (uv)^\star(f) = u_A\epsilon_C$. Hence, the map $u^\star(f)$ has a right $\star$-inverse.

[56] See Exercise 1.7.13 below for the proofs of these properties.

Also, from $u = 1 + T$, we obtain

$$\begin{aligned} u^{\star}(f) &= (1+T)^{\star}(f) = \underbrace{1^{\star}(f)}_{=u_A\epsilon_C} + \underbrace{T^{\star}(f)}_{\substack{=f^{\star 1} \\ \text{(by (1.7.1), applied to } k=1\text{)}}} \qquad \text{(by (1.7.2))} \\ &= u_A\epsilon_C + \underbrace{f^{\star 1}}_{=f} = u_A\epsilon_C + f. \end{aligned}$$

Thus, the map $u_A\epsilon_C + f$ has a right $\star$-inverse (since the map $u^{\star}(f)$ has a right $\star$-inverse). A similar argument shows that this map $u_A\epsilon_C + f$ has a left $\star$-inverse. Consequently, the map $u_A\epsilon_C + f$ is $\star$-invertible.

Exercise 1.7.13. Prove Proposition 1.7.11.

Definition 1.7.14. (a) For the rest of Section 1.7, we assume that $\mathbf{k}$ is a commutative $\mathbb{Q}$-algebra. Thus, the two formal power series $\exp = \sum_{n\geq 0} \dfrac{1}{n!} T^n \in \mathbf{k}[[T]]$ and $\log(1+T) = \sum_{n\geq 1} \dfrac{(-1)^{n-1}}{n} T^n \in \mathbf{k}[[T]]$ are well-defined.

(b) Define two power series $\overline{\exp} \in \mathbf{k}[[T]]$ and $\overline{\log} \in \mathbf{k}[[T]]$ by $\overline{\exp} = \exp - 1$ and $\overline{\log} = \log(1+T)$.

(c) If u and v are two power series in $\mathbf{k}[[T]]$ such that v has constant term 0, then $u[v]$ denotes the *composition* of u with v; this is the power series obtained by substituting v for T in u.

The following proposition is just a formal analogue of the well-known fact that the exponential function and the logarithm are mutually inverse (on their domains of definition):[57]

Proposition 1.7.15. *Both power series $\overline{\exp}$ and $\overline{\log}$ have constant term 0 and satisfy $\overline{\exp}\left[\overline{\log}\right] = T$ and $\overline{\log}\left[\overline{\exp}\right] = T$.*

For any map $f \in \mathfrak{n}(C, A)$, the power series $\exp$, $\overline{\exp}$ and $\overline{\log}$ give rise to three further maps $\exp^{\star} f$, $\overline{\exp}^{\star} f$ and $\overline{\log}^{\star} f$. We can also define a map $\log^{\star} g$ whenever g is a map in $\operatorname{Hom}(C, A)$ satisfying $g - u_A\epsilon_C \in \mathfrak{n}(C, A)$ (but we cannot define $\log^{\star} f$ for $f \in \mathfrak{n}(C, A)$, since $\log$ is not per se a power series); in order to do this, we need a simple lemma:

Lemma 1.7.16. *Let $g \in \operatorname{Hom}(C, A)$ be such that $g - u_A\epsilon_C \in \mathfrak{n}(C, A)$. Then, $\overline{\log}^{\star}(g - u_A\epsilon_C)$ is a well-defined element of $\mathfrak{n}(C, A)$.*

Definition 1.7.17. If $g \in \operatorname{Hom}(C, A)$ is a map satisfying $g - u_A\epsilon_C \in \mathfrak{n}(C, A)$, then we define a map $\log^{\star} g \in \mathfrak{n}(C, A)$ by $\log^{\star} g = \overline{\log}^{\star}(g - u_A\epsilon_C)$. (This is well-defined, according to Lemma 1.7.16.)

Proposition 1.7.18. (a) *Each $f \in \mathfrak{n}(C, A)$ satisfies $\exp^{\star} f - u_A\epsilon_C \in \mathfrak{n}(C, A)$ and*

$$\log^{\star}(\exp^{\star} f) = f.$$

(b) *Each $g \in \operatorname{Hom}(C, A)$ satisfying $g - u_A\epsilon_C \in \mathfrak{n}(C, A)$ satisfies*

$$\exp^{\star}(\log^{\star} g) = g.$$

(c) *If $f, g \in \mathfrak{n}(C, A)$ satisfy $f \star g = g \star f$, then $f + g \in \mathfrak{n}(C, A)$ and $\exp^{\star}(f+g) = (\exp^{\star} f) \star (\exp^{\star} g)$.*

(d) *The $\mathbf{k}$-linear map $0 : C \to A$ satisfies $0 \in \mathfrak{n}(C, A)$ and $\exp^{\star} 0 = u_A\epsilon_C$.*

(e) *If $f \in \mathfrak{n}(C, A)$ and $n \in \mathbb{N}$, then $nf \in \mathfrak{n}(C, A)$ and $\exp^{\star}(nf) = (\exp^{\star} f)^{\star n}$.*

(f) *If $f \in \mathfrak{n}(C, A)$, then*

$$\log^{\star}(f + u_A\epsilon_C) = \sum_{n\geq 1} \frac{(-1)^{n-1}}{n} f^{\star n}. \tag{1.7.8}$$

Example 1.7.19. Consider again the Hopf algebra $\mathbf{k}[x]$ from Exercise 1.6.4. Let $c_1 : \mathbf{k}[x] \to \mathbf{k}$ be the $\mathbf{k}$-linear map sending each polynomial $p \in \mathbf{k}[x]$ to the coefficient of x^1 in p. (In other words, c_1 sends each polynomial $p \in \mathbf{k}[x]$ to its derivative at 0.)

Then, $c_1\left((\mathbf{k}[x])_0\right) = 0$ (as can easily be seen). Hence, Proposition 1.7.11(h) shows that $c_1 \in \mathfrak{n}(\mathbf{k}[x], \mathbf{k})$. Thus, a map $\exp^{\star}(c_1) : \mathbf{k}[x] \to \mathbf{k}$ is well-defined. It is not hard to see that this map is explicitly given by

$$(\exp^{\star}(c_1))(p) = p(1) \qquad \text{for every } p \in \mathbf{k}[x].$$

[57] See Exercise 1.7.20 below for the proof of this proposition, as well as of the lemma and proposition that follow afterwards.

(In fact, this follows easily after showing that each $n \in \mathbb{N}$ satisfies

$$(c_1)^{\star n}(p) = n! \cdot (\text{the coefficient of } x^n \text{ in } p) \qquad \text{for every } p \in \mathbf{k}[x],$$

which in turn is easily seen by induction.)

Note that the equality $(\exp^{\star}(c_1))(p) = p(1)$ shows that the map $\exp^{\star}(c_1)$ is a $\mathbf{k}$-algebra homomorphism. This is a particular case of a fact that we will soon see (Proposition 1.7.23).

Exercise 1.7.20. Prove Proposition 1.7.15, Lemma 1.7.16 and Proposition 1.7.18.

Next, we state another sequence of facts (some of which have nothing to do with Hopf algebras), beginning with a fact about convolutions which is similar to Proposition 1.4.3:[58]

Proposition 1.7.21. *Let C and C' be two $\mathbf{k}$-coalgebras, and let A and A' be two $\mathbf{k}$-algebras. Let $\gamma : C \to C'$ be a $\mathbf{k}$-coalgebra morphism. Let $\alpha : A \to A'$ be a $\mathbf{k}$-algebra morphism.*

(a) *If $f \in \operatorname{Hom}(C, A)$, $g \in \operatorname{Hom}(C, A)$, $f' \in \operatorname{Hom}(C', A')$ and $g' \in \operatorname{Hom}(C', A')$ satisfy $f' \circ \gamma = \alpha \circ f$ and $g' \circ \gamma = \alpha \circ g$, then $(f' \star g') \circ \gamma = \alpha \circ (f \star g)$.*

(b) *If $f \in \operatorname{Hom}(C, A)$ and $f' \in \operatorname{Hom}(C', A')$ satisfy $f' \circ \gamma = \alpha \circ f$, then each $n \in \mathbb{N}$ satisfies $(f')^{\star n} \circ \gamma = \alpha \circ f^{\star n}$.*

Proposition 1.7.22. *Let C be a $\mathbf{k}$-bialgebra. Let A be a commutative $\mathbf{k}$-algebra. Let $f \in \operatorname{Hom}(C, A)$ be such that $f\left((\ker \epsilon)^2\right) = 0$ and $f(1) = 0$. Then, any $x, y \in C$ and $n \in \mathbb{N}$ satisfy*

$$f^{\star n}(xy) = \sum_{i=0}^{n} \binom{n}{i} f^{\star i}(x) f^{\star(n-i)}(y).$$

Proposition 1.7.23. *Let C be a $\mathbf{k}$-bialgebra. Let A be a commutative $\mathbf{k}$-algebra. Let $f \in \mathfrak{n}(C, A)$ be such that $f\left((\ker \epsilon)^2\right) = 0$ and $f(1) = 0$. Then, $\exp^{\star} f : C \to A$ is a $\mathbf{k}$-algebra homomorphism.*

Lemma 1.7.24. *Let V be any torsionfree abelian group (written additively). Let $N \in \mathbb{N}$. For every $k \in \{0, 1, \ldots, N\}$, let w_k be an element of V. Assume that*

$$\sum_{k=0}^{N} w_k n^k = 0 \qquad \text{for all } n \in \mathbb{N}. \tag{1.7.9}$$

Then, $w_k = 0$ for every $k \in \{0, 1, \ldots, N\}$.

Lemma 1.7.25. *Let V be a torsionfree abelian group (written additively). Let $(w_k)_{k \in \mathbb{N}} \in V^{\mathbb{N}}$ be a finitely supported family of elements of V. Assume that*

$$\sum_{k \in \mathbb{N}} w_k n^k = 0 \qquad \text{for all } n \in \mathbb{N}.$$

Then, $w_k = 0$ for every $k \in \mathbb{N}$.

Proposition 1.7.26. *Let C be a graded $\mathbf{k}$-bialgebra. Let A be a commutative $\mathbf{k}$-algebra. Let $f \in \operatorname{Hom}(C, A)$ be such that $f(C_0) = 0$. Assume that*[59] *$\exp^{\star} f : C \to A$ is a $\mathbf{k}$-algebra homomorphism. Then, $f\left((\ker \epsilon)^2\right) = 0$.*

Proposition 1.7.27. *Let C be a connected graded $\mathbf{k}$-bialgebra. Let A be a commutative $\mathbf{k}$-algebra. Let $f \in \mathfrak{n}(C, A)$ be such that $f\left((\ker \epsilon)^2\right) = 0$ and $f(1) = 0$. Assume further that $f(C)$ generates the $\mathbf{k}$-algebra A. Then, $\exp^{\star} f : C \to A$ is a surjective $\mathbf{k}$-algebra homomorphism.*

Exercise 1.7.28. Prove Lemmas 1.7.24 and 1.7.25 and Propositions 1.7.21, 1.7.22, 1.7.23, 1.7.26 and 1.7.27.

[**Hint:** For Proposition 1.7.26, show first that $\exp^{\star}(nf) = (\exp^{\star} f)^{\star n}$ is a $\mathbf{k}$-algebra homomorphism for each $n \in \mathbb{N}$. Turn this into an equality between polynomials in n, and use Lemma 1.7.25.]

With these preparations, we can state our version of Leray's theorem:

[58] See Exercise 1.7.28 below for their proofs.

[59] Notice that $\exp^{\star} f$ is well-defined, since Proposition 1.7.11(h) yields $f \in \mathfrak{n}(C, A)$.

Theorem 1.7.29. *Let A be a commutative connected graded $\mathbf{k}$-bialgebra.*[60]

(a) *We have $\operatorname{id}_A - u_A\epsilon_A \in \mathfrak{n}(A, A)$; thus, the map $\log^\star(\operatorname{id}_A) \in \mathfrak{n}(A, A)$ is well-defined. We denote this map $\log^\star(\operatorname{id}_A)$ by $\mathfrak{e}$.*

(b) *We have $\ker \mathfrak{e} = \mathbf{k} \cdot 1_A + (\ker \epsilon)^2$ and $\mathfrak{e}(A) \cong (\ker \epsilon) / (\ker \epsilon)^2$ (as $\mathbf{k}$-modules).*

(c) *For each $\mathbf{k}$-module V, let ι_V be the canonical inclusion $V \to \operatorname{Sym} V$. Let $\mathfrak{q}$ be the map*

$$A \xrightarrow{\mathfrak{e}} \mathfrak{e}(A) \xrightarrow{\iota_{\mathfrak{e}(A)}} \operatorname{Sym}(\mathfrak{e}(A)) .$$

Then, $\mathfrak{q} \in \mathfrak{n}(A, \operatorname{Sym}(\mathfrak{e}(A)))$ [61].

(d) *Let $\mathbf{i}$ be the canonical inclusion $\mathfrak{e}(A) \to A$. Recall the universal property of the symmetric algebra: If V is a $\mathbf{k}$-module, if W is a commutative $\mathbf{k}$-algebra, and if $\varphi : V \to W$ is any $\mathbf{k}$-linear map, then there exists a unique $\mathbf{k}$-algebra homomorphism $\Phi : \operatorname{Sym} V \to W$ satisfying $\varphi = \Phi \circ \iota_V$. Applying this to $V = \mathfrak{e}(A)$, $W = A$ and $\varphi = \mathbf{i}$, we conclude that there exists a unique $\mathbf{k}$-algebra homomorphism $\Phi : \operatorname{Sym}(\mathfrak{e}(A)) \to A$ satisfying $\mathbf{i} = \Phi \circ \iota_{\mathfrak{e}(A)}$. Denote this Φ by $\mathfrak{s}$. Then, the maps $\exp^\star \mathfrak{q} : A \to \operatorname{Sym}(\mathfrak{e}(A))$ and $\mathfrak{s} : \operatorname{Sym}(\mathfrak{e}(A)) \to A$ are mutually inverse $\mathbf{k}$-algebra isomorphisms.*

(e) *We have $A \cong \operatorname{Sym}\left((\ker \epsilon) / (\ker \epsilon)^2\right)$ as $\mathbf{k}$-algebras.*

(f) *The map $\mathfrak{e} : A \to A$ is a projection (i.e., it satisfies $\mathfrak{e} \circ \mathfrak{e} = \mathfrak{e}$).*

Remark 1.7.30. (a) The main upshot of Theorem 1.7.29 is that any commutative connected graded $\mathbf{k}$-bialgebra A (where $\mathbf{k}$ is a commutative $\mathbb{Q}$-algebra) is isomorphic **as a k-algebra** to the symmetric algebra $\operatorname{Sym} W$ of some $\mathbf{k}$-module W. (Specifically, Theorem 1.7.29(e) claims this for $W = (\ker \epsilon) / (\ker \epsilon)^2$, whereas Theorem 1.7.29(d) claims this for $W = \mathfrak{e}(A)$; these two modules W are isomorphic by Theorem 1.7.29(b).) This is a useful statement even without any specific knowledge about W, since symmetric algebras are a far tamer class of algebras than arbitrary commutative algebras. For example, if $\mathbf{k}$ is a field, then symmetric algebras are just polynomial algebras (up to isomorphism). This can be applied, for example, to the case of the shuffle algebra $\operatorname{Sh}(V)$ of a $\mathbf{k}$-module V. The consequence is that the shuffle algebra $\operatorname{Sh}(V)$ of any $\mathbf{k}$-module V (where $\mathbf{k}$ is a commutative $\mathbb{Q}$-algebra) is isomorphic **as a k-algebra** to a symmetric algebra $\operatorname{Sym} W$. When V is a free $\mathbf{k}$-module, one can actually show that $\operatorname{Sh}(V)$ is isomorphic **as a k-algebra** to the symmetric algebra of a **free** $\mathbf{k}$-module W (that is, to a polynomial ring over $\mathbf{k}$); however, this W is not easy to characterize. Such a characterization is given by *Radford's theorem* (Theorem 6.3.4 below) using the concept of *Lyndon words*. Notice that if V has rank ≥ 2, then W is not finitely generated.

(b) The isomorphism in Theorem 1.7.29(e) is generally not an isomorphism of Hopf algebras. However, with a little (rather straightforward) work, it reveals to be an isomorphism of **graded k**-algebras. Actually, all maps mentioned in Theorem 1.7.29 are graded, provided that we use the appropriate gradings for $\mathfrak{e}(A)$ and $\operatorname{Sym}(\mathfrak{e}(A))$. (To define the appropriate grading for $\mathfrak{e}(A)$, we must show that $\mathfrak{e}$ is a graded map, whence $\mathfrak{e}(A)$ is a homogeneous submodule of A; this provides $\mathfrak{e}(A)$ with the grading we seek. The grading on $\operatorname{Sym}(\mathfrak{e}(A))$ then follows from the usual definition of the grading on the symmetric algebra $\operatorname{Sym} V$ of a graded $\mathbf{k}$-module V: Namely, if V is a graded $\mathbf{k}$-module, then the n-th graded component of $\operatorname{Sym} V$ is defined to be the span of all products of the form $v_1 v_2 \cdots v_k \in \operatorname{Sym} V$, where $v_1, v_2, \ldots, v_k \in V$ are homogeneous elements satisfying $\deg(v_1) + \deg(v_2) + \cdots + \deg(v_k) = n$.)

(c) The map $\mathfrak{e} : A \to A$ from Theorem 1.7.29 is called the *Eulerian idempotent* of A.

(d) Theorem 1.7.29 is concerned with commutative bialgebras. Most of its claims have a "dual version", concerning cocommutative bialgebras. Again, the Eulerian idempotent plays a crucial role; but the result characterizes not the $\mathbf{k}$-algebra structure on A, but the $\mathbf{k}$-coalgebra structure on A. This leads to the Cartier-Milnor-Moore theorem; see [35, §3.8] and [60, §3.2]. We shall say a bit about the Eulerian idempotent for a cocommutative bialgebra in Exercises 5.4.6 and 5.4.8.

Example 1.7.31. Consider the symmetric algebra $\operatorname{Sym} V$ of a $\mathbf{k}$-module V. Then, $\operatorname{Sym} V$ is a commutative connected graded $\mathbf{k}$-bialgebra, and thus Theorem 1.7.29 can be applied to $A = \operatorname{Sym} V$. What is the projection $\mathfrak{e} : A \to A$ obtained in this case?

[60] Keep in mind that $\mathbf{k}$ is assumed to be a commutative $\mathbb{Q}$-algebra.

[61] Do not mistake the map $\mathfrak{q}$ for $\mathfrak{e}$. While every $a \in A$ satisfies $\mathfrak{q}(a) = \mathfrak{e}(a)$, the two maps $\mathfrak{q}$ and $\mathfrak{e}$ have different target sets, and thus we do **not** have $(\exp^\star \mathfrak{q})(a) = (\exp^\star \mathfrak{e})(a)$ for every $a \in A$.

Theorem 1.7.29(b) shows that its kernel is

$$\operatorname{Ker}\mathfrak{e} = \underbrace{\mathbf{k}\cdot 1_A}_{=\operatorname{Sym}^0 V} + \underbrace{(\ker\epsilon)^2}_{=\sum_{n\geq 2}\operatorname{Sym}^n V} = \operatorname{Sym}^0 V + \sum_{n\geq 2}\operatorname{Sym}^n V = \sum_{n\neq 1}\operatorname{Sym}^n V. \tag{1.7.10}$$

This does not yet characterize $\mathfrak{e}$ completely, because we have yet to determine the action of $\mathfrak{e}$ on $\operatorname{Sym}^1 V$. Fortunately, the elements of $\operatorname{Sym}^1 V$ are all primitive (recall that $\Delta_{\operatorname{Sym} V}(v) = 1\otimes v + v\otimes 1$ for each $v\in V$), and it can easily be shown that the map $\mathfrak{e}$ fixes any primitive element of A [62]. Therefore, the map $\mathfrak{e}$ fixes all elements of $\operatorname{Sym}^1 V$. Since we also know that $\mathfrak{e}$ annihilates all elements of $\sum_{n\neq 1}\operatorname{Sym}^n V$ (by (1.7.10)), we thus conclude that $\mathfrak{e}$ is the canonical projection from the direct sum $\operatorname{Sym} V = \bigoplus_{n\in\mathbb{N}}\operatorname{Sym}^n V$ onto its addend $\operatorname{Sym}^1 V$.

Example 1.7.32. For this example, let A be the shuffle algebra $\operatorname{Sh}(V)$ of a $\mathbf{k}$-module V. (See Proposition 1.6.7 for its definition, and keep in mind that its product is being denoted by $\sqcup\!\sqcup$, whereas the notation uv is still being used for the product of two elements u and v in the **tensor** algebra $T(V)$.)

Theorem 1.7.29 can be applied to $A = \operatorname{Sh}(V)$. What is the projection $\mathfrak{e} : A\to A$ obtained in this case?

Let us compute $\mathfrak{e}(v_1v_2)$ for two elements $v_1, v_2\in V$. Indeed, define a map $\widetilde{\operatorname{id}} : A\to A$ by $\widetilde{\operatorname{id}} = \operatorname{id}_A - u_A\epsilon_A$. Then, $\widetilde{\operatorname{id}}\in\mathfrak{n}(A,A)$ and $\log^\star\left(\underbrace{\widetilde{\operatorname{id}} + u_A\epsilon_A}_{=\operatorname{id}_A}\right) = \log^\star(\operatorname{id}_A) = \mathfrak{e}$. Hence, (1.7.8) (applied to $C = A$ and $f = \widetilde{\operatorname{id}}$) shows that

$$\mathfrak{e} = \sum_{n\geq 1}\frac{(-1)^{n-1}}{n}\widetilde{\operatorname{id}}^{\star n}. \tag{1.7.11}$$

Thus, we need to compute $\widetilde{\operatorname{id}}^{\star n}(v_1v_2)$ for each $n\geq 1$.

Notice that the map $\widetilde{\operatorname{id}}$ annihilates A_0, but fixes any element of A_k for $k > 0$. Thus,

$$\widetilde{\operatorname{id}}(w_1w_2\cdots w_k) = \begin{cases} w_1w_2\cdots w_k, & \text{if } k>0;\\ 0, & \text{if } k=0\end{cases}\qquad\text{for any } w_1,w_2,\ldots,w_k\in V.$$

But it is easy to see that the map $\widetilde{\operatorname{id}}^{\star n} : A\to A$ annihilates A_k whenever $n > k$. In particular, for every $n > 2$, the map $\widetilde{\operatorname{id}}^{\star n} : A\to A$ annihilates A_2, and therefore satisfies

$$\widetilde{\operatorname{id}}^{\star n}(v_1v_2) = 0 \qquad (\text{since } v_1v_2\in A_2). \tag{1.7.12}$$

It remains to find $\widetilde{\operatorname{id}}^{\star n}(v_1v_2)$ for $n\in\{1,2\}$.

We have $\widetilde{\operatorname{id}}^{\star 1} = \widetilde{\operatorname{id}}$ and thus

$$\widetilde{\operatorname{id}}^{\star 1}(v_1v_2) = \widetilde{\operatorname{id}}(v_1v_2) = v_1v_2$$

and

$$\begin{aligned}\widetilde{\operatorname{id}}^{\star 2}(v_1v_2) &= \underbrace{\widetilde{\operatorname{id}}(1)}_{=0}\sqcup\!\sqcup\underbrace{\widetilde{\operatorname{id}}(v_1v_2)}_{=v_1v_2} + \underbrace{\widetilde{\operatorname{id}}(v_1)}_{=v_1}\sqcup\!\sqcup\underbrace{\widetilde{\operatorname{id}}(v_2)}_{=v_2} + \underbrace{\widetilde{\operatorname{id}}(v_1v_2)}_{=v_1v_2}\sqcup\!\sqcup\underbrace{\widetilde{\operatorname{id}}(1)}_{=0}\\ &\qquad(\text{since } \Delta_{\operatorname{Sh} V}(v_1v_2) = 1\otimes v_1v_2 + v_1\otimes v_2 + v_1v_2\otimes 1)\\ &= \underbrace{0\sqcup\!\sqcup(v_1v_2)}_{=0} + \underbrace{v_1\sqcup\!\sqcup v_2}_{=v_1v_2+v_2v_1} + \underbrace{(v_1v_2)\sqcup\!\sqcup 0}_{=0}\\ &= v_1v_2 + v_2v_1.\end{aligned}$$

[62] See Exercise 5.4.6(f) further below for this proof. (While Exercise 5.4.6 requires A to be cocommutative, this requirement is not used in the solution to Exercise 5.4.6(f). That said, this requirement is actually satisfied for $A = \operatorname{Sym} V$, so we do not even need to avoid it here.)

Now, applying both sides of (1.7.11) to v_1v_2, we find

$$\begin{aligned}
&\mathfrak{e}\left(v_1v_2\right)\\
&= \sum_{n\geq 1} \frac{(-1)^{n-1}}{n} \widetilde{\operatorname{id}}^{\star n}\left(v_1v_2\right) = \underbrace{\frac{(-1)^{1-1}}{1}}_{=1} \underbrace{\widetilde{\operatorname{id}}^{\star 1}\left(v_1v_2\right)}_{=v_1v_2} + \underbrace{\frac{(-1)^{2-1}}{2}}_{=\frac{-1}{2}} \underbrace{\widetilde{\operatorname{id}}^{\star 2}\left(v_1v_2\right)}_{=v_1v_2+v_2v_1} + \sum_{n\geq 3} \frac{(-1)^{n-1}}{n} \underbrace{\widetilde{\operatorname{id}}^{\star n}\left(v_1v_2\right)}_{\substack{=0\\ \text{(by (1.7.12))}}}\\
&= v_1v_2 + \frac{-1}{2}\left(v_1v_2+v_2v_1\right) + \underbrace{\sum_{n\geq 3} \frac{(-1)^{n-1}}{n} 0}_{=0} = \frac{1}{2}\left(v_1v_2 - v_2v_1\right).
\end{aligned}$$

This describes the action of $\mathfrak{e}$ on the graded component A_2 of $A = \operatorname{Sh}(V)$.

Similarly, we can describe $\mathfrak{e}$ acting on any other graded component:

$$\begin{aligned}
\mathfrak{e}(1) &= 0;\\
\mathfrak{e}\left(v_1\right) &= v_1 \qquad \text{for each } v_1 \in V;\\
\mathfrak{e}\left(v_1v_2\right) &= \frac{1}{2}\left(v_1v_2 - v_2v_1\right) \qquad \text{for any } v_1, v_2 \in V;\\
\mathfrak{e}\left(v_1v_2v_3\right) &= \frac{1}{6}\left(2v_1v_2v_3 - v_1v_3v_2 - v_2v_1v_3 - v_2v_3v_1 - v_3v_1v_2 + 2v_3v_2v_1\right) \qquad \text{for any } v_1, v_2, v_3 \in V,\\
&\ldots.
\end{aligned}$$

With some more work, one can show the following formula for the action of $\mathfrak{e}$ on any nontrivial pure tensor:

$$\begin{aligned}
\mathfrak{e}\left(v_1v_2\cdots v_n\right) &= \sum_{\sigma\in\mathfrak{S}_n} \left(\sum_{k=1+\operatorname{des}\left(\sigma^{-1}\right)}^{n} \frac{(-1)^{k-1}}{k} \binom{n-1-\operatorname{des}\left(\sigma^{-1}\right)}{k-1-\operatorname{des}\left(\sigma^{-1}\right)} \right) v_{\sigma(1)}v_{\sigma(2)}\cdots v_{\sigma(n)}\\
&= \sum_{\sigma\in\mathfrak{S}_n} \frac{(-1)^{\operatorname{des}\left(\sigma^{-1}\right)}}{\operatorname{des}\left(\sigma^{-1}\right)+1} \binom{n}{\operatorname{des}\left(\sigma^{-1}\right)+1}^{-1} v_{\sigma(1)}v_{\sigma(2)}\cdots v_{\sigma(n)}\\
&\qquad \text{for any } n \geq 1 \text{ and } v_1, v_2, \ldots, v_n \in V,
\end{aligned}$$

where we use the notation $\operatorname{des}\pi$ for the number of descents[63] of any permutation $\pi \in \mathfrak{S}_n$. (A statement essentially dual to this appears in [191, Theorem 9.5].)

Theorem 1.7.29(b) yields $\ker\mathfrak{e} = \mathbf{k}\cdot 1_A + (\ker\epsilon)^2$. Notice, however, that $(\ker\epsilon)^2$ means the square of the ideal $\ker\epsilon$ with respect to the shuffle multiplication $⧢$; thus, $(\ker\epsilon)^2$ is the $\mathbf{k}$-linear span of all shuffle products of the form $a ⧢ b$ with $a \in \ker\epsilon$ and $b \in \ker\epsilon$.

Exercise 1.7.33. Prove Theorem 1.7.29.

[**Hint:** (a) is easy. For (b), define an element $\widetilde{\operatorname{id}}$ of $\mathfrak{n}(A,A)$ by $\widetilde{\operatorname{id}} = \operatorname{id}_A - u_A\epsilon_A$. Observe that $\mathfrak{e} = \sum_{n\geq 1} \frac{(-1)^{n-1}}{n} \widetilde{\operatorname{id}}^{\star n}$, and draw the conclusions that $\mathfrak{e}(1_A) = 0$ and that each $x \in A$ satisfies $\widetilde{\operatorname{id}}(x) - \mathfrak{e}(x) \in (\ker\epsilon)^2$ (because $\widetilde{\operatorname{id}}^{\star n}(x) \in (\ker\epsilon)^2$ for every $n \geq 2$). Use this to prove $\ker\mathfrak{e} \subset \mathbf{k}\cdot 1_A + (\ker\epsilon)^2$. On the other hand, prove $\mathfrak{e}\left((\ker\epsilon)^2\right) = 0$ by applying Proposition 1.7.26. Combine to obtain $\ker\mathfrak{e} = \mathbf{k}\cdot 1_A + (\ker\epsilon)^2$. Finish (b) by showing that $A/\left(\mathbf{k}\cdot 1_A + (\ker\epsilon)^2\right) \cong (\ker\epsilon)/(\ker\epsilon)^2$ as $\mathbf{k}$-modules. Part (c) is easy again. For (d), first apply Proposition 1.7.11(i) to show that $\exp^\star(\mathfrak{s}\circ\mathfrak{q}) = \mathfrak{s}\circ(\exp^\star\mathfrak{q})$. In light of $\mathfrak{s}\circ\mathfrak{q} = \mathfrak{e}$ and $\exp^\star\mathfrak{e} = \operatorname{id}_A$, this becomes $\operatorname{id}_A = \mathfrak{s}\circ(\exp^\star\mathfrak{q})$. To obtain part (d), it remains to show that $\exp^\star\mathfrak{q}$ is a surjective $\mathbf{k}$-algebra homomorphism; but this follows from Proposition 1.7.27. For (e), combine (d) and (b). For (f), use once again the observation that each $x \in A$ satisfies $\widetilde{\operatorname{id}}(x) - \mathfrak{e}(x) \in (\ker\epsilon)^2$.]

[63] A *descent* of a permutation $\pi \in \mathfrak{S}_n$ means an $i \in \{1, 2, \ldots, n-1\}$ satisfying $\pi(i) > \pi(i+1)$.

2. Review of symmetric functions Λ as Hopf algebra

Here we review the ring of symmetric functions, borrowing heavily from standard treatments, such as Macdonald [142, Chap. I], Sagan [186, Chap. 4], Stanley [206, Chap. 7], and Mendes and Remmel [154], but emphasizing the Hopf structure early on. Other recent references for this subject are [224], [189], [63], [153, Chapters 2–3] and [187, Chapter 7].

2.1. **Definition of Λ.**

As before, $\mathbf{k}$ here is a commutative ring (hence could be a field or the integers $\mathbb{Z}$; these are the usual choices).

Given an infinite variable set $\mathbf{x} = (x_1, x_2, \ldots)$, a monomial $\mathbf{x}^\alpha := x_1^{\alpha_1} x_2^{\alpha_2} \cdots$ is indexed by a sequence $\alpha = (\alpha_1, \alpha_2, \ldots)$ in $\mathbb{N}^\infty$ having finite support[64]; such sequences α are called *weak compositions*. The nonzero entries of the sequence $\alpha = (\alpha_1, \alpha_2, \ldots)$ are called the *parts* of the weak composition α.

The sum $\alpha_1 + \alpha_2 + \alpha_3 + \cdots$ of all entries of a weak composition $\alpha = (\alpha_1, \alpha_2, \alpha_3, \ldots)$ (or, equivalently, the sum of all parts of α) is called the *size* of α and denoted by $|\alpha|$.

Consider the $\mathbf{k}$-algebra $\mathbf{k}[[\mathbf{x}]] := \mathbf{k}[[x_1, x_2, x_3, \ldots]]$ of all formal power series in the indeterminates $x_1, x_2, x_3, \ldots$ over $\mathbf{k}$; these series are infinite $\mathbf{k}$-linear combinations $\sum_\alpha c_\alpha \mathbf{x}^\alpha$ (with c_α in $\mathbf{k}$) of the monomials $\mathbf{x}^\alpha$ where α ranges over all weak compositions. The product of two such formal power series is well-defined by the usual multiplication rule.

The *degree* of a monomial $\mathbf{x}^\alpha$ is defined to be the number $\deg(\mathbf{x}^\alpha) := \sum_i \alpha_i \in \mathbb{N}$. Given a number $d \in \mathbb{N}$, we say that a formal power series $f(\mathbf{x}) = \sum_\alpha c_\alpha \mathbf{x}^\alpha \in \mathbf{k}[[\mathbf{x}]]$ (with c_α in $\mathbf{k}$) is *homogeneous of degree d* if every weak composition α satisfying $\deg(\mathbf{x}^\alpha) \neq d$ must satisfy $c_\alpha = 0$. In other words, a formal power series is homogeneous of degree d if it is an infinite $\mathbf{k}$-linear combination of monomials of degree d. Every formal power series $f \in \mathbf{k}[[\mathbf{x}]]$ can be uniquely represented as an infinite sum $f_0 + f_1 + f_2 + \cdots$, where each f_d is homogeneous of degree d; in this case, we refer to each f_d as the *d-th homogeneous component of f*. Note that this does not make $\mathbf{k}[[\mathbf{x}]]$ into a graded $\mathbf{k}$-module, since these sums $f_0 + f_1 + f_2 + \cdots$ can have infinitely many nonzero addends. Nevertheless, if f and g are homogeneous power series of degrees d and e, then fg is homogeneous of degree $d + e$.

A formal power series $f(\mathbf{x}) = \sum_\alpha c_\alpha \mathbf{x}^\alpha \in \mathbf{k}[[\mathbf{x}]]$ (with c_α in $\mathbf{k}$) is said to be *of bounded degree* if there exists some bound $d = d(f) \in \mathbb{N}$ such that every weak composition $\alpha = (\alpha_1, \alpha_2, \alpha_3, \ldots)$ satisfying $\deg(\mathbf{x}^\alpha) > d$ must satisfy $c_\alpha = 0$. Equivalently, a formal power series $f \in \mathbf{k}[[\mathbf{x}]]$ is of bounded degree if all but finitely many of its homogeneous components are zero. (For example, $x_1^2 + x_2^2 + x_3^2 + \cdots$ and $1 + x_1 + x_2 + x_3 + \cdots$ are of bounded degree, while $x_1 + x_1x_2 + x_1x_2x_3 + \cdots$ and $1 + x_1 + x_1^2 + x_1^3 + \cdots$ are not.) It is easy to see that the sum and the product of two power series of bounded degree also have bounded degree. Thus, the formal power series of bounded degree form a $\mathbf{k}$-subalgebra of $\mathbf{k}[[\mathbf{x}]]$, which we call $R(\mathbf{x})$. This subalgebra $R(\mathbf{x})$ is graded (by degree).

The symmetric group $\mathfrak{S}_n$ permuting the first n variables $x_1, \ldots, x_n$ acts as a group of automorphisms on $R(\mathbf{x})$, as does the union $\mathfrak{S}_{(\infty)} = \bigcup_{n\geq 0} \mathfrak{S}_n$ of the infinite ascending chain $\mathfrak{S}_0 \subset \mathfrak{S}_1 \subset \mathfrak{S}_2 \subset \cdots$ of symmetric groups[65]. This group $\mathfrak{S}_{(\infty)}$ can also be described as the group of all permutations of the set $\{1, 2, 3, \ldots\}$ which leave all but finitely many elements invariant. It is known as the *finitary symmetric group* on $\{1, 2, 3, \ldots\}$.

The group $\mathfrak{S}_{(\infty)}$ also acts on the set of all weak compositions by permuting their entries:

$$\sigma(\alpha_1, \alpha_2, \alpha_3, \ldots) = \left(\alpha_{\sigma^{-1}(1)}, \alpha_{\sigma^{-1}(2)}, \alpha_{\sigma^{-1}(3)}, \ldots\right)$$
$$\text{for any weak composition } (\alpha_1, \alpha_2, \alpha_3, \ldots) \text{ and any } \sigma \in \mathfrak{S}_{(\infty)}.$$

These two actions are connected by the equality $\sigma(\mathbf{x}^\alpha) = \mathbf{x}^{\sigma\alpha}$ for any weak composition α and any $\sigma \in \mathfrak{S}_{(\infty)}$.

[64] The *support* of a sequence $\alpha = (\alpha_1, \alpha_2, \alpha_3, \ldots) \in \mathbb{N}^\infty$ is defined to be the set of all positive integers i for which $\alpha_i \neq 0$.

[65] This ascending chain is constructed as follows: For every $n \in \mathbb{N}$, there is an injective group homomorphism $\iota_n : \mathfrak{S}_n \to \mathfrak{S}_{n+1}$ which sends every permutation $\sigma \in \mathfrak{S}_n$ to the permutation $\iota_n(\sigma) = \tau \in \mathfrak{S}_{n+1}$ defined by

$$\tau(i) = \begin{cases} \sigma(i), & \text{if } i \leq n; \\ i, & \text{if } i = n+1 \end{cases} \qquad \text{for all } i \in \{1, 2, \ldots, n+1\}.$$

These homomorphisms ι_n for all n form a chain $\mathfrak{S}_0 \xrightarrow{\iota_0} \mathfrak{S}_1 \xrightarrow{\iota_1} \mathfrak{S}_2 \xrightarrow{\iota_2} \cdots$, which is often regarded as a chain of inclusions.

Definition 2.1.1. The *ring of symmetric functions in* $\mathbf{x}$ *with coefficients in* $\mathbf{k}$, denoted $\Lambda = \Lambda_{\mathbf{k}} = \Lambda(\mathbf{x}) = \Lambda_{\mathbf{k}}(\mathbf{x})$, is the $\mathfrak{S}_{(\infty)}$-invariant subalgebra $R(\mathbf{x})^{\mathfrak{S}_{(\infty)}}$ of $R(\mathbf{x})$:

$$\begin{aligned}\Lambda &:= \left\{f \in R(\mathbf{x}) : \sigma(f) = f \text{ for all } \sigma \in \mathfrak{S}_{(\infty)}\right\} \\ &= \left\{f = \sum_{\alpha} c_\alpha \mathbf{x}^\alpha \in R(\mathbf{x}) : c_\alpha = c_\beta \text{ if } \alpha, \beta \text{ lie in the same } \mathfrak{S}_{(\infty)}\text{-orbit}\right\}.\end{aligned}$$

We refer to the elements of Λ as *symmetric functions* (over $\mathbf{k}$); however, despite this terminology, they are not functions in the usual sense.[66]

Note that Λ is a graded $\mathbf{k}$-algebra, since $\Lambda = \bigoplus_{n \geq 0} \Lambda_n$ where Λ_n are the symmetric functions $f = \sum_\alpha c_\alpha \mathbf{x}^\alpha$ which are *homogeneous of degree* n, meaning $\deg(\mathbf{x}^\alpha) = n$ for all $c_\alpha \neq 0$.

Exercise 2.1.2. Let $f \in R(\mathbf{x})$. Let A be a commutative $\mathbf{k}$-algebra, and $a_1, a_2, \ldots, a_k$ be finitely many elements of A. Show that substituting $a_1, a_2, \ldots, a_k, 0, 0, 0, \ldots$ for $x_1, x_2, x_3, \ldots$ in f yields an infinite sum in which all but finitely many addends are zero. Hence, this sum has a value in A, which is commonly denoted by $f(a_1, a_2, \ldots, a_k)$.

Definition 2.1.3. A *partition* $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_\ell, 0, 0, \ldots)$ is a weak composition whose entries weakly decrease: $\lambda_1 \geq \cdots \geq \lambda_\ell > 0$. The (uniquely defined) ℓ is said to be the *length* of the partition λ and denoted by $\ell(\lambda)$. Thus, $\ell(\lambda)$ is the number of parts[67] of λ. One sometimes omits trailing zeroes from a partition: e.g., one can write the partition $(3, 1, 0, 0, 0, \ldots)$ as $(3, 1)$. We will often (but not always) write λ_i for the i-th entry of the partition λ (for instance, if $\lambda = (5, 3, 1, 1)$, then $\lambda_2 = 3$ and $\lambda_5 = 0$). If λ_i is nonzero, we will also call it the *i-th part* of λ. The sum $\lambda_1 + \lambda_2 + \cdots + \lambda_\ell = \lambda_1 + \lambda_2 + \cdots$ (where $\ell = \ell(\lambda)$) of all entries of λ (or, equivalently, of all parts of λ) is the size $|\lambda|$ of λ. For a given integer n, the partitions of size n are referred to as the *partitions of* n. The empty partition $() = (0, 0, 0, \ldots)$ is denoted by $\varnothing$.

Partitions (as defined above) are sometimes called *integer partitions* in order to distinguish them from set partitions.

Every weak composition α lies in the $\mathfrak{S}_{(\infty)}$-orbit of a unique partition $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_\ell, 0, 0, \ldots)$ with $\lambda_1 \geq \cdots \geq \lambda_\ell > 0$. For any partition λ, define the *monomial symmetric function*

$$m_\lambda := \sum_{\alpha \in \mathfrak{S}_{(\infty)} \lambda} \mathbf{x}^\alpha. \tag{2.1.1}$$

Letting λ run through the set Par of all partitions, this gives the *monomial* $\mathbf{k}$*-basis* $\{m_\lambda\}$ of Λ. Letting λ run only through the set Par_n of partitions of n gives the monomial $\mathbf{k}$-basis for Λ_n.

Example 2.1.4. For $n = 3$, one has

$$\begin{aligned}m_{(3)} &= x_1^3 + x_2^3 + x_3^3 + \cdots, \\ m_{(2,1)} &= x_1^2 x_2 + x_1 x_2^2 + x_1^2 x_3 + x_1 x_3^2 + \cdots, \\ m_{(1,1,1)} &= x_1 x_2 x_3 + x_1 x_2 x_4 + x_1 x_3 x_4 + x_2 x_3 x_4 + x_1 x_2 x_5 + \cdots.\end{aligned}$$

The monomial basis $\{m_\lambda\}_{\lambda \in \mathrm{Par}}$ of Λ is thus a graded basis[68] of the graded $\mathbf{k}$-module Λ. (Here and in the following, when we say that a basis $\{u_\lambda\}_{\lambda \in \mathrm{Par}}$ indexed by Par is a graded basis of Λ, we tacitly understand that Par is partitioned into $\mathrm{Par}_0, \mathrm{Par}_1, \mathrm{Par}_2, \ldots$, so that for each $n \in \mathbb{N}$, the subfamily $\{u_\lambda\}_{\lambda \in \mathrm{Par}_n}$ should be a basis for Λ_n.)

Remark 2.1.5. We have defined the symmetric functions as the elements of $R(\mathbf{x})$ invariant under the group $\mathfrak{S}_{(\infty)}$. However, they also are the elements of $R(\mathbf{x})$ invariant under the group $\mathfrak{S}_\infty$ of *all* permutations of the set $\{1, 2, 3, \ldots\}$ (which acts on $R(\mathbf{x})$ in the same way as its subgroup $\mathfrak{S}_{(\infty)}$ does).[69]

[66] Being power series, they can be evaluated at appropriate families of variables. But this does not make them functions (no more than polynomials are functions). The terminology "symmetric function" is thus not well-chosen; but it is standard.

[67] Recall that a *part* of a partition means a nonzero entry of the partition.

[68] See Definition 1.3.21 for the meaning of "graded basis".

[69] *Proof.* We need to show that $\Lambda = R(\mathbf{x})^{\mathfrak{S}_\infty}$. Since

$$\Lambda = \left\{f = \sum_\alpha c_\alpha \mathbf{x}^\alpha \in R(\mathbf{x}) : c_\alpha = c_\beta \text{ if } \alpha, \beta \text{ lie in the same } \mathfrak{S}_{(\infty)}\text{-orbit}\right\}$$

Remark 2.1.6. It is sometimes convenient to work with finite variable sets $x_1, \ldots, x_n$, which one justifies as follows. Note that the algebra homomorphism

$$R(\mathbf{x}) \to R(x_1, \ldots, x_n) = \mathbf{k}[x_1, \ldots, x_n]$$

which sends $x_{n+1}, x_{n+2}, \ldots$ to 0 restricts to an algebra homomorphism

$$\Lambda_{\mathbf{k}}(\mathbf{x}) \to \Lambda_{\mathbf{k}}(x_1, \ldots, x_n) = \mathbf{k}[x_1, \ldots, x_n]^{\mathfrak{S}_n}.$$

Furthermore, this last homomorphism is a $\mathbf{k}$-module isomorphism when restricted to Λ_i for $0 \leq i \leq n$, since it sends the monomial basis elements $m_\lambda(\mathbf{x})$ to the monomial basis elements $m_\lambda(x_1, \ldots, x_n)$. Thus, when one proves identities in $\Lambda_{\mathbf{k}}(x_1, \ldots, x_n)$ for all n, they are valid in Λ, that is, Λ is the inverse limit of the $\Lambda(x_1, \ldots, x_n)$ in the category of graded $\mathbf{k}$-algebras.[70]

This characterization of Λ as an inverse limit of the graded $\mathbf{k}$-algebras $\Lambda(x_1, \ldots, x_n)$ can be used as an alternative definition of Λ. The definitions used by Macdonald [142] and Wildon [224] are closely related (see [142, §1.2, p. 19, Remark 1], [90, §A.11] and [224, §1.7] for discussions of this definition). It also suggests that much of the theory of symmetric functions can be rewritten in terms of the $\Lambda(x_1, \ldots, x_n)$ (at the cost of extra complexity); and this indeed is possible[71].

One can also define a comultiplication on Λ as follows.

Consider the countably infinite set of variables $(\mathbf{x}, \mathbf{y}) = (x_1, x_2, \ldots, y_1, y_2, \ldots)$. Although it properly contains $\mathbf{x}$, there are nevertheless bijections between $\mathbf{x}$ and $(\mathbf{x}, \mathbf{y})$, since these two variable sets have the same cardinality.

Let $R(\mathbf{x}, \mathbf{y})$ denote the $\mathbf{k}$-algebra of formal power series in $(\mathbf{x}, \mathbf{y})$ of bounded degree. Let $\mathfrak{S}_{(\infty,\infty)}$ be the group of all permutations of $\{x_1, x_2, \ldots, y_1, y_2, \ldots\}$ leaving all but finitely many variables invariant. Then, $\mathfrak{S}_{(\infty,\infty)}$ acts on $R(\mathbf{x}, \mathbf{y})$ by permuting variables, in the same way as $\mathfrak{S}_{(\infty)}$ acts on $R(\mathbf{x})$. The fixed space $R(\mathbf{x}, \mathbf{y})^{\mathfrak{S}_{(\infty,\infty)}}$ is a $\mathbf{k}$-algebra, which we denote by $\Lambda(\mathbf{x}, \mathbf{y})$. This $\mathbf{k}$-algebra $\Lambda(\mathbf{x}, \mathbf{y})$ is isomorphic to $\Lambda = \Lambda(\mathbf{x})$, since there is a bijection between the two sets of variables $(\mathbf{x}, \mathbf{y})$ and $\mathbf{x}$. More explicitly: The map

$$\begin{array}{rcl} \Lambda = \Lambda(\mathbf{x}) & \xrightarrow{\Delta} & \Lambda(\mathbf{x}, \mathbf{y}), \\ f(\mathbf{x}) = f(x_1, x_2, \ldots) & \longmapsto & f(\mathbf{x}, \mathbf{y}) = f(x_1, x_2, \ldots, y_1, y_2, \ldots) \end{array} \tag{2.1.2}$$

is a graded $\mathbf{k}$-algebra isomorphism. Here, $f(x_1, x_2, \ldots, y_1, y_2, \ldots)$ means the result of choosing some bijection $\phi : \{x_1, x_2, x_3, \ldots\} \to \{x_1, x_2, \ldots, y_1, y_2, \ldots\}$ and substituting $\phi(x_i)$ for every x_i in f. (The choice of ϕ is irrelevant since f is symmetric.[72])

The group $\mathfrak{S}_{(\infty)} \times \mathfrak{S}_{(\infty)}$ is a subgroup of the group $\mathfrak{S}_{(\infty,\infty)}$ (via the obvious injection, which lets each $(\sigma, \tau) \in \mathfrak{S}_{(\infty)} \times \mathfrak{S}_{(\infty)}$ act by separately permuting the $x_1, x_2, x_3, \ldots$ using σ and permuting the $y_1, y_2, y_3, \ldots$ using τ), and thus also acts on $R(\mathbf{x}, \mathbf{y})$. Hence, we have an inclusion of $\mathbf{k}$-algebras $\Lambda(\mathbf{x}, \mathbf{y}) = R(\mathbf{x}, \mathbf{y})^{\mathfrak{S}_{(\infty,\infty)}} \subset R(\mathbf{x}, \mathbf{y})^{\mathfrak{S}_{(\infty)} \times \mathfrak{S}_{(\infty)}} \subset R(\mathbf{x}, \mathbf{y})$. The $\mathbf{k}$-module $R(\mathbf{x}, \mathbf{y})^{\mathfrak{S}_{(\infty)} \times \mathfrak{S}_{(\infty)}}$ has $\mathbf{k}$-basis $\{m_\lambda(\mathbf{x}) m_\mu(\mathbf{y})\}_{\lambda, \mu \in \operatorname{Par}}$, since $m_\lambda(\mathbf{x}) m_\mu(\mathbf{y})$ is just the sum of all monomials in the $\mathfrak{S}_{(\infty)} \times \mathfrak{S}_{(\infty)}$-orbit of $\mathbf{x}^\lambda \mathbf{y}^\mu$ (and since any $\mathfrak{S}_{(\infty)} \times \mathfrak{S}_{(\infty)}$-orbit of monomials has exactly one representative of the form $\mathbf{x}^\lambda \mathbf{y}^\mu$ with $\lambda, \mu \in \operatorname{Par}$). Here, of course, $\mathbf{y}$ stands for the set of variables $(y_1, y_2, y_3, \ldots)$, and we define $\mathbf{y}^\mu$ to be $y_1^{\mu_1} y_2^{\mu_2} \cdots$.

On the other hand, the map

$$\begin{array}{rcl} R(\mathbf{x}) \otimes R(\mathbf{x}) & \longrightarrow & R(\mathbf{x}, \mathbf{y}), \\ f(\mathbf{x}) \otimes g(\mathbf{x}) & \longmapsto & f(\mathbf{x}) g(\mathbf{y}) \end{array}$$

and

$$R(\mathbf{x})^{\mathfrak{S}_\infty} = \left\{ f = \sum_\alpha c_\alpha \mathbf{x}^\alpha \in R(\mathbf{x}) : c_\alpha = c_\beta \text{ if } \alpha, \beta \text{ lie in the same } \mathfrak{S}_\infty\text{-orbit} \right\},$$

this will follow immediately if we can show that two weak compositions α and β lie in the same $\mathfrak{S}_{(\infty)}$-orbit if and only if they lie in the same $\mathfrak{S}_\infty$-orbit. But this is straightforward to check (in fact, two weak compositions α and β lie in the same orbit under either group if and only if they have the same multiset of nonzero entries).

[70] *Warning:* The word "graded" here is crucial. Indeed, Λ is **not** the inverse limit of the $\Lambda(x_1, \ldots, x_n)$ in the category of $\mathbf{k}$-algebras. In fact, the latter limit is the $\mathbf{k}$-algebra of all symmetric power series f in $\mathbf{k}[\mathbf{x}]$ with the following property: For each $g \in \mathbb{N}$, there exists a $d \in \mathbb{N}$ such that every monomial in f that involves exactly g distinct indeterminates has degree at most d. For example, the power series $(1 + x_1)(1 + x_2)(1 + x_3) \cdots$ and $m_{(1)} + m_{(2,2)} + m_{(3,3,3)} + \cdots$ satisfy this property, although they do not lie in Λ (unless $\mathbf{k}$ is a trivial ring).

[71] See, for example, [119, Chapter SYM], [174] and [138, Chapters 10–11] for various results of this present chapter rewritten in terms of symmetric polynomials in finitely many variables.

[72] To be more precise, the choice of ϕ is irrelevant because f is $\mathfrak{S}_\infty$-invariant, with the notations of Remark 2.1.5.

is a $\mathbf{k}$-algebra homomorphism. Restricting it to $R(\mathbf{x})^{\mathfrak{S}_{(\infty)}} \otimes R(\mathbf{x})^{\mathfrak{S}_{(\infty)}}$, we obtain a $\mathbf{k}$-algebra homomorphism

$$\Lambda \otimes \Lambda = R(\mathbf{x})^{\mathfrak{S}_{(\infty)}} \otimes R(\mathbf{x})^{\mathfrak{S}_{(\infty)}} \longrightarrow R(\mathbf{x},\mathbf{y})^{\mathfrak{S}_{(\infty)} \times \mathfrak{S}_{(\infty)}}, \tag{2.1.3}$$

which is an isomorphism because it sends the basis $\{m_\lambda \otimes m_\mu\}_{\lambda,\mu \in \mathrm{Par}}$ of the $\mathbf{k}$-module $\Lambda \otimes \Lambda$ to the basis $\{m_\lambda(\mathbf{x}) m_\mu(\mathbf{y})\}_{\lambda,\mu \in \mathrm{Par}}$ of the $\mathbf{k}$-module $R(\mathbf{x},\mathbf{y})^{\mathfrak{S}_{(\infty)} \times \mathfrak{S}_{(\infty)}}$. Thus, we get an inclusion of graded $\mathbf{k}$-algebras

$$\Lambda(\mathbf{x},\mathbf{y}) = R(\mathbf{x},\mathbf{y})^{\mathfrak{S}_{(\infty,\infty)}} \hookrightarrow R(\mathbf{x},\mathbf{y})^{\mathfrak{S}_{(\infty)} \times \mathfrak{S}_{(\infty)}} \cong \Lambda \otimes \Lambda$$

where the last isomorphism is the inverse of the one in (2.1.3). This gives a comultiplication

$$\begin{array}{rcl} \Lambda = \Lambda(\mathbf{x}) & \xrightarrow{\Delta} & \Lambda(\mathbf{x},\mathbf{y}) \hookrightarrow \Lambda \otimes \Lambda, \\ f(\mathbf{x}) = f(x_1, x_2, \ldots) & \longmapsto & f(\mathbf{x},\mathbf{y}) = f(x_1, x_2, \ldots, y_1, y_2, \ldots). \end{array}$$

Here, $f(x_1, x_2, \ldots, y_1, y_2, \ldots)$ is understood as in (2.1.2).

Example 2.1.7. One has

$$\begin{aligned} \Delta m_{(2,1)} &= m_{(2,1)}(x_1, x_2, \ldots, y_1, y_2, \ldots) \\ &= x_1^2 x_2 + x_1 x_2^2 + \cdots \\ &\quad + x_1^2 y_1 + x_1^2 y_2 + \cdots \\ &\quad + x_1 y_1^2 + x_1 y_2^2 + \cdots \\ &\quad + y_1^2 y_2 + y_1 y_2^2 + \cdots \\ &= m_{(2,1)}(\mathbf{x}) + m_{(2)}(\mathbf{x}) m_{(1)}(\mathbf{y}) + m_{(1)}(\mathbf{x}) m_{(2)}(\mathbf{y}) + m_{(2,1)}(\mathbf{y}) \\ &= m_{(2,1)} \otimes 1 + m_{(2)} \otimes m_{(1)} + m_{(1)} \otimes m_{(2)} + 1 \otimes m_{(2,1)}. \end{aligned}$$

This example generalizes easily to the following formula:

$$\Delta m_\lambda = \sum_{\substack{(\mu,\nu): \\ \mu \sqcup \nu = \lambda}} m_\mu \otimes m_\nu, \tag{2.1.4}$$

in which $\mu \sqcup \nu$ is the partition obtained by taking the multiset union of the parts of μ and ν, and then reordering them to make them weakly decreasing.

Checking that Δ is coassociative amounts to checking that

$$(\Delta \otimes \mathrm{id}) \circ \Delta f = f(\mathbf{x},\mathbf{y},\mathbf{z}) = (\mathrm{id} \otimes \Delta) \circ \Delta f$$

inside $\Lambda(\mathbf{x},\mathbf{y},\mathbf{z})$ as a subring of $\Lambda \otimes \Lambda \otimes \Lambda$.

The counit $\Lambda \xrightarrow{\epsilon} \mathbf{k}$ is defined in the usual fashion for connected graded coalgebras, namely ϵ annihilates $I = \bigoplus_{n>0} \Lambda_n$, and ϵ is the identity on $\Lambda_0 = \mathbf{k}$; alternatively ϵ sends a symmetric function $f(\mathbf{x})$ to its constant term $f(0, 0, \ldots)$.

Note that Δ is an algebra morphism $\Lambda \to \Lambda \otimes \Lambda$ because it is a composition of maps which are all algebra morphisms. As the unit and counit axioms are easily checked, Λ becomes a connected graded $\mathbf{k}$-bialgebra of finite type, and hence also a Hopf algebra by Proposition 1.4.16. We will identify its antipode more explicitly in Section 2.4 below.

2.2. Other Bases.

We introduce the usual other bases of Λ, and explain their significance later.

Definition 2.2.1. Define the families of *power sum symmetric functions* p_n, *elementary symmetric functions* e_n, and *complete homogeneous symmetric functions* h_n, for $n = 1, 2, 3, \ldots$ by

$$p_n := x_1^n + x_2^n + \cdots = m_{(n)}, \tag{2.2.1}$$

$$e_n := \sum_{i_1 < \cdots < i_n} x_{i_1} \cdots x_{i_n} = m_{(1^n)}, \tag{2.2.2}$$

$$h_n := \sum_{i_1 \le \cdots \le i_n} x_{i_1} \cdots x_{i_n} = \sum_{\lambda \in \mathrm{Par}_n} m_\lambda. \tag{2.2.3}$$

Here, we are using the *multiplicative notation* for partitions: whenever $(m_1, m_2, m_3, \ldots)$ is a weak composition, $(1^{m_1} 2^{m_2} 3^{m_3} \cdots)$ denotes the partition λ such that for every i, the multiplicity of the part i in λ is m_i. The i^{m_i} satisfying $m_i = 0$ are often omitted from this notation, and so the (1^n) in (2.2.2) means

$\Bigg(\underbrace{1,1,\ldots,1}_{n \text{ ones}}\Bigg)$. (For another example, $\left(1^2 3^1 4^3\right) = \left(1^2 2^0 3^1 4^3 5^0 6^0 7^0 \cdots\right)$ means the partition $(4,4,4,3,1,1)$.) By convention, also define $h_0 = e_0 = 1$, and $h_n = e_n = 0$ if $n < 0$. Extend these multiplicatively to partitions $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_\ell)$ with $\lambda_1 \geq \cdots \geq \lambda_\ell > 0$ by setting

$$\begin{aligned} p_\lambda &:= p_{\lambda_1} p_{\lambda_2} \cdots p_{\lambda_\ell}, \\ e_\lambda &:= e_{\lambda_1} e_{\lambda_2} \cdots e_{\lambda_\ell}, \\ h_\lambda &:= h_{\lambda_1} h_{\lambda_2} \cdots h_{\lambda_\ell}. \end{aligned}$$

Also define the *Schur function*

$$s_\lambda := \sum_T \mathbf{x}^{\operatorname{cont}(T)} \qquad (2.2.4)$$

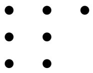
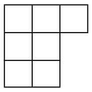
where T runs through all *column-strict tableaux* of shape λ, that is, T is an assignment of entries in $\{1,2,3,\ldots\}$ to the cells of the *Ferrers diagram*[73] for λ, weakly increasing left-to-right in rows, and strictly increasing top-to-bottom in columns. Here $\operatorname{cont}(T)$ denotes the weak composition $\left(|T^{-1}(1)|, |T^{-1}(2)|, |T^{-1}(3)|, \ldots\right)$, so that $\mathbf{x}^{\operatorname{cont}(T)} = \prod_i x_i^{|T^{-1}(i)|}$. For example,[74]

$$T = \begin{array}{ccccc} 1 & 1 & 1 & 4 & 7 \\ 2 & 3 & 3 & & \\ 4 & 4 & 6 & & \\ 6 & 7 & & & \end{array}$$

is a column-strict tableau of shape $\lambda = (5,3,3,2)$ with $\mathbf{x}^{\operatorname{cont}(T)} = x_1^3 x_2^1 x_3^2 x_4^3 x_5^0 x_6^2 x_7^2$. If T is a column-strict tableau, then the weak composition $\operatorname{cont}(T)$ is called the *content* of T.

Column-strict tableaux are also known as *semistandard tableaux*, and some authors even omit the adjective and just call them *tableaux* (e.g., Fulton in [73], a book entirely devoted to them).

[73] The *Ferrers diagram* of a partition λ is defined as the set of all pairs $(i,j) \in \{1,2,3,\ldots\}^2$ satisfying $j \leq \lambda_i$. This is a set of cardinality $|\lambda|$. Usually, one visually represents a Ferrers diagram by drawing its elements (i,j) as points on the plane, although (unlike the standard convention for drawing points on the plane) one lets the x-axis go top-to-bottom (i.e., the point $(i+1,j)$ is one step below the point (i,j)), and the y-axis go left-to-right (i.e., the point $(i,j+1)$ is one step to the right of the point (i,j)). (This is the so-called *English notation*, also known as the *matrix notation* because it is precisely the way one labels the entries of a matrix. Other notations appear in literature, such as the French notation used, e.g., in Malvenuto's [145], and the Russian notation used, e.g., in parts of Kerov's [108].) These points are drawn either as dots or as square boxes; in the latter case, the boxes are centered at the points they represent, and they have sidelength 1 so that the boxes centered around (i,j) and $(i,j+1)$ touch each other along a sideline. For example, the Ferrers diagram of the partition $(3,2,2)$ is represented as

 (using dots) or as  (using boxes).

The Ferrers diagram of a partition λ uniquely determines λ. One refers to the elements of the Ferrers diagram of λ as the *cells* (or *boxes*) of this diagram (which is particularly natural when one represents them by boxes) or, briefly, as the cells of λ. Notation like "west", "north", "left", "right", "row" and "column" concerning cells of Ferrers diagrams normally refers to their visual representation.

Ferrers diagrams are also known as *Young diagrams*.

One can characterize the Ferrers diagrams of partitions as follows: A finite subset S of $\{1,2,3,\ldots\}^2$ is the Ferrers diagram of some partition if and only if for every $(i,j) \in S$ and every $(i',j') \in \{1,2,3,\ldots\}^2$ satisfying $i' \leq i$ and $j' \leq j$, we have $(i',j') \in S$. In other words, a finite subset S of $\{1,2,3,\ldots\}^2$ is the Ferrers diagram of some partition if and only if it is a lower set of the poset $\{1,2,3,\ldots\}^2$ with respect to the componentwise order.

[74] To visually represent a column-strict tableau T of shape λ, we draw the same picture as when representing the Ferrers diagram of λ, but with a little difference: a cell (i,j) is no longer represented by a dot or box, but instead is represented by the entry of T assigned to this cell. Accordingly, the entry of T assigned to a given cell c is often referred to as *the entry of T in c*.

Example 2.2.2. One has

$$\begin{aligned} m_{(1)} &= p_{(1)} = e_{(1)} = h_{(1)} = s_{(1)} = x_1 + x_2 + x_3 + \cdots , \\ s_{(n)} &= h_n, \\ s_{(1^n)} &= e_n. \end{aligned}$$

Example 2.2.3. One has for $\lambda = (2,1)$ that

$$\begin{aligned} p_{(2,1)} &= p_2 p_1 = (x_1^2 + x_2^2 + \cdots)(x_1 + x_2 + \cdots) \\ &= m_{(2,1)} + m_{(3)}, \\[1em] e_{(2,1)} &= e_2 e_1 = (x_1x_2 + x_1x_3 + \cdots)(x_1 + x_2 + \cdots) \\ &= m_{(2,1)} + 3m_{(1,1,1)}, \\[1em] h_{(2,1)} &= h_2 h_1 = (x_1^2 + x_2^2 + \cdots + x_1x_2 + x_1x_3 + \cdots)(x_1 + x_2 + \cdots) \\ &= m_{(3)} + 2m_{(2,1)} + 3m_{(1,1,1)}, \end{aligned}$$

and

$$\begin{array}{llllllllll} s_{(2,1)} & = x_1^2 x_2 & +x_1^2 x_3 & +x_1 x_2^2 & +x_1 x_3^2 & +x_1 x_2 x_3 & +x_1 x_2 x_3 & +x_1 x_2 x_4 & + \cdots \\ & \quad 11 & \quad 11 & \quad 12 & \quad 13 & \quad 12 & \quad 13 & \quad 12 & \\ & \quad 2 & \quad 3 & \quad 2 & \quad 3 & \quad 3 & \quad 2 & \quad 4 & \\ & = m_{(2,1)} + 2m_{(1,1,1)}. & & & & & & & \end{array}$$

In fact, one has these transition matrices for $n = 3$ expressing elements in terms of the monomial basis m_λ:

$$\begin{array}{c} \\ m_{(3)} \\ m_{(2,1)} \\ m_{(1,1,1)} \end{array} \begin{array}{c} \begin{array}{ccc} p_{(3)} & p_{(2,1)} & p_{(1,1,1)} \end{array} \\ \left(\begin{array}{ccc} 1 & 1 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 6 \end{array} \right) \end{array}, \qquad \begin{array}{c} \\ m_{(3)} \\ m_{(2,1)} \\ m_{(1,1,1)} \end{array} \begin{array}{c} \begin{array}{ccc} e_{(3)} & e_{(2,1)} & e_{(1,1,1)} \end{array} \\ \left(\begin{array}{ccc} 0 & 0 & 1 \\ 0 & 1 & 3 \\ 1 & 3 & 6 \end{array} \right) \end{array},$$

$$\begin{array}{c} \\ m_{(3)} \\ m_{(2,1)} \\ m_{(1,1,1)} \end{array} \begin{array}{c} \begin{array}{ccc} h_{(3)} & h_{(2,1)} & h_{(1,1,1)} \end{array} \\ \left(\begin{array}{ccc} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 6 \end{array} \right) \end{array}, \qquad \begin{array}{c} \\ m_{(3)} \\ m_{(2,1)} \\ m_{(1,1,1)} \end{array} \begin{array}{c} \begin{array}{ccc} s_{(3)} & s_{(2,1)} & s_{(1,1,1)} \end{array} \\ \left(\begin{array}{ccc} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 2 & 1 \end{array} \right) \end{array}.$$

Our next goal is to show that $e_\lambda, s_\lambda, h_\lambda$ (and, under some conditions, the p_λ as well) all give bases for Λ. However at the moment it is not yet even clear that s_λ are symmetric!

Proposition 2.2.4. *Schur functions s_λ are symmetric, that is, they lie in Λ.*

Proof. It suffices to show s_λ is symmetric under swapping the variables x_i, x_{i+1}, by providing an involution ι on the set of all column-strict tableaux T of shape λ which switches the $\operatorname{cont}(T)$ for $(i, i+1)\operatorname{cont}(T)$. Restrict attention to the entries $i, i+1$ in T, which must look something like this:

$$\begin{array}{cccccccccccccc} & & & & & & & & i & i & i & i & i+1 & i+1 \\ & i & i & i & i & i & i+1 & i+1 & i+1 & i+1 & i+1 & & & \\ i+1 & i+1 & i+1 & & & & & & & & & & & \end{array}$$

One finds several vertically aligned pairs $\begin{smallmatrix} i \\ i+1 \end{smallmatrix}$. If one were to remove all such pairs, the remaining entries would be a sequence of rows, each looking like this:

$$\underbrace{i, i, \ldots, i}_{r \text{ occurrences}}, \underbrace{i+1, i+1, \ldots, i+1}_{s \text{ occurrences}}. \tag{2.2.5}$$

An involution due to Bender and Knuth tells us to leave fixed all the vertically aligned pairs $\begin{smallmatrix} i \\ i+1 \end{smallmatrix}$, but change each sequence of remaining entries as in (2.2.5) to this:

$$\underbrace{i, i, \ldots, i}_{s \text{ occurrences}}, \underbrace{i+1, i+1, \ldots, i+1}_{r \text{ occurrences}}.$$

For example, the above configuration in T would change to

$$\begin{array}{cccccccccccccc} &&&&&&&& i & i & i & i & i & i+1 \\ & i & i & i & i & i+1 & i+1 & i+1 & i+1 & i+1 & i+1 &&& \\ i & i+1 & i+1 &&&&&&&&&&& \end{array}$$

It is easily checked that this map is an involution, and that it has the effect of swapping $(i, i+1)$ in $\operatorname{cont}(T)$. □

Remark 2.2.5. The symmetry of Schur functions allows one to reformulate them via column-strict tableaux defined with respect to *any* total ordering $\mathcal{L}$ on the positive integers, rather than the usual $1 < 2 < 3 < \cdots$. For example, one can use the *reverse order*[75] $\cdots < 3 < 2 < 1$, or even more exotic orders, such as

$$1 < 3 < 5 < 7 < \cdots < 2 < 4 < 6 < 8 < \cdots .$$

Say that an assignment T of entries in $\{1, 2, 3, \ldots\}$ to the cells of the Ferrers diagram of λ is an *$\mathcal{L}$-column-strict tableau* if it is weakly $\mathcal{L}$-increasing left-to-right in rows, and strictly $\mathcal{L}$-increasing top-to-bottom in columns.

Proposition 2.2.6. *For any total order $\mathcal{L}$ on the positive integers,*

$$s_\lambda = \sum_T \mathbf{x}^{\operatorname{cont}(T)} \tag{2.2.6}$$

as T runs through all $\mathcal{L}$-column-strict tableaux of shape λ.

Proof. Given a weak composition $\alpha = (\alpha_1, \alpha_2, \ldots)$ with $\alpha_{n+1} = \alpha_{n+2} = \cdots = 0$, assume that the integers $1, 2, \ldots, n$ are totally ordered by $\mathcal{L}$ as $w(1) <_{\mathcal{L}} \cdots <_{\mathcal{L}} w(n)$ for some w in $\mathfrak{S}_n$. Then the coefficient of $\mathbf{x}^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ on the right side of (2.2.6) is the same as the coefficient of $\mathbf{x}^{w^{-1}(\alpha)}$ on the right side of (2.2.4) defining s_λ, which by symmetry of s_λ is the same as the coefficient of $\mathbf{x}^\alpha$ on the right side of (2.2.4). □

It is now not hard to show that $p_\lambda, e_\lambda, s_\lambda$ give bases by a triangularity argument[76]. For this purpose, let us introduce a useful partial order on partitions.

Definition 2.2.7. The *dominance* or *majorization* order on Par_n is the partial order on the set Par_n whose greater-or-equal relation $\trianglerighteq$ is defined as follows: For two partitions λ and μ of n, we set $\lambda \trianglerighteq \mu$ (and say that λ *dominates*, or *majorizes*, μ) if and only if

$$\lambda_1 + \lambda_2 + \cdots + \lambda_k \geq \mu_1 + \mu_2 + \cdots + \mu_k \qquad \text{for } k = 1, 2, \ldots, n.$$

(The definition of dominance would not change if we would replace "for $k = 1, 2, \ldots, n$" by "for every positive integer k" or by "for every $k \in \mathbb{N}$".)

Definition 2.2.8. For a partition λ, its *conjugate* or *transpose* partition λ^t is the one whose Ferrers diagram is obtained from that of λ by exchanging rows for columns (i.e., by flipping the diagram across the "main", i.e., top-right-to-bottom-left, diagonal)[77]. Alternatively, one has this formula for its i-th entry:

$$(\lambda^t)_i := |\{j : \lambda_j \geq i\}|. \tag{2.2.7}$$

For example, $(4,3,1)^t = (3,2,2,1)$, which can be easily verified by flipping the Ferrers diagram of $(4,3,1)$ across the "main diagonal":

$$\underbrace{\begin{array}{cccc} \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \\ \bullet &&& \end{array}}_{\text{Ferrers diagram of } (4,3,1)} \qquad \longmapsto \qquad \underbrace{\begin{array}{ccc} \bullet & \bullet & \bullet \\ \bullet & \bullet & \\ \bullet & \bullet & \\ \bullet && \end{array}}_{\text{Ferrers diagram of } (4,2,2,1)}$$

(or simply counting the boxes in each column of this diagram).

[75] This reverse order is what one uses when one defines a Schur function as a generating function for *reverse semistandard tableaux* or *column-strict plane partitions*; see Stanley [206, Proposition 7.10.4].

[76] See Section 11.1 for some notions and notations that will be used in this argument.

[77] In more rigorous terms: The cells of the Ferrers diagram of λ^t are the pairs (j, i), where (i, j) ranges over all cells of λ. It is easy to see that this indeed uniquely determines a partition λ^t.

Exercise 2.2.9. Let $\lambda, \mu \in \operatorname{Par}_n$. Show that $\lambda \triangleright \mu$ if and only if $\mu^t \triangleright \lambda^t$.

Proposition 2.2.10. *The families $\{e_\lambda\}$ and $\{s_\lambda\}$, as λ runs through all partitions, are graded bases for the graded $\mathbf{k}$-module $\Lambda_{\mathbf{k}}$ whenever $\mathbf{k}$ is a commutative ring. The same holds for the family $\{p_\lambda\}$ when $\mathbb{Q}$ is a subring of $\mathbf{k}$.*

Our proof of this proposition will involve three separate arguments, one for each of the three alleged bases $\{s_\lambda\}$, $\{e_\lambda\}$ and $\{p_\lambda\}$; however, all these three arguments fit the same mold: Each one shows that the alleged basis expands invertibly triangularly[78] in the basis $\{m_\lambda\}$ (possibly after reindexing), with an appropriately chosen partial order on the indexing set. We will simplify our life by restricting ourselves to Par_n for a given $n \in \mathbb{N}$, and by stating the common part of the three arguments in a greater generality (so that we won't have to repeat it thrice):

Lemma 2.2.11. *Let S be a finite poset. We write $\leq$ for the smaller-or-equal relation of S.*

Let M be a free $\mathbf{k}$-module with a basis $(b_\lambda)_{\lambda \in S}$. Let $(a_\lambda)_{\lambda \in S}$ be a further family of elements of M.

For each $\lambda \in S$, let $(g_{\lambda,\mu})_{\mu \in S}$ be the family of the coefficients in the expansion of $a_\lambda \in M$ in the basis $(b_\mu)_{\mu \in S}$; in other words, let $(g_{\lambda,\mu})_{\mu \in S} \in \mathbf{k}^S$ be such that $a_\lambda = \sum\limits_{\mu \in S} g_{\lambda,\mu} b_\mu$. Assume that:

- *Assumption A1: Any $\lambda \in S$ and $\mu \in S$ satisfy $g_{\lambda,\mu} = 0$ unless $\mu \leq \lambda$.*
- *Assumption A2: For any $\lambda \in S$, the element $g_{\lambda,\lambda}$ of $\mathbf{k}$ is invertible.*

Then, the family $(a_\lambda)_{\lambda \in S}$ is a basis of the $\mathbf{k}$-module M.

Proof of Lemma 2.2.11. Use the notations of Section 11.1. Assumptions A1 and A2 yield that the $S \times S$-matrix $(g_{\lambda,\mu})_{(\lambda,\mu) \in S \times S} \in \mathbf{k}^{S \times S}$ is invertibly triangular. But the definition of the $g_{\lambda,\mu}$ yields that the family $(a_\lambda)_{\lambda \in S}$ expands in the family $(b_\lambda)_{\lambda \in S}$ through this matrix $(g_{\lambda,\mu})_{(\lambda,\mu) \in S \times S}$. Since the latter matrix is invertibly triangular, this shows that the family $(a_\lambda)_{\lambda \in S}$ expands invertibly triangularly in the family $(b_\lambda)_{\lambda \in S}$. Therefore, Corollary 11.1.19(e) (applied to $(e_s)_{s \in S} = (a_\lambda)_{\lambda \in S}$ and $(f_s)_{s \in S} = (b_\lambda)_{\lambda \in S}$) shows that $(a_\lambda)_{\lambda \in S}$ is a basis of the $\mathbf{k}$-module M (since $(b_\lambda)_{\lambda \in S}$ is a basis of the $\mathbf{k}$-module M). □

Proof of Proposition 2.2.10. We can restrict our attention to each homogeneous component Λ_n and partitions λ of n. Thus, we have to prove that, for each $n \in \mathbb{N}$, the families $(e_\lambda)_{\lambda \in \operatorname{Par}_n}$ and $(s_\lambda)_{\lambda \in \operatorname{Par}_n}$ are bases of the $\mathbf{k}$-module Λ_n, and that the same holds for $(p_\lambda)_{\lambda \in \operatorname{Par}_n}$ if $\mathbb{Q}$ is a subring of $\mathbf{k}$.

Fix $n \in \mathbb{N}$. We already know that $(m_\lambda)_{\lambda \in \operatorname{Par}_n}$ is a basis of the $\mathbf{k}$-module Λ_n.

1. We shall first show that the family $(s_\lambda)_{\lambda \in \operatorname{Par}_n}$ is a basis of the $\mathbf{k}$-module Λ_n.

   For every partition λ, we have $s_\lambda = \sum_{\mu \in \operatorname{Par}} K_{\lambda,\mu} m_\mu$, where the coefficient $K_{\lambda,\mu}$ is the *Kostka number* counting the column-strict tableaux T of shape λ having $\operatorname{cont}(T) = \mu$; this follows because both sides are symmetric functions, and $K_{\lambda,\mu}$ is the coefficient of $\mathbf{x}^\mu$ on both sides[79]. Thus, for every $\lambda \in \operatorname{Par}_n$, one has

$$s_\lambda = \sum_{\mu \in \operatorname{Par}_n} K_{\lambda,\mu} m_\mu \tag{2.2.8}$$

   (since s_λ is homogeneous of degree n). [80] But if λ and μ are partitions satisfying $K_{\lambda,\mu} \neq 0$, then there exists a column-strict tableau T of shape λ having $\operatorname{cont}(T) = \mu$ (since $K_{\lambda,\mu}$ counts such tableaux), and therefore we must have $\lambda_1 + \lambda_2 + \cdots + \lambda_k \geq \mu_1 + \mu_2 + \cdots + \mu_k$ for each positive integer k (since the entries $1, 2, \ldots, k$ in T must all lie within the first k rows of λ); in other words, $\lambda \triangleright \mu$ (if $K_{\lambda,\mu} \neq 0$) [81]. In other words,

$$\text{any } \lambda \in \operatorname{Par}_n \text{ and } \mu \in \operatorname{Par}_n \text{ satisfy } K_{\lambda,\mu} = 0 \text{ unless } \lambda \triangleright \mu. \tag{2.2.9}$$

[78] i.e., triangularly, with all diagonal coefficients being invertible

[79] In general, in order to prove that two symmetric functions f and g are equal, it suffices to show that, for every $\mu \in \operatorname{Par}$, the coefficients of $\mathbf{x}^\mu$ in f and in g are equal. (Indeed, all other coefficients are determined by these coefficients because of the symmetry.)

[80] See Exercise 2.2.13(c) below for a detailed proof of (2.2.8).

[81] See Exercise 2.2.13(d) below for a detailed proof of this fact.

One can also check that $K_{\lambda,\lambda} = 1$ for any $\lambda \in \operatorname{Par}_n$ [82]. Hence,

$$\text{for any } \lambda \in \operatorname{Par}_n \text{, the element } K_{\lambda,\lambda} \text{ of } \mathbf{k} \text{ is invertible.} \tag{2.2.10}$$

Now, let us regard the set Par_n as a poset, whose greater-or-equal relation is $\triangleright$. Lemma 2.2.11 (applied to $S = \operatorname{Par}_n$, $M = \Lambda_n$, $a_\lambda = s_\lambda$, $b_\lambda = m_\lambda$ and $g_{\lambda,\mu} = K_{\lambda,\mu}$) shows that the family $(s_\lambda)_{\lambda \in \operatorname{Par}_n}$ is a basis of the $\mathbf{k}$-module Λ_n (because the Assumptions A1 and A2 of Lemma 2.2.11 are satisfied[83]).

2. Before we show that $(e_\lambda)_{\lambda \in \operatorname{Par}_n}$ is a basis, we define a few notations regarding integer matrices. A $\{0,1\}$-*matrix* means a matrix whose entries belong to the set $\{0,1\}$. If $A \in \mathbb{N}^{\ell \times m}$ is a matrix, then the *row sums* of A means the ℓ-tuple $(r_1, r_2, \ldots, r_\ell)$, where each r_i is the sum of all entries in the i-th row of A; similarly, the *column sums* of A means the m-tuple $(c_1, c_2, \ldots, c_m)$, where each c_j is the sum of all entries in the j-th column of A. (For instance, the row sums of the $\{0,1\}$-matrix $\begin{pmatrix} 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \end{pmatrix}$ is $(2,3)$, whereas its column sums is $(1,2,1,1,0)$.) We identify any k-tuple of nonnegative integers $(a_1, a_2, \ldots, a_k)$ with the weak composition $(a_1, a_2, \ldots, a_k, 0, 0, 0, \ldots)$; thus, the row sums and the column sums of a matrix in $\mathbb{N}^{\ell \times m}$ can be viewed as weak compositions. (For example, the column sums of the matrix $\begin{pmatrix} 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \end{pmatrix}$ is the 5-tuple $(1,2,1,1,0)$, and can be viewed as the weak composition $(1,2,1,1,0,0,0,\ldots)$.)

For every $\lambda \in \operatorname{Par}_n$, one has

$$e_\lambda = \sum_{\mu \in \operatorname{Par}_n} a_{\lambda,\mu} m_\mu, \tag{2.2.11}$$

where $a_{\lambda,\mu}$ counts $\{0,1\}$-matrices (of size $\ell(\lambda) \times \ell(\mu)$) having row sums λ and column sums μ: indeed, when one expands $e_{\lambda_1} e_{\lambda_2} \cdots$, choosing the monomial $x_{j_1} \ldots x_{j_{\lambda_i}}$ in the e_{λ_i} factor corresponds to putting 1's in the i-th row and columns $j_1, \ldots, j_{\lambda_i}$ of the $\{0,1\}$-matrix [84]. Applying (2.2.11) to λ^t instead of λ, we see that

$$e_{\lambda^t} = \sum_{\mu \in \operatorname{Par}_n} a_{\lambda^t,\mu} m_\mu \tag{2.2.12}$$

for every $\lambda \in \operatorname{Par}_n$.

It is not hard to check[85] that $a_{\lambda,\mu}$ vanishes unless $\lambda^t \triangleright \mu$. Applying this to λ^t instead of λ, we conclude that

$$\text{any } \lambda \in \operatorname{Par}_n \text{ and } \mu \in \operatorname{Par}_n \text{ satisfy } a_{\lambda^t,\mu} = 0 \text{ unless } \lambda \triangleright \mu. \tag{2.2.13}$$

Moreover, one can show that $a_{\lambda^t,\lambda} = 1$ for each $\lambda \in \operatorname{Par}_n$ [86]. Hence,

$$\text{for any } \lambda \in \operatorname{Par}_n \text{, the element } a_{\lambda^t,\lambda} \text{ of } \mathbf{k} \text{ is invertible.} \tag{2.2.14}$$

Now, let us regard the set Par_n as a poset, whose greater-or-equal relation is $\triangleright$. Lemma 2.2.11 (applied to $S = \operatorname{Par}_n$, $M = \Lambda_n$, $a_\lambda = e_{\lambda^t}$, $b_\lambda = m_\lambda$ and $g_{\lambda,\mu} = a_{\lambda^t,\mu}$) shows that the family $(e_{\lambda^t})_{\lambda \in \operatorname{Par}_n}$ is a basis of the $\mathbf{k}$-module Λ_n (because the Assumptions A1 and A2 of Lemma 2.2.11 are satisfied[87]). Hence, $(e_\lambda)_{\lambda \in \operatorname{Par}_n}$ is a basis of Λ_n.

3. Assume now that $\mathbb{Q}$ is a subring of $\mathbf{k}$. For every $\lambda \in \operatorname{Par}_n$, one has

$$p_\lambda = \sum_{\mu \in \operatorname{Par}_n} b_{\lambda,\mu} m_\mu, \tag{2.2.15}$$

[82] See Exercise 2.2.13(e) below for a proof of this.

[83] Indeed, they follow from (2.2.9) and (2.2.10), respectively.

[84] See Exercise 2.2.13(g) below for a detailed proof of (2.2.11).

[85] See Exercise 2.2.13(h) below for a proof of this. This is the easy implication in the *Gale-Ryser Theorem*. (The hard implication is the converse: It says that if $\lambda, \mu \in \operatorname{Par}_n$ satisfy $\lambda^t \triangleright \mu$, then there exists a $\{0,1\}$-matrix having row sums λ and column sums μ, so that $a_{\lambda,\mu}$ is a positive integer. This is proven, e.g., in [114], in [46, Theorem 2.4] and in [224, Section 5.2].)

[86] See Exercise 2.2.13(i) below for a proof of this.

[87] Indeed, they follow from (2.2.13) and (2.2.14), respectively.

where $b_{\lambda,\mu}$ counts the ways to partition the nonzero parts $\lambda_1, \ldots, \lambda_\ell$ (where $\ell = \ell(\lambda)$) into blocks such that the sums of the blocks give μ; more formally, $b_{\lambda,\mu}$ is the number of maps $\varphi : \{1, 2, \ldots, \ell\} \to \{1, 2, 3, \ldots\}$ having

$$\mu_j = \sum_{i:\varphi(i)=j} \lambda_i \qquad \text{for all } j = 1, 2, \ldots$$

[88]. Again it is not hard to check that

$$\text{any } \lambda \in \operatorname{Par}_n \text{ and } \mu \in \operatorname{Par}_n \text{ satisfy } b_{\lambda,\mu} = 0 \text{ unless } \mu \triangleright \lambda. \tag{2.2.16}$$

[89] Furthermore, for any $\lambda \in \operatorname{Par}_n$, the element $b_{\lambda,\lambda}$ is a positive integer[90], and thus invertible in $\mathbf{k}$ (since $\mathbb{Q}$ is a subring of $\mathbf{k}$). Thus,

$$\text{for any } \lambda \in \operatorname{Par}_n, \text{ the element } b_{\lambda,\lambda} \text{ of } \mathbf{k} \text{ is invertible} \tag{2.2.17}$$

(although we don't always have $b_{\lambda,\lambda} = 1$ this time).

Now, let us regard the set Par_n as a poset, whose smaller-or-equal relation is $\triangleright$. Lemma 2.2.11 (applied to $S = \operatorname{Par}_n$, $M = \Lambda_n$, $a_\lambda = p_\lambda$, $b_\lambda = m_\lambda$ and $g_{\lambda,\mu} = b_{\lambda,\mu}$) shows that the family $(p_\lambda)_{\lambda \in \operatorname{Par}_n}$ is a basis of the $\mathbf{k}$-module Λ_n (because the Assumptions A1 and A2 of Lemma 2.2.11 are satisfied[91]). □

Remark 2.2.12. When $\mathbb{Q}$ is not a subring of $\mathbf{k}$, the family $\{p_\lambda\}$ is not (in general) a basis of $\Lambda_{\mathbf{k}}$; for instance, $e_2 = \frac{1}{2}\left(p_{(1,1)} - p_2\right) \in \Lambda_{\mathbb{Q}}$ is not in the $\mathbb{Z}$-span of this family. However, if we define $b_{\lambda,\mu}$ as in the above proof, then the $\mathbb{Z}$-linear span of all p_λ equals the $\mathbb{Z}$-linear span of all $b_{\lambda,\lambda} m_\lambda$. Indeed, if $\mu = (\mu_1, \mu_2, \ldots, \mu_k)$ with $k = \ell(\mu)$, then $b_{\mu,\mu}$ is the size of the subgroup of $\mathfrak{S}_k$ consisting of all permutations $\sigma \in \mathfrak{S}_k$ having each i satisfy $\mu_{\sigma(i)} = \mu_i$ [92]. As a consequence, $b_{\mu,\mu}$ divides $b_{\lambda,\mu}$ for every partition μ of the same size as λ (because this group acts[93] freely on the set which is enumerated by $b_{\lambda,\mu}$) [94]. Hence, the $\operatorname{Par}_n \times \operatorname{Par}_n$-matrix $\left(\frac{b_{\lambda,\mu}}{b_{\mu,\mu}}\right)_{(\lambda,\mu) \in \operatorname{Par}_n \times \operatorname{Par}_n}$ has integer entries. Furthermore, this matrix is unitriangular[95] (indeed, (2.2.16) shows that it is triangular, but its diagonal entries are clearly 1) and thus invertibly triangular. But (2.2.15) shows that the family $(p_\lambda)_{\lambda \in \operatorname{Par}_n}$ expands in the family $(b_{\lambda,\lambda} m_\lambda)_{\lambda \in \operatorname{Par}_n}$ through this matrix. Hence, the family $(p_\lambda)_{\lambda \in \operatorname{Par}_n}$ expands invertibly triangularly in the family $(b_{\lambda,\lambda} m_\lambda)_{\lambda \in \operatorname{Par}_n}$. Thus, Corollary 11.1.19(b) (applied to $\mathbb{Z}$, Λ_n, Par_n, $(p_\lambda)_{\lambda \in \operatorname{Par}_n}$ and $(b_{\lambda,\lambda} m_\lambda)_{\lambda \in \operatorname{Par}_n}$ instead of $\mathbf{k}$, M, S, $(e_s)_{s \in S}$ and $(f_s)_{s \in S}$) shows that the $\mathbb{Z}$-submodule of Λ_n spanned by $(p_\lambda)_{\lambda \in \operatorname{Par}_n}$ is the $\mathbb{Z}$-submodule of Λ_n spanned by $(b_{\lambda,\lambda} m_\lambda)_{\lambda \in \operatorname{Par}_n}$.

The purpose of the following exercise is to fill in some details omitted from the proof of Proposition 2.2.10.

Exercise 2.2.13. Let $n \in \mathbb{N}$.

(a) Show that every $f \in \Lambda_n$ satisfies

$$f = \sum_{\mu \in \operatorname{Par}_n} \left([\mathbf{x}^\mu] f\right) m_\mu.$$

Here, $[\mathbf{x}^\mu] f$ denotes the coefficient of the monomial $\mathbf{x}^\mu$ in the power series f.

Now, we introduce a notation (which generalizes the notation $K_{\lambda,\mu}$ from the proof of Proposition 2.2.10): For any partition λ and any weak composition μ, we let $K_{\lambda,\mu}$ denote the number of all column-strict tableaux T of shape λ having $\operatorname{cont}(T) = \mu$.

(b) Prove that this number $K_{\lambda,\mu}$ is well-defined (i.e., there are only finitely many column-strict tableaux T of shape λ having $\operatorname{cont}(T) = \mu$).

[88] See Exercise 2.2.13(k) below for a detailed proof of (2.2.15) (and see Exercise 2.2.13(j) for a proof that the numbers $b_{\lambda,\mu}$ are well-defined).

[89] See Exercise 2.2.13(l) below for a proof of this.

[90] This is proven in Exercise 2.2.13(m) below.

[91] Indeed, they follow from (2.2.16) and (2.2.17), respectively.

[92] See Exercise 2.2.13(n) below for a proof of this.

[93] Specifically, an element σ of the group takes $\varphi : \{1, 2, \ldots, \ell\} \to \{1, 2, 3, \ldots\}$ to $\sigma \circ \varphi$.

[94] See Exercise 2.2.13(o) below for a detailed proof of this.

[95] Here, we are using the terminology defined in Section 11.1, and we are regarding Par_n as a poset whose smaller-or-equal relation is $\triangleright$.

(c) Show that $s_\lambda = \sum_{\mu \in \operatorname{Par}_n} K_{\lambda,\mu} m_\mu$ for every $\lambda \in \operatorname{Par}_n$.
(d) Show that $K_{\lambda,\mu} = 0$ for any partitions $\lambda \in \operatorname{Par}_n$ and $\mu \in \operatorname{Par}_n$ that don't satisfy $\lambda \triangleright \mu$.
(e) Show that $K_{\lambda,\lambda} = 1$ for any $\lambda \in \operatorname{Par}_n$.

Next, we recall a further notation: For any two partitions λ and μ, we let $a_{\lambda,\mu}$ denote the number of all $\{0,1\}$-matrices of size $\ell(\lambda) \times \ell(\mu)$ having row sums λ and column sums μ. (See the proof of Proposition 2.2.10 for the concepts of $\{0,1\}$-matrices and of row sums and column sums.)

(f) Prove that this number $a_{\lambda,\mu}$ is well-defined (i.e., there are only finitely many $\{0,1\}$-matrices of size $\ell(\lambda) \times \ell(\mu)$ having row sums λ and column sums μ).
(g) Show that $e_\lambda = \sum_{\mu \in \operatorname{Par}_n} a_{\lambda,\mu} m_\mu$ for every $\lambda \in \operatorname{Par}_n$.
(h) Show that $a_{\lambda,\mu} = 0$ for any partitions $\lambda \in \operatorname{Par}_n$ and $\mu \in \operatorname{Par}_n$ that don't satisfy $\lambda^t \triangleright \mu$.
(i) Show that $a_{\lambda^t,\lambda} = 1$ for any $\lambda \in \operatorname{Par}_n$.

Next, we introduce a further notation (which generalizes the notation $b_{\lambda,\mu}$ from the proof of Proposition 2.2.10): For any partition λ and any weak composition μ, we let $b_{\lambda,\mu}$ be the number of all maps $\varphi : \{1, 2, \ldots, \ell\} \to \{1, 2, 3, \ldots\}$ satisfying $\left(\mu_j = \sum_{\substack{i \in \{1,2,\ldots,\ell\};\\ \varphi(i)=j}} \lambda_i \text{ for all } j \geq 1 \right)$, where $\ell = \ell(\lambda)$.

(j) Prove that this number $b_{\lambda,\mu}$ is well-defined (i.e., there are only finitely many maps $\varphi : \{1, 2, \ldots, \ell\} \to \{1, 2, 3, \ldots\}$ satisfying $\left(\mu_j = \sum_{\substack{i \in \{1,2,\ldots,\ell\};\\ \varphi(i)=j}} \lambda_i \text{ for all } j \geq 1 \right)$).
(k) Show that $p_\lambda = \sum_{\mu \in \operatorname{Par}_n} b_{\lambda,\mu} m_\mu$ for every $\lambda \in \operatorname{Par}_n$.
(l) Show that $b_{\lambda,\mu} = 0$ for any partitions $\lambda \in \operatorname{Par}_n$ and $\mu \in \operatorname{Par}_n$ that don't satisfy $\mu \triangleright \lambda$.
(m) Show that $b_{\lambda,\lambda}$ is a positive integer for any $\lambda \in \operatorname{Par}_n$.
(n) Show that for any partition $\mu = (\mu_1, \mu_2, \ldots, \mu_k) \in \operatorname{Par}_n$ with $k = \ell(\mu)$, the integer $b_{\mu,\mu}$ is the size of the subgroup of $\mathfrak{S}_k$ consisting of all permutations $\sigma \in \mathfrak{S}_k$ having each i satisfy $\mu_{\sigma(i)} = \mu_i$. (In particular, show that this subgroup is indeed a subgroup.)
(o) Show that $b_{\mu,\mu} \mid b_{\lambda,\mu}$ for every $\lambda \in \operatorname{Par}_n$ and $\mu \in \operatorname{Par}_n$.

The bases $\{p_\lambda\}$ and $\{e_\lambda\}$ of Λ are two examples of *multiplicative bases*: these are bases constructed from a sequence $v_1, v_2, v_3, \ldots$ of symmetric functions by taking all possible finite products. We will soon encounter another example. First, let us observe that the finite products of a sequence $v_1, v_2, v_3, \ldots$ of symmetric functions form a basis of Λ if and only if the sequence is an algebraically independent generating set of Λ. This holds more generally for any commutative algebra, as the following simple exercise shows:

Exercise 2.2.14. Let A be a commutative $\mathbf{k}$-algebra. Let $v_1, v_2, v_3, \ldots$ be some elements of A.

For every partition λ, define an element $v_\lambda \in A$ by $v_\lambda = v_{\lambda_1} v_{\lambda_2} \cdots v_{\lambda_{\ell(\lambda)}}$. Prove the following:

(a) The $\mathbf{k}$-subalgebra of A generated by $v_1, v_2, v_3, \ldots$ is the $\mathbf{k}$-submodule of A spanned by the family $(v_\lambda)_{\lambda \in \operatorname{Par}}$.
(b) The elements $v_1, v_2, v_3, \ldots$ generate the $\mathbf{k}$-algebra A if and only if the family $(v_\lambda)_{\lambda \in \operatorname{Par}}$ spans the $\mathbf{k}$-module A.
(c) The elements $v_1, v_2, v_3, \ldots$ are algebraically independent over $\mathbf{k}$ if and only if the family $(v_\lambda)_{\lambda \in \operatorname{Par}}$ is $\mathbf{k}$-linearly independent.

The next exercise states two well-known identities for the *generating functions* of the sequences $(e_0, e_1, e_2, \ldots)$ and $(h_0, h_1, h_2, \ldots)$, which will be used several times further below:

Exercise 2.2.15. In the ring of formal power series $(\mathbf{k}[[\mathbf{x}]])[[t]]$, prove the two identities

$$\prod_{i=1}^{\infty} (1 - x_i t)^{-1} = 1 + h_1(\mathbf{x}) t + h_2(\mathbf{x}) t^2 + \cdots = \sum_{n \geq 0} h_n(\mathbf{x}) t^n \tag{2.2.18}$$

and

$$\prod_{i=1}^{\infty} (1 + x_i t) = 1 + e_1(\mathbf{x}) t + e_2(\mathbf{x}) t^2 + \cdots = \sum_{n \geq 0} e_n(\mathbf{x}) t^n. \tag{2.2.19}$$

2.3. **Comultiplications.** Thinking about comultiplication $\Lambda \overset{\Delta}{\to} \Lambda \otimes \Lambda$ on Schur functions forces us to immediately confront the following.

Definition 2.3.1. For partitions μ and λ say that $\mu \subseteq \lambda$ if $\mu_i \leq \lambda_i$ for $i = 1, 2, \ldots$. In other words, two partitions μ and λ satisfy $\mu \subseteq \lambda$ if and only if the Ferrers diagram for μ is a subset of the Ferrers diagram of λ. In this case, define the *skew (Ferrers) diagram* λ/μ to be their set difference.[96]

Then define the *skew Schur function* $s_{\lambda/\mu}(\mathbf{x})$ to be the sum $s_{\lambda/\mu} := \sum_T \mathbf{x}^{\operatorname{cont}(T)}$, where the sum ranges over all *column-strict tableaux* T of shape λ/μ, that is, assignments of a value in $\{1, 2, 3, \ldots\}$ to each cell of λ/μ, weakly increasing left-to-right in rows, and strictly increasing top-to-bottom in columns.

Example 2.3.2. Let $\lambda = (5, 3, 3, 2)$ and $\mu = (3, 1, 1, 0)$. Then, $\mu \subseteq \lambda$. The Ferrers diagrams for λ and μ and the skew Ferrers diagram for λ/μ look as follows:

$$\underbrace{\begin{array}{ccccc} \bullet & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & & \\ \bullet & \bullet & \bullet & & \\ \bullet & \bullet & & & \end{array}}_{\text{Ferrers diagram of } \lambda} \qquad \underbrace{\begin{array}{ccc} \bullet & \bullet & \bullet \\ \bullet & & \\ \bullet & & \\ & & \end{array}}_{\text{Ferrers diagram of } \mu} \qquad \underbrace{\begin{array}{ccccc} \cdot & \cdot & \cdot & \bullet & \bullet \\ \cdot & \bullet & \bullet & & \\ \cdot & \bullet & \bullet & & \\ \bullet & \bullet & & & \end{array}}_{\text{skew Ferrers diagram of } \lambda/\mu}$$

(where the small dots represent boxes removed from the diagram). The filling

$$T = \begin{array}{ccccc} \cdot & \cdot & \cdot & 2 & 5 \\ \cdot & 1 & 1 & & \\ 2 & 2 & 4 & & \\ 4 & 5 & & & \end{array}$$

is a column-strict tableau of shape $\lambda/\mu = (5, 3, 3, 2)/(3, 1, 0, 0)$ and it has $\mathbf{x}^{\operatorname{cont}(T)} = x_1^2 x_2^3 x_3^0 x_4^2 x_5^2$.

On the other hand, if we took $\lambda = (5, 3, 1)$ and $\mu = (1, 1, 1, 1)$, then we wouldn't have $\mu \subseteq \lambda$, since $\mu_4 = 1 > 0 = \lambda_4$.

Remark 2.3.3. If μ and λ are partitions such that $\mu \subseteq \lambda$, then $s_{\lambda/\mu} \in \Lambda$. (This is proven similarly as Proposition 2.2.4.) Actually, if $\mu \subseteq \lambda$, then $s_{\lambda/\mu} \in \Lambda_{|\lambda/\mu|}$, where $|\lambda/\mu|$ denotes the number of cells of the skew shape λ/μ (so $|\lambda/\mu| = |\lambda| - |\mu|$).

It is customary to define $s_{\lambda/\mu}$ to be 0 if we don't have $\mu \subseteq \lambda$. This can also be seen by a literal reading of the definition $s_{\lambda/\mu} := \sum_T \mathbf{x}^{\operatorname{cont}(T)}$, as long as we understand that there are no column-strict tableaux of shape λ/μ when λ/μ is not defined.

Clearly, every partition λ satisfies $s_\lambda = s_{\lambda/\varnothing}$.

It is easy to see that two partitions λ and μ satisfy $\mu \subseteq \lambda$ if and only if they satisfy $\mu^t \subseteq \lambda^t$.

Exercise 2.3.4. (a) State and prove an analogue of Proposition 2.2.6 for skew Schur functions.

(b) Let λ, μ, λ' and μ' be partitions such that $\mu \subseteq \lambda$ and $\mu' \subseteq \lambda'$. Assume that the skew Ferrers diagram λ'/μ' can be obtained from the skew Ferrers diagram λ/μ by a 180° rotation.[97] Prove that $s_{\lambda/\mu} = s_{\lambda'/\mu'}$.

Exercise 2.3.5. Let λ and μ be two partitions, and let $k \in \mathbb{N}$ be such that[98] $\mu_k \geq \lambda_{k+1}$. Let F be the skew Ferrers diagram λ/μ. Let $F_{\text{rows}\leq k}$ denote the subset of F consisting of all $(i, j) \in F$ satisfying $i \leq k$. Let $F_{\text{rows}>k}$ denote the subset of F consisting of all $(i, j) \in F$ satisfying $i > k$. Let α and β be two partitions such that $\beta \subseteq \alpha$ and such that the skew Ferrers diagram α/β can be obtained from $F_{\text{rows}\leq k}$ by parallel

[96] In other words, the skew Ferrers diagram λ/μ is the set of all $(i, j) \in \{1, 2, 3, \ldots\}^2$ satisfying $\mu_i < j \leq \lambda_i$.

While the Ferrers diagram for a single partition λ uniquely determines λ, the skew Ferrers diagram λ/μ does not uniquely determine μ and λ. (For instance, it is empty whenever $\lambda = \mu$.) When one wants to keep μ and λ in memory, one speaks of the *skew shape* λ/μ; this simply means the pair (μ, λ). Every notion defined for skew Ferrers diagrams also makes sense for skew shapes, because to any skew shape λ/μ we can assign the skew Ferrers diagram λ/μ (even if not injectively). For instance, the *cells* of the skew shape λ/μ are the cells of the skew Ferrers diagram λ/μ.

One can characterize the skew Ferrers diagrams as follows: A finite subset S of $\{1, 2, 3, \ldots\}^2$ is a skew Ferrers diagram (i.e., there exist two partitions λ and μ such that $\mu \subseteq \lambda$ and such that S is the skew Ferrers diagram λ/μ) if and only if for every $(i, j) \in S$, every $(i', j') \in \{1, 2, 3, \ldots\}^2$ and every $(i'', j'') \in S$ satisfying $i'' \leq i' \leq i$ and $j'' \leq j' \leq j$, we have $(i', j') \in S$.

[97] For example, this happens when $\lambda = (3, 2)$, $\mu = (1)$, $\lambda' = (5, 4)$ and $\mu' = (3, 1)$.

[98] As usual, we write ν_k for the k-th entry of a partition ν.

translation. Let γ and δ be two partitions such that $\delta \subseteq \gamma$ and such that the skew Ferrers diagram γ/δ can be obtained from $F_{\mathrm{rows}>k}$ by parallel translation.[99] Prove that $s_{\lambda/\mu} = s_{\alpha/\beta} s_{\gamma/\delta}$.

Proposition 2.3.6. *The comultiplication $\Lambda \overset{\Delta}{\to} \Lambda \otimes \Lambda$ has the following effect on the symmetric functions discussed so far*[100]*:*

(i) $\Delta p_n = 1 \otimes p_n + p_n \otimes 1$ *for every* $n \geq 1$, *that is, the power sums* p_n *are primitive.*
(ii) $\Delta e_n = \sum_{i+j=n} e_i \otimes e_j$ *for every* $n \in \mathbb{N}$.
(iii) $\Delta h_n = \sum_{i+j=n} h_i \otimes h_j$ *for every* $n \in \mathbb{N}$.
(iv) $\Delta s_\lambda = \sum_{\mu \subseteq \lambda} s_\mu \otimes s_{\lambda/\mu}$ *for any partition* λ.
(v) $\Delta s_{\lambda/\nu} = \sum\limits_{\substack{\mu \in \mathrm{Par}: \\ \nu \subseteq \mu \subseteq \lambda}} s_{\mu/\nu} \otimes s_{\lambda/\mu}$ *for any partitions* λ *and* ν.

Proof. Recall that Δ sends $f(\mathbf{x}) \mapsto f(\mathbf{x}, \mathbf{y})$, and one can easily check that

(i) $p_n(\mathbf{x}, \mathbf{y}) = \sum_i x_i^n + \sum_i y_i^n = p_n(\mathbf{x}) \cdot 1 + 1 \cdot p_n(\mathbf{y})$ for every $n \geq 1$;
(ii) $e_n(\mathbf{x}, \mathbf{y}) = \sum_{i+j=n} e_i(\mathbf{x}) e_j(\mathbf{y})$ for every $n \in \mathbb{N}$;
(iii) $h_n(\mathbf{x}, \mathbf{y}) = \sum_{i+j=n} h_i(\mathbf{x}) h_j(\mathbf{y})$ for every $n \in \mathbb{N}$.

For assertion (iv), note that by (2.2.6), one has

$$s_\lambda(\mathbf{x}, \mathbf{y}) = \sum_T (\mathbf{x}, \mathbf{y})^{\mathrm{cont}(T)}, \tag{2.3.1}$$

where the sum is over column-strict tableaux T of shape λ having entries in the linearly ordered alphabet

$$x_1 < x_2 < \cdots < y_1 < y_2 < \cdots . \tag{2.3.2}$$

[101] For example,

$$T = \begin{array}{ccccc} x_1 & x_1 & x_1 & y_2 & y_5 \\ x_2 & y_1 & y_1 & & \\ y_2 & y_2 & y_4 & & \\ y_4 & y_5 & & & \end{array}$$

is such a tableau of shape $\lambda = (5, 3, 3, 2)$. Note that the restriction of T to the alphabet $\mathbf{x}$ gives a column-strict tableau $T_{\mathbf{x}}$ of some shape $\mu \subseteq \lambda$, and the restriction of T to the alphabet $\mathbf{y}$ gives a column-strict tableau $T_{\mathbf{y}}$ of shape λ/μ (e.g. for T in the example above, the tableau $T_{\mathbf{y}}$ appeared in Example 2.3.2). Consequently, one has

$$\begin{aligned} s_\lambda(\mathbf{x}, \mathbf{y}) &= \sum_T \mathbf{x}^{\mathrm{cont}(T_{\mathbf{x}})} \cdot \mathbf{y}^{\mathrm{cont}(T_{\mathbf{y}})} \\ &= \sum_{\mu \subseteq \lambda} \left(\sum_{T_{\mathbf{x}}} \mathbf{x}^{\mathrm{cont}(T_{\mathbf{x}})} \right) \left(\sum_{T_{\mathbf{y}}} \mathbf{y}^{\mathrm{cont}(T_{\mathbf{y}})} \right) = \sum_{\mu \subseteq \lambda} s_\mu(\mathbf{x}) s_{\lambda/\mu}(\mathbf{y}). \end{aligned} \tag{2.3.3}$$

[99] Here is an example of the situation: $\lambda = (6, 5, 5, 2, 2)$, $\mu = (4, 4, 3, 1)$, $k = 3$ (satisfying $\mu_k = \mu_3 = 3 \geq 2 = \lambda_4 = \lambda_{k+1}$), $\alpha = (3, 2, 2)$, $\beta = (1, 1)$, $\gamma = (2, 2)$, and $\delta = (1)$.

[100] The abbreviated summation indexing $\sum_{i+j=n} t_{i,j}$ used here is intended to mean

$$\sum_{\substack{(i,j) \in \mathbb{N}^2; \\ i+j=n}} t_{i,j}.$$

[101] Here, $(\mathbf{x}, \mathbf{y})^{\mathrm{cont}(T)}$ means the monomial $\prod_{a \in \mathfrak{A}} a^{|T^{-1}(a)|}$, where $\mathfrak{A}$ denotes the totally ordered alphabet $x_1 < x_2 < \cdots < y_1 < y_2 < \cdots$. In other words, $(\mathbf{x}, \mathbf{y})^{\mathrm{cont}(T)}$ is the product of all entries of the tableau T (which is a monomial, since the entries of T are not numbers but variables).

The following rather formal argument should allay any doubts as to why (2.3.1) holds: Let $\mathcal{L}$ denote the totally ordered set which is given by the set $\{1, 2, 3, \ldots\}$ of positive integers, equipped with the total order $1 <_{\mathcal{L}} 3 <_{\mathcal{L}} 5 <_{\mathcal{L}} 7 <_{\mathcal{L}} \cdots <_{\mathcal{L}} 2 <_{\mathcal{L}} 4 <_{\mathcal{L}} 6 <_{\mathcal{L}} 8 <_{\mathcal{L}} \cdots$. Then, (2.2.6) yields $s_\lambda = \sum_T \mathbf{x}^{\mathrm{cont}(T)}$ as T runs through all $\mathcal{L}$-column-strict tableaux of shape λ. Substituting the variables $x_1, y_1, x_2, y_2, x_3, y_3, \ldots$ for $x_1, x_2, x_3, x_4, x_5, x_6, \ldots$ (that is, substituting x_i for x_{2i-1} and y_i for x_{2i}) in this equality, we obtain (2.3.1).

Assertion (v) is obvious in the case when we don't have $\nu \subseteq \lambda$ (in fact, in this case, both $s_{\lambda/\nu}$ and $\sum_{\substack{\mu \in \mathrm{Par}: \\ \nu \subseteq \mu \subseteq \lambda}} s_{\mu/\nu} \otimes s_{\lambda/\mu}$ are clearly zero). In the remaining case, the proof of assertion (v) is similar to that of (iv). (Of course, the tableaux T and $T_{\mathbf{x}}$ now have skew shapes λ/ν and μ/ν, and instead of (2.2.6), we need to use the answer to Exercise 2.3.4(a).) □

Notice that parts (ii) and (iii) of Proposition 2.3.6 are particular cases of part (iv), since $h_n = s_{(n)}$ and $e_n = s_{(1^n)}$.

Exercise 2.3.7. (a) Show that the Hopf algebra Λ is cocommutative.

(b) Show that $\Delta s_{\lambda/\nu} = \sum_{\substack{\mu \in \mathrm{Par}: \\ \nu \subseteq \mu \subseteq \lambda}} s_{\lambda/\mu} \otimes s_{\mu/\nu}$ for any partitions λ and ν.

Exercise 2.3.8. Let $n \in \mathbb{N}$. Consider the finite variable set $(x_1, x_2, \ldots, x_n)$ as a subset of $\mathbf{x} = (x_1, x_2, x_3, \ldots)$. Recall that $f(x_1, x_2, \ldots, x_n)$ is a well-defined element of $\mathbf{k}[x_1, x_2, \ldots, x_n]$ for every $f \in R(\mathbf{x})$ (and therefore also for every $f \in \Lambda$, since $\Lambda \subset R(\mathbf{x})$), according to Exercise 2.1.2.

(a) Show that any two partitions λ and μ satisfy

$$s_{\lambda/\mu}(x_1, x_2, \ldots, x_n) = \sum_{\substack{T \text{ is a column-strict} \\ \text{tableau of shape } \lambda/\mu; \\ \text{all entries of } T \text{ belong} \\ \text{to } \{1,2,\ldots,n\}}} \mathbf{x}^{\mathrm{cont}(T)}.$$

(b) If λ is a partition having more than n parts[102], then show that $s_\lambda(x_1, x_2, \ldots, x_n) = 0$.

Remark 2.3.9. An analogue of Proposition 2.2.10 holds for symmetric polynomials in finitely many variables: Let $N \in \mathbb{N}$. Then, we have

(a) The family $\{m_\lambda(x_1, x_2, \ldots, x_N)\}$, as λ runs through all partitions having length $\leq N$, is a graded basis of the graded $\mathbf{k}$-module $\Lambda(x_1, x_2, \ldots, x_N) = \mathbf{k}[x_1, x_2, \ldots, x_N]^{\mathfrak{S}_N}$.

(b) For any partition λ having length $> N$, we have $m_\lambda(x_1, x_2, \ldots, x_N) = 0$.

(c) The family $\{e_\lambda(x_1, x_2, \ldots, x_N)\}$, as λ runs through all partitions whose parts are all $\leq N$, is a graded basis of the graded $\mathbf{k}$-module $\Lambda(x_1, x_2, \ldots, x_N)$.

(d) The family $\{s_\lambda(x_1, x_2, \ldots, x_N)\}$, as λ runs through all partitions having length $\leq N$, is a graded basis of the graded $\mathbf{k}$-module $\Lambda(x_1, x_2, \ldots, x_N)$.

(e) If $\mathbb{Q}$ is a subring of $\mathbf{k}$, then the family $\{p_\lambda(x_1, x_2, \ldots, x_N)\}$, as λ runs through all partitions having length $\leq N$, is a graded basis of the graded $\mathbf{k}$-module $\Lambda(x_1, x_2, \ldots, x_N)$.

(f) If $\mathbb{Q}$ is a subring of $\mathbf{k}$, then the family $\{p_\lambda(x_1, x_2, \ldots, x_N)\}$, as λ runs through all partitions whose parts are all $\leq N$, is a graded basis of the graded $\mathbf{k}$-module $\Lambda(x_1, x_2, \ldots, x_N)$.

Indeed, the claims (a) and (b) are obvious, while the claims (c), (d) and (e) are proven similarly to our proof of Proposition 2.2.10. We leave the proof of (f) to the reader; this proof can also be found in [138, Theorem 10.86][103].

Claim (c) can be rewritten as follows: The elementary symmetric polynomials $e_i(x_1, x_2, \ldots, x_N)$, for $i \in \{1, 2, \ldots, N\}$, form an algebraically independent generating set of $\Lambda(x_1, x_2, \ldots, x_N)$. This is precisely the well-known theorem (due to Gauss)[104] that every symmetric polynomial in N variables $x_1, x_2, \ldots, x_N$ can be written uniquely as a polynomial in the N elementary symmetric polynomials.

2.4. The antipode, the involution ω, and algebra generators.

Since Λ is a connected graded $\mathbf{k}$-bialgebra, it will have an antipode $\Lambda \xrightarrow{S} \Lambda$ making it a Hopf algebra by Proposition 1.4.16. However, we can identify S more explicitly now.

Proposition 2.4.1. *Each of the families $\{e_n\}_{n=1,2,\ldots}$ and $\{h_n\}_{n=1,2,\ldots}$ are algebraically independent, and generate $\Lambda_{\mathbf{k}}$ as a polynomial algebra for any commutative ring $\mathbf{k}$. The same holds for $\{p_n\}_{n=1,2,\ldots}$ when $\mathbb{Q}$ is a subring of $\mathbf{k}$.*

Furthermore, the antipode S acts as follows:

[102] Recall that the *parts* of a partition are its nonzero entries.

[103] See [138, Remark 10.76] for why [138, Theorem 10.86] is equivalent to our claim (f).

[104] See, e.g., [40, *Symmetric Polynomials*, Theorem 5 and Remark 17] or [221, §5.3] or [26, Theorem 1]. In a slightly different form, it also appears in [119, Theorem (5.10)].

(i) $S(p_n) = -p_n$ *for every positive integer* n.
(ii) $S(e_n) = (-1)^n h_n$ *for every* $n \in \mathbb{N}$.
(iii) $S(h_n) = (-1)^n e_n$ *for every* $n \in \mathbb{N}$.

Proof. The assertion that $\{e_n\}_{n \geq 1}$ are algebraically independent and generate Λ is equivalent to Proposition 2.2.10 asserting that $\{e_\lambda\}_{\lambda \in \mathrm{Par}}$ is a basis for Λ. (Indeed, this equivalence follows from parts (b) and (c) of Exercise 2.2.14, applied to $v_n = e_n$ and $v_\lambda = e_\lambda$.) Thus, the former assertion is true. If $\mathbb{Q}$ is a subring of $\mathbf{k}$, then a similar argument (using p_n and p_λ instead of e_n and e_λ) shows that $\{p_n\}_{n \geq 1}$ are algebraically independent and generate Λ.

The assertion $S(p_n) = -p_n$ follows from Proposition 1.4.17 since p_n is primitive by Proposition 2.3.6(i).

For the remaining assertions, start with the easy generating function identities[105]

$$H(t) := \prod_{i=1}^{\infty} (1 - x_i t)^{-1} = 1 + h_1(\mathbf{x})t + h_2(\mathbf{x})t^2 + \cdots = \sum_{n \geq 0} h_n(\mathbf{x}) t^n; \tag{2.4.1}$$

$$E(t) := \prod_{i=1}^{\infty} (1 + x_i t) = 1 + e_1(\mathbf{x})t + e_2(\mathbf{x})t^2 + \cdots = \sum_{n \geq 0} e_n(\mathbf{x}) t^n. \tag{2.4.2}$$

These show that

$$1 = E(-t)H(t) = \left(\sum_{n \geq 0} e_n(\mathbf{x})(-t)^n \right) \left(\sum_{n \geq 0} h_n(\mathbf{x}) t^n \right). \tag{2.4.3}$$

Hence, equating coefficients of powers of t, we see that for $n = 0, 1, 2, \ldots$ we have

$$\sum_{i+j=n} (-1)^i e_i h_j = \delta_{0,n}. \tag{2.4.4}$$

This lets us recursively express the e_n in terms of h_n and vice-versa:

$$e_0 = 1 = h_0; \tag{2.4.5}$$

$$e_n = e_{n-1}h_1 - e_{n-2}h_2 + e_{n-3}h_3 - \cdots; \tag{2.4.6}$$

$$h_n = h_{n-1}e_1 - h_{n-2}e_2 + h_{n-3}e_3 - \cdots \tag{2.4.7}$$

for $n = 1, 2, 3, \ldots$ Now, let us use the algebraic independence of the generators $\{e_n\}$ for Λ to define a $\mathbf{k}$-algebra endomorphism

$$\begin{array}{ccc} \Lambda & \overset{\omega}{\to} & \Lambda, \\ e_n & \longmapsto & h_n \end{array} \qquad \text{(for positive integers } n\text{)}.$$

Then,

$$\omega(e_n) = h_n \qquad \text{for each } n \geq 0 \tag{2.4.8}$$

(indeed, this holds for $n > 0$ by definition, and for $n = 0$ because $\omega(e_0) = \omega(1) = 1 = h_0$). Hence, the identical form of the two recursions (2.4.6) and (2.4.7) shows that

$$\omega(h_n) = e_n \qquad \text{for each } n \geq 0 \tag{2.4.9}$$

[106]. Combining this with (2.4.8), we conclude that $(\omega \circ \omega)(e_n) = e_n$ for each $n \geq 0$. Therefore, the two $\mathbf{k}$-algebra homomorphisms $\omega \circ \omega : \Lambda \to \Lambda$ and $\mathrm{id} : \Lambda \to \Lambda$ agree on each element of the generating set

[105] See the solution to Exercise 2.2.15 for the proofs of the identities.

[106] Here is this argument in more detail: We must show that $\omega(h_n) = e_n$ for each $n \geq 0$. We shall prove this by strong induction on n. Thus, we fix an $n \geq 0$, and assume as induction hypothesis that $\omega(h_m) = e_m$ for each $m < n$. We must then prove that $\omega(h_n) = e_n$. If $n = 0$, then this is obvious; thus, assume WLOG that $n > 0$. Hence,

$$\begin{aligned} \omega(h_n) &= \omega(h_{n-1}e_1 - h_{n-2}e_2 + h_{n-3}e_3 - \cdots) \qquad (\text{by (2.4.7)}) \\ &= \omega(h_{n-1})\,\omega(e_1) - \omega(h_{n-2})\,\omega(e_2) + \omega(h_{n-3})\,\omega(e_3) - \cdots \qquad (\text{since } \omega \text{ is a } \mathbf{k}\text{-algebra homomorphism}) \\ &= e_{n-1}\omega(e_1) - e_{n-2}\omega(e_2) + e_{n-3}\omega(e_3) - \cdots \qquad (\text{since } \omega(h_m) = e_m \text{ for each } m < n) \\ &= e_{n-1}h_1 - e_{n-2}h_2 + e_{n-3}h_3 - \cdots \qquad (\text{since (2.4.8) shows that } \omega(e_m) = h_m \text{ for each } m \geq 0) \\ &= e_n \qquad (\text{by (2.4.6)}), \end{aligned}$$

as desired. This completes the induction step.

$\{e_n\}$ of Λ. Hence, they are equal, i.e., we have $\omega \circ \omega = \mathrm{id}$. Therefore ω is an involution and therefore a $\mathbf{k}$-algebra automorphism of Λ. This, in turn, yields that the $\{h_n\}$ (being the images of the $\{e_n\}$ under this automorphism) are another algebraically independent generating set for Λ.

For the assertion about the antipode S applied to e_n or h_n, note that the coproduct formulas for e_n, h_n in Proposition 2.3.6(ii),(iii) show that the defining relations for their antipodes (1.4.4) will in this case be

$$\sum_{i+j=n} S(e_i)e_j = \delta_{0,n} = \sum_{i+j=n} e_i S(e_j),$$
$$\sum_{i+j=n} S(h_i)h_j = \delta_{0,n} = \sum_{i+j=n} h_i S(h_j)$$

because $u\epsilon(e_n) = u\epsilon(h_n) = \delta_{0,n}$. Comparing these to (2.4.4), one concludes via induction on n that $S(e_n) = (-1)^n h_n$ and $S(h_n) = (-1)^n e_n$. □

The $\mathbf{k}$-algebra endomorphism ω of Λ defined in the proof of Proposition 2.4.1 is sufficiently important that we record its definition and a selection of fundamental properties:

Definition 2.4.2. Let ω be the $\mathbf{k}$-algebra homomorphism

$$\begin{array}{ccc} \Lambda & \to & \Lambda, \\ e_n & \longmapsto & h_n \end{array} \qquad \text{(for positive integers } n). \tag{2.4.10}$$

This homomorphism ω is known as the *fundamental involution* of Λ.

Proposition 2.4.3. *Consider the fundamental involution ω and the antipode S of the Hopf algebra Λ.*

(a) *We have*
$$\omega(e_n) = h_n \qquad \text{for each } n \in \mathbb{Z}.$$

(b) *We have*
$$\omega(h_n) = e_n \qquad \text{for each } n \in \mathbb{Z}.$$

(c) *We have*
$$\omega(p_n) = (-1)^{n-1} p_n \qquad \text{for each positive integer } n.$$

(d) *The map ω is a $\mathbf{k}$-algebra automorphism of Λ and an involution.*

(e) *If $n \in \mathbb{N}$, then*
$$S(f) = (-1)^n \omega(f) \qquad \text{for all } f \in \Lambda_n. \tag{2.4.11}$$

(f) *The map ω is a Hopf algebra automorphism of Λ.*

(g) *The map S is a Hopf algebra automorphism of Λ.*

(h) *Every partition λ satisfies the three equalities*
$$\omega(h_\lambda) = e_\lambda; \tag{2.4.12}$$
$$\omega(e_\lambda) = h_\lambda; \tag{2.4.13}$$
$$\omega(p_\lambda) = (-1)^{|\lambda| - \ell(\lambda)} p_\lambda. \tag{2.4.14}$$

(i) *The map ω is an isomorphism of graded $\mathbf{k}$-modules.*

(j) *The family $(h_\lambda)_{\lambda \in \mathrm{Par}}$ is a graded basis of the graded $\mathbf{k}$-module Λ.*

Exercise 2.4.4. Prove Proposition 2.4.3.

[**Hint:** Parts (a), (b) and (d) have been shown in the proof of Proposition 2.4.1 above. For part (e), let $D_{-1} : \Lambda \to \Lambda$ be the $\mathbf{k}$-algebra morphism sending each homogeneous $f \in \Lambda_n$ to $(-1)^n f$; then argue that $\omega \circ D_{-1}$ and S are two $\mathbf{k}$-algebra morphisms that agree on all elements of the generating set $\{e_n\}$. Derive part (c) from (d) and Proposition 2.4.1. Part (h) then follows by multiplicativity. For parts (f) and (g), check the coalgebra homomorphism axioms on the e_n. Parts (i) and (j) are easy consequences.]

Proposition 2.4.3(e) shows that the antipode S on Λ is, up to sign, the same as the fundamental involution ω. Thus, studying ω is essentially equivalent to studying S.

Remark 2.4.5. Up to now we have not yet derived how the involution ω and the antipode S act on (skew) Schur functions, which is quite beautiful: If λ and μ are partitions satisfying $\mu \subseteq \lambda$, then

$$\begin{aligned} \omega(s_{\lambda/\mu}) &= s_{\lambda^t/\mu^t}, \\ S(s_{\lambda/\mu}) &= (-1)^{|\lambda/\mu|} s_{\lambda^t/\mu^t} \end{aligned} \tag{2.4.15}$$

where recall that λ^t is the transpose or conjugate partition to λ, and $|\lambda/\mu|$ is the number of squares in the skew diagram λ/μ, that is, $|\lambda/\mu| = n - k$ if λ, μ lie in $\mathrm{Par}_n, \mathrm{Par}_k$ respectively.

We will deduce this later in three ways (once as an exercise using the Pieri rules in Exercise 2.7.11, once again using skewing operators in Exercise 2.8.7, and for the third time from the action of the antipode in QSym on P-partition enumerators in Corollary 5.2.22). However, one could also deduce it immediately from our knowledge of the action of ω and S on e_n, h_n, if we were to prove the following famous *Jacobi-Trudi* and *dual Jacobi-Trudi* formulas[107]:

Theorem 2.4.6. *Skew Schur functions are the following polynomials in* $\{h_n\}, \{e_n\}$*:*

$$s_{\lambda/\mu} = \det(h_{\lambda_i - \mu_j - i + j})_{i,j=1,2,\ldots,\ell}, \tag{2.4.16}$$

$$s_{\lambda^t/\mu^t} = \det(e_{\lambda_i - \mu_j - i + j})_{i,j=1,2,\ldots,\ell} \tag{2.4.17}$$

for any two partitions λ *and* μ *and any* $\ell \in \mathbb{N}$ *satisfying* $\ell(\lambda) \leq \ell$ *and* $\ell(\mu) \leq \ell$.

Since we appear not to need these formulas in the sequel, we will not prove them right away. However, a proof is sketched in the solution to Exercise 2.7.13, and various proofs are well-explained in [126, (39) and (41)], [142, §I.5], [184, Thm. 7.1], [186, §4.5], [206, §7.16], [220, Thms. 3.5 and 3.5*]; also, a simultaneous generalization of both formulas is shown in [83, Theorem 11], and three others in [181, 1.9], [88, Thm. 3.1] and [105]. An elegant treatment of Schur polynomials taking the Jacobi-Trudi formula (2.4.16) as the *definition* of s_λ is given by Tamvakis [215].

2.5. **Cauchy product, Hall inner product, self-duality.** The Schur functions, although a bit unmotivated right now, have special properties with regard to the Hopf structure. One property is intimately connected with the following *Cauchy identity*.

Theorem 2.5.1. *In the power series ring* $\mathbf{k}[[\mathbf{x}, \mathbf{y}]] := \mathbf{k}[[x_1, x_2, \ldots, y_1, y_2, \ldots]]$, *one has the following expansion:*

$$\prod_{i,j=1}^{\infty} (1 - x_i y_j)^{-1} = \sum_{\lambda \in \mathrm{Par}} s_\lambda(\mathbf{x}) s_\lambda(\mathbf{y}). \tag{2.5.1}$$

Remark 2.5.2. The left hand side of (2.5.1) is known as the *Cauchy product*, or *Cauchy kernel*.

An equivalent version of the equality (2.5.1) is obtained by replacing each x_i by $x_i t$, and writing the resulting identity in the power series ring $R(\mathbf{x}, \mathbf{y})[[t]]$:

$$\prod_{i,j=1}^{\infty} (1 - t x_i y_j)^{-1} = \sum_{\lambda \in \mathrm{Par}} t^{|\lambda|} s_\lambda(\mathbf{x}) s_\lambda(\mathbf{y}). \tag{2.5.2}$$

(Recall that $|\lambda| = \lambda_1 + \lambda_2 + \cdots + \lambda_\ell$ for any partition $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_\ell)$.)

Proof of Theorem 2.5.1. We follow the standard combinatorial proof (see [186, §4.8],[206, §7.11,7.12]), which rewrites the left and right sides of (2.5.2), and then compares them with the *Robinson-Schensted-Knuth* (RSK) bijection.[108] On the left side, expanding out each geometric series

$$(1 - t x_i y_j)^{-1} = 1 + t x_i y_j + (t x_i y_j)^2 + (t x_i y_j)^3 + \cdots$$

[107] The second of the following identities is also known as the *von Nägelsbach-Kostka identity*.

[108] The RSK bijection has been introduced by Knuth [111], where what we call "biletters" is referred to as "two-line arrays". The most important ingredient of this algorithm – the RS-insertion operation – however goes back to Schensted. The special case of the RSK algorithm where the biword has to be a permutation (written in two-line notation) and the two tableaux have to be *standard* (i.e., each of them has content (1^n), where n is the size of their shape) is the famous *Robinson-Schensted correspondence* [130]. More about these algorithms can be found in [186, Chapter 3], [154, Chapter 5], [206, §7.11-7.12], [138, Sections 10.9–10.22], [73, Chapters 1 and A], [28, §3, §6] and various other places.

and thinking of $(x_i y_j)^m$ as m occurrences of a *biletter*[109] $\binom{i}{j}$, we see that the left hand side can be rewritten as the sum of $t^\ell \left(x_{i_1} y_{j_1}\right) \left(x_{i_2} y_{j_2}\right) \cdots \left(x_{i_\ell} y_{j_\ell}\right)$ over all multisets $\left\{\binom{i_1}{j_1}, \ldots, \binom{i_\ell}{j_\ell}\right\}_{\text{multiset}}$ of biletters. Order the biletters in such a multiset in the lexicographic order $\leq_{lex}$, which is the total order on the set of all biletters defined by

$$\binom{i_1}{j_1} \leq_{lex} \binom{i_2}{j_2} \qquad \Longleftrightarrow \qquad (\text{we have } i_1 \leq i_2, \text{ and if } i_1 = i_2, \text{ then } j_1 \leq j_2)\,.$$

Defining a *biword* to be an array $\binom{\mathbf{i}}{\mathbf{j}} = \binom{i_1 \cdots i_\ell}{j_1 \cdots j_\ell}$ in which the biletters are ordered $\binom{i_1}{j_1} \leq_{lex} \cdots \leq_{lex} \binom{i_\ell}{j_\ell}$, then the left side of (2.5.2) is the sum $\sum t^\ell \mathbf{x}^{\operatorname{cont}(\mathbf{i})} \mathbf{y}^{\operatorname{cont}(\mathbf{j})}$ over all biwords $\binom{\mathbf{i}}{\mathbf{j}}$, where ℓ stands for the number of biletters in the biword. On the right side, expanding out the Schur functions as sums of tableaux gives $\sum_{(P,Q)} t^\ell \mathbf{x}^{\operatorname{cont}(Q)} \mathbf{y}^{\operatorname{cont}(P)}$ in which the sum is over all ordered pairs (P, Q) of column-strict tableaux *having the same shape*[110], with ℓ cells. (We shall refer to such pairs as *tableau pairs* from now on.)

The *Robinson-Schensted-Knuth algorithm* gives us a bijection between the biwords $\binom{\mathbf{i}}{\mathbf{j}}$ and the tableau pairs (P, Q), which has the property that

$$\begin{aligned} \operatorname{cont}(\mathbf{i}) &= \operatorname{cont}(Q), \\ \operatorname{cont}(\mathbf{j}) &= \operatorname{cont}(P) \end{aligned}$$

(and that the length ℓ of the biword $\binom{\mathbf{i}}{\mathbf{j}}$ equals the size $|\lambda|$ of the common shape of P and Q; but this follows automatically from $\operatorname{cont}(\mathbf{i}) = \operatorname{cont}(Q)$). Clearly, once such a bijection is constructed, the equality (2.5.2) will follow.

Before we define this algorithm, we introduce a simpler operation known as *RS-insertion* (short for Robinson-Schensted insertion). RS-insertion takes as input a column-strict tableau P and a letter j, and returns a new column-strict tableau P' along with a corner cell[111] c of P', which is constructed as follows: Start out by setting $P' = P$. The letter j tries to insert itself into the first row of P' by either bumping out the leftmost letter in the first row strictly larger than j, or else placing itself at the right end of the row if no such larger letter exists. If a letter was bumped from the first row, this letter follows the same rules to insert itself into the second row, and so on[112]. This series of bumps must eventually come to an end[113]. At the end of the bumping, the tableau P' created has an extra corner cell not present in P. If we call this corner cell c, then P' (in its final form) and c are what the RS-insertion operation returns. One says that P' is the result of *inserting*[114] j into the tableau P. It is straightforward to see that this resulting filling P' is a column-strict tableau[115].

Example 2.5.3. To give an example of this operation, let us insert the letter $j = 3$ into the column-strict tableau $\begin{array}{ccccc} 1 & 1 & 3 & 3 & 4 \\ 2 & 2 & 4 & 6 & \\ 3 & 4 & 7 & & \\ 5 & & & & \end{array}$ (we are showing all intermediate states of P'; the underlined letter is always the one

[109] A *biletter* here simply means a pair of letters, written as a column vector. A *letter* means a positive integer.

[110] And this shape should be the Ferrers diagram of a partition (not just a skew diagram).

[111] A *corner cell* of a tableau or of a Ferrers diagram is defined to be a cell c which belongs to the tableau (resp. diagram) but whose immediate neighbors to the east and to the south don't. For example, the cell $(3, 2)$ is a corner cell of the Ferrers diagram of the partition $(3, 2, 2, 1)$, and thus also of any tableau whose shape is this partition. But the cell $(2, 2)$ is not a corner cell of this Ferrers diagram, since its immediate neighbor to the south is still in the diagram.

[112] Here, rows are allowed to be empty – so it is possible that a letter is bumped from the last nonempty row of P' and settles in the next, initially empty, row.

[113] since we can only bump out entries from nonempty rows

[114] This terminology is reminiscent of insertion into binary search trees, a basic operation in theoretical computer science. This is more than superficial similarity; there are, in fact, various analogies between Ferrers diagrams (and their fillings) and unlabelled plane binary trees (resp. their labellings), and one of them is the analogy between RS-insertion and binary search tree insertion. See [97, §4.1].

[115] Indeed, the reader can check that P' remains a column-strict tableau throughout the algorithm that defines RS-insertion. (The only part of this that isn't obvious is showing that when a letter t bumped out of some row k is inserted into row $k + 1$, the property that the letters increase strictly down columns is preserved. Argue that the bumping-out of t from row k was caused by the insertion of another letter $u < t$, and that the cell of row $k + 1$ into which t is then being inserted is in the same column as this u, or in a column further left than it.)

that is going to be bumped out at the next step):

$$\begin{array}{lllll} 1 & 1 & 3 & 3 & \underline{4} \\ 2 & 2 & 4 & 6 & \\ 3 & 4 & 7 & & \\ 5 & & & & \end{array} \quad \underset{\substack{\text{insert 3;} \\ \text{bump out 4}}}{\longmapsto} \quad \begin{array}{lllll} 1 & 1 & 3 & 3 & 3 \\ 2 & 2 & 4 & \underline{6} & \\ 3 & 4 & 7 & & \\ 5 & & & & \end{array} \quad \underset{\substack{\text{insert 4;} \\ \text{bump out 6}}}{\longmapsto} \quad \begin{array}{lllll} 1 & 1 & 3 & 3 & 3 \\ 2 & 2 & 4 & 4 & \\ 3 & 4 & \underline{7} & & \\ 5 & & & & \end{array}$$

$$\underset{\substack{\text{insert 6;} \\ \text{bump out 7}}}{\longmapsto} \quad \begin{array}{lllll} 1 & 1 & 3 & 3 & 3 \\ 2 & 2 & 4 & 4 & \\ 3 & 4 & 6 & & \\ 5 & & & & \end{array} \quad \underset{\substack{\text{insert 7;} \\ \text{done}}}{\longmapsto} \quad \begin{array}{lllll} 1 & 1 & 3 & 3 & 3 \\ 2 & 2 & 4 & 4 & \\ 3 & 4 & 6 & & \\ 5 & 7 & & & \end{array} .$$

The last tableau in this sequence is the column-strict tableau that is returned. The corner cell that is returned is the second cell of the fourth row (the one containing 7).

RS-insertion will be used as a step in the RSK algorithm; the construction will rely on a simple fact known as the *row bumping lemma*. Let us first define the notion of a *bumping path* (or *bumping route*): If P is a column-strict tableau, and j is a letter, then some letters are inserted into some cells when RS-insertion is applied to P and j. The sequence of these cells (in the order in which they see letters inserted into them) is called the *bumping path* for P and j. This bumping path always ends with the corner cell c which is returned by RS-insertion. As an example, when $j = 1$ is inserted into the tableau P shown below, the result P' is shown with all entries on the bumping path underlined:

$$P = \begin{array}{lllll} 1 & 1 & 2 & 2 & 3 \\ 2 & 2 & 4 & 4 & \\ 3 & 4 & 5 & & \\ 4 & 6 & 6 & & \end{array} \quad \underset{j=1}{\overset{\text{insert}}{\longmapsto}} \quad P' = \begin{array}{lllll} 1 & 1 & \underline{1} & 2 & 3 \\ 2 & 2 & \underline{2} & 4 & \\ 3 & 4 & \underline{4} & & \\ 4 & \underline{5} & 6 & & \\ \underline{6} & & & & \end{array}$$

A first simple observation about bumping paths is that bumping paths *trend weakly left* – that is, if the bumping path of P and j is $(c_1, c_2, \ldots, c_k)$, then, for each $1 \leq i < k$, the cell c_{i+1} lies in the same column as c_i or in a column further left.[116] A subtler property of bumping paths is the following *row bumping lemma* ([73, p. 9]):

> **Row bumping lemma:** Let P be a column-strict tableau, and let j and j' be two letters. Applying RS-insertion to the tableau P and the letter j yields a new column-strict tableau P' and a corner cell c. Applying RS-insertion to the tableau P' and the letter j' yields a new column-strict tableau P'' and a corner cell c'.
>
> (a) Assume that $j \leq j'$. Then, the bumping path for P' and j' stays strictly to the right, within each row, of the bumping path for P and j. The cell c' (in which the bumping path for P' and j' ends) is in the same row as the cell c (in which the bumping path for P and j ends) or in a row further up; it is also in a column further right than c.
>
> (b) Assume instead that $j > j'$. Then, the bumping path for P' and j' stays weakly to the left, within each row, of the bumping path for P and j. The cell c' (in which the bumping path for P' and j' ends) is in a row further down than the cell c (in which the bumping path for P and j ends); it is also in the same column as c or in a column further left.

This lemma can be easily proven by induction over the row.[117]

[116] This follows easily from the preservation of column-strictness during RS-insertion.

[117] We leave the details to the reader, only giving the main idea for (a) (the proof of (b) is similar). To prove the first claim of (a), it is enough to show that for every i, if any letter is inserted into row i during RS-insertion for P' and j', then some letter is also inserted into row i during RS-insertion for P and j, and the former insertion happens in a cell strictly to the right of the cell where the latter insertion happens. This follows by induction over i. In the induction step, we need to show that if, for a positive integer i, we try to consecutively insert two letters k and k', in this order, into the i-th row of a column-strict tableau, possibly bumping out existing letters in the process, and if we have $k \leq k'$, then the cell into which k is inserted is strictly to the left of the cell into which k' is inserted, and the letter bumped out by the insertion of k is $\leq$ to the letter bumped out by the insertion of k' (or else the insertion of k' bumps out no letter at all – but it cannot happen that k' bumps out a letter but k does not). This statement is completely straightforward to check (by only studying the i-th row). This way, the first claim of (a) is proven, and this entails that the cell c' (being the last cell of the bumping path for P' and j') is in the same row as the cell c or in a row further up. It only remains to show that c' is in a column further right than c. This follows by noticing

We can now define the actual RSK algorithm. Let $\binom{\mathbf{i}}{\mathbf{j}}$ be a biword. Starting with the pair $(P_0, Q_0) = (\varnothing, \varnothing)$ and $m = 0$, the algorithm applies the following steps (see Example 2.5.4 below):

- If i_{m+1} does not exist (that is, m is the length of $\mathbf{i}$), stop.
- Apply RS-insertion to the column-strict tableau P_m and the letter j_{m+1} (the bottom letter of $\binom{i_{m+1}}{j_{m+1}}$). Let P_{m+1} be the resulting column-strict tableau, and let c_{m+1} be the resulting corner cell.
- Create Q_{m+1} from Q_m by adding the top letter i_{m+1} of $\binom{i_{m+1}}{j_{m+1}}$ to Q_m in the cell c_{m+1} (which, as we recall, is the extra corner cell of P_{m+1} not present in P_m).
- Set m to $m + 1$.

After all of the biletters have been thus processed, the result of the RSK algorithm is $(P_\ell, Q_\ell) =: (P, Q)$.

Example 2.5.4. The term in the expansion of the left side of (2.5.1) corresponding to

$$(x_1y_2)^1(x_1y_4)^1(x_2y_1)^1(x_4y_1)^1(x_4y_3)^2(x_5y_2)^1$$

is the biword $\binom{\mathbf{i}}{\mathbf{j}} = \binom{1124445}{2411332}$, whose RSK algorithm goes as follows:

$$\begin{array}{rcl|rcl}
P_0 & = & \varnothing & Q_0 & = & \varnothing \\[1ex]
P_1 & = & 2 & Q_1 & = & 1 \\[1ex]
P_2 & = & 2\ \ 4 & Q_2 & = & 1\ \ 1 \\[1ex]
P_3 & = & \begin{array}{ll} 1 & 4 \\ 2 & \end{array} & Q_3 & = & \begin{array}{ll} 1 & 1 \\ 2 & \end{array} \\[1ex]
P_4 & = & \begin{array}{ll} 1 & 1 \\ 2 & 4 \end{array} & Q_4 & = & \begin{array}{ll} 1 & 1 \\ 2 & 4 \end{array} \\[1ex]
P_5 & = & \begin{array}{lll} 1 & 1 & 3 \\ 2 & 4 & \end{array} & Q_5 & = & \begin{array}{lll} 1 & 1 & 4 \\ 2 & 4 & \end{array} \\[1ex]
P_6 & = & \begin{array}{llll} 1 & 1 & 3 & 3 \\ 2 & 4 & & \end{array} & Q_6 & = & \begin{array}{llll} 1 & 1 & 4 & 4 \\ 2 & 4 & & \end{array} \\[1ex]
P := P_7 & = & \begin{array}{llll} 1 & 1 & 2 & 3 \\ 2 & 3 & & \\ 4 & & & \end{array} & Q := Q_7 & = & \begin{array}{llll} 1 & 1 & 4 & 4 \\ 2 & 4 & & \\ 5 & & & \end{array}
\end{array}$$

The bumping rule obviously maintains the property that P_m is a column-strict tableau of some Ferrers shape throughout. It should be clear that (P_m, Q_m) have the same shape at each stage. Also, the construction of Q_m shows that it is at least weakly increasing in rows and weakly increasing in columns throughout. What is perhaps least clear is that Q_m remains strictly increasing down columns. That is, when one has a string of equal letters on top $i_m = i_{m+1} = \cdots = i_{m+r}$, so that on bottom one bumps in $j_m \leq j_{m+1} \leq \cdots \leq j_{m+r}$, one needs to know that the new cells form a *horizontal strip*, that is, no two of them lie in the same column[118]. This follows from (the last claim of) part (a) of the row bumping lemma. Hence, the result (P, Q) of the RSK algorithm is a tableau pair.

To see that the RSK map is a bijection, we show how to recover $\binom{\mathbf{i}}{\mathbf{j}}$ from (P, Q). This is done by *reverse bumping* from (P_{m+1}, Q_{m+1}) to recover both the biletter $\binom{i_{m+1}}{j_{m+1}}$ and the tableaux (P_m, Q_m), as follows. Firstly, i_{m+1} is the maximum entry of Q_{m+1}, and Q_m is obtained by removing the rightmost occurrence of

that, if k is the row in which the cell c' lies, then c' is in a column further right than the entry of the bumping path for P and j in row k (by the first claim of (a)), and this latter entry is further right than or in the same column as the ultimate entry c of this bumping path (since bumping paths trend weakly left).

[118] Actually, each of these new cells (except for the first one) is in a column further right than the previous one. We will use this stronger fact further below.

this letter i_{m+1} from Q_{m+1}. [119] To produce P_m and j_{m+1}, find the position of the rightmost occurrence of i_{m+1} in Q_{m+1}, and start *reverse bumping* in P_{m+1} from the entry in this same position, where reverse bumping an entry means inserting it into one row higher by having it bump out the rightmost entry which is strictly smaller.[120] The entry bumped out of the first row is j_{m+1}, and the resulting tableau is P_m.

Finally, to see that the RSK map is surjective, one needs to show that the reverse bumping procedure can be applied to any pair (P, Q) of column-strict tableaux of the same shape, and will result in a (lexicographically ordered) biword $\binom{\mathbf{i}}{\mathbf{j}}$. We leave this verification to the reader.[121] □

This is by far not the only known proof of Theorem 2.5.1. Two further proofs will be sketched in Exercise 2.7.10 and Exercise 2.7.8.

Before we move on to extracting identities in Λ from Theorem 2.5.1, let us state (as an exercise) a simple technical fact that will be useful:

Exercise 2.5.5. Let $(q_\lambda)_{\lambda \in \operatorname{Par}}$ be a basis of the $\mathbf{k}$-module Λ. Assume that for each partition λ, the element $q_\lambda \in \Lambda$ is homogeneous of degree $|\lambda|$.

(a) If two families $(a_\lambda)_{\lambda \in \operatorname{Par}} \in \mathbf{k}^{\operatorname{Par}}$ and $(b_\lambda)_{\lambda \in \operatorname{Par}} \in \mathbf{k}^{\operatorname{Par}}$ satisfy

$$\sum_{\lambda \in \operatorname{Par}} a_\lambda q_\lambda (\mathbf{x}) = \sum_{\lambda \in \operatorname{Par}} b_\lambda q_\lambda (\mathbf{x}) \tag{2.5.3}$$

in $\mathbf{k}[[\mathbf{x}]]$, then $(a_\lambda)_{\lambda \in \operatorname{Par}} = (b_\lambda)_{\lambda \in \operatorname{Par}}$. [122]

(b) Consider a further infinite family $\mathbf{y} = (y_1, y_2, y_3, \ldots)$ of indeterminates (disjoint from $\mathbf{x}$). If two families $(a_{\mu,\nu})_{(\mu,\nu) \in \operatorname{Par}^2} \in \mathbf{k}^{\operatorname{Par}^2}$ and $(b_{\mu,\nu})_{(\mu,\nu) \in \operatorname{Par}^2} \in \mathbf{k}^{\operatorname{Par}^2}$ satisfy

$$\sum_{(\mu,\nu) \in \operatorname{Par}^2} a_{\mu,\nu} q_\mu (\mathbf{x}) q_\nu (\mathbf{y}) = \sum_{(\mu,\nu) \in \operatorname{Par}^2} b_{\mu,\nu} q_\mu (\mathbf{x}) q_\nu (\mathbf{y}) \tag{2.5.4}$$

[119] It necessarily has to be the rightmost occurrence, since (according to the previous footnote) the cell into which i_{m+1} was filled at the step from Q_m to Q_{m+1} lies further right than any existing cell of Q_m containing the letter i_{m+1}.

[120] Let us give a few more details on this "reverse bumping" procedure. Reverse bumping (also known as *RS-deletion* or *reverse RS-insertion*) is an operation which takes a column-strict tableau P' and a corner cell c of P', and constructs a column-strict tableau P and a letter j such that RS-insertion for P and j yields P' and c. It starts by setting $P = P'$, and removing the entry in the cell c from P. This removed entry is then denoted by k, and is inserted into the row of P above c, bumping out the rightmost entry which is smaller than k. The letter which is bumped out – say, ℓ –, in turn, is inserted into the row above it, bumping out the rightmost entry which is smaller than ℓ. This procedure continues in the same way until an entry is bumped out of the first row (which will eventually happen). The reverse bumping operation returns the resulting tableau P and the entry which is bumped out of the first row.

It is straightforward to check that the reverse bumping operation is well-defined (i.e., P does stay a column-strict tableau throughout the procedure) and is the inverse of the RS-insertion operation. (In fact, these two operations undo each other step by step.)

[121] It is easy to see that repeatedly applying reverse bumping to (P, Q) will result in a sequence $\binom{i_\ell}{j_\ell}, \binom{i_{\ell-1}}{j_{\ell-1}}, \ldots, \binom{i_1}{j_1}$ of biletters such that applying the RSK algorithm to $\binom{i_1 \cdots i_\ell}{j_1 \cdots j_\ell}$ gives back (P, Q). The question is why we have $\binom{i_1}{j_1} \leq_{lex} \cdots \leq_{lex} \binom{i_\ell}{j_\ell}$. Since the chain of inequalities $i_1 \leq i_2 \leq \cdots \leq i_\ell$ is clear from the choice of entry to reverse-bump, it only remains to show that for every string $i_m = i_{m+1} = \cdots = i_{m+r}$ of equal top letters, the corresponding bottom letters weakly increase (that is, $j_m \leq j_{m+1} \leq \cdots \leq j_{m+r}$). One way to see this is the following:

Assume the contrary; i.e., assume that the bottom letters corresponding to some string $i_m = i_{m+1} = \cdots = i_{m+r}$ of equal top letters do not weakly increase. Thus, $j_{m+p} > j_{m+p+1}$ for some $p \in \{0, 1, \ldots, r-1\}$. Consider this p.

Let us consider the cells containing the equal letters $i_m = i_{m+1} = \cdots = i_{m+r}$ in the tableau Q_{m+r}. Label these cells as $c_m, c_{m+1}, \ldots, c_{m+r}$ from left to right (noticing that no two of them lie in the same column, since Q_{m+r} is column-strict). By the definition of reverse bumping, the first entry to be reverse bumped from P_{m+r} is the entry in position c_{m+r} (since this is the rightmost occurrence of the letter i_{m+r} in Q_{m+r}); then, the next entry to be reverse bumped is the one in position c_{m+r-1}, etc., moving further and further left. Thus, for each $q \in \{0, 1, \ldots, r\}$, the tableau P_{m+q-1} is obtained from P_{m+q} by reverse bumping the entry in position c_{m+q}. Hence, conversely, the tableau P_{m+q} is obtained from P_{m+q-1} by RS-inserting the entry j_{m+q}, which creates the corner cell c_{m+q}.

But recall that $j_{m+p} > j_{m+p+1}$. Hence, part (b) of the row bumping lemma (applied to P_{m+p-1}, j_{m+p}, j_{m+p+1}, P_{m+p}, c_{m+p}, P_{m+p+1} and c_{m+p+1} instead of P, j, j', P', c, P'' and c') shows that the cell c_{m+p+1} is in the same column as the cell c_{m+p} or in a column further left. But this contradicts the fact that the cell c_{m+p+1} is in a column further right than the cell c_{m+p} (since we have labeled our cells as $c_m, c_{m+1}, \ldots, c_{m+r}$ from left to right, and no two of them lied in the same column). This contradiction completes our proof.

[122] Note that this does not immediately follow from the linear independence of the basis $(q_\lambda)_{\lambda \in \operatorname{Par}}$. Indeed, linear independence would help if the sums in (2.5.3) were finite, but they are not. A subtler argument (involving the homogeneity of the q_λ) thus has to be used.

in $\mathbf{k}[[\mathbf{x},\mathbf{y}]]$, then $(a_{\mu,\nu})_{(\mu,\nu)\in\mathrm{Par}^2} = (b_{\mu,\nu})_{(\mu,\nu)\in\mathrm{Par}^2}$.

(c) Consider a further infinite family $\mathbf{z} = (z_1, z_2, z_3, \ldots)$ of indeterminates (disjoint from $\mathbf{x}$ and $\mathbf{y}$). If two families $(a_{\lambda,\mu,\nu})_{(\mu,\nu,\lambda)\in\mathrm{Par}^3} \in \mathbf{k}^{\mathrm{Par}^3}$ and $(b_{\lambda,\mu,\nu})_{(\mu,\nu,\lambda)\in\mathrm{Par}^3} \in \mathbf{k}^{\mathrm{Par}^3}$ satisfy

$$\sum_{(\mu,\nu,\lambda)\in\mathrm{Par}^3} a_{\lambda,\mu,\nu} q_\mu(\mathbf{x}) q_\nu(\mathbf{y}) q_\lambda(\mathbf{z}) = \sum_{(\mu,\nu,\lambda)\in\mathrm{Par}^3} b_{\lambda,\mu,\nu} q_\mu(\mathbf{x}) q_\nu(\mathbf{y}) q_\lambda(\mathbf{z}) \tag{2.5.5}$$

in $\mathbf{k}[[\mathbf{x},\mathbf{y},\mathbf{z}]]$, then $(a_{\lambda,\mu,\nu})_{(\mu,\nu,\lambda)\in\mathrm{Par}^3} = (b_{\lambda,\mu,\nu})_{(\mu,\nu,\lambda)\in\mathrm{Par}^3}$.

Remark 2.5.6. Clearly, for any $n \in \mathbb{N}$, we can state an analogue of Exercise 2.5.5 for n infinite families $\mathbf{x}_i = (x_{i,1}, x_{i,2}, x_{i,3}, \ldots)$ of indeterminates (with $i \in \{1, 2, \ldots, n\}$). The three parts of Exercise 2.5.5 are the particular cases of this analogue for $n = 1$, for $n = 2$ and for $n = 3$. We have shied away from stating this analogue in full generality because these particular cases are the only ones we will need.

Corollary 2.5.7. *In the Schur function basis $\{s_\lambda\}$ for Λ, the structure constants for multiplication and comultiplication are the same, that is, if one defines scalars $c^\lambda_{\mu,\nu}, \hat{c}^\lambda_{\mu,\nu}$ via the unique expansions*

$$s_\mu s_\nu = \sum_\lambda c^\lambda_{\mu,\nu} s_\lambda, \tag{2.5.6}$$

$$\Delta(s_\lambda) = \sum_{\mu,\nu} \hat{c}^\lambda_{\mu,\nu} s_\mu \otimes s_\nu, \tag{2.5.7}$$

then $c^\lambda_{\mu,\nu} = \hat{c}^\lambda_{\mu,\nu}$.

Proof. Work in the ring $\mathbf{k}[[\mathbf{x},\mathbf{y},\mathbf{z}]]$, where $\mathbf{y} = (y_1, y_2, y_3, \ldots)$ and $\mathbf{z} = (z_1, z_2, z_3, \ldots)$ are two new sets of variables. The identity (2.5.1) lets one interpret both $c^\lambda_{\mu,\nu}, \hat{c}^\lambda_{\mu,\nu}$ as the coefficient[123] of $s_\mu(\mathbf{x})s_\nu(\mathbf{y})s_\lambda(\mathbf{z})$ in the product

$$\begin{aligned}\prod_{i,j=1}^{\infty}(1-x_iz_j)^{-1}\prod_{i,j=1}^{\infty}(1-y_iz_j)^{-1} &\overset{(2.5.1)}{=} \left(\sum_\mu s_\mu(\mathbf{x})s_\mu(\mathbf{z})\right)\left(\sum_\nu s_\nu(\mathbf{y})s_\nu(\mathbf{z})\right)\\ &= \sum_{\mu,\nu} s_\mu(\mathbf{x})s_\nu(\mathbf{y}) \cdot s_\mu(\mathbf{z})s_\nu(\mathbf{z})\\ &= \sum_{\mu,\nu} s_\mu(\mathbf{x})s_\nu(\mathbf{y}) \left(\sum_\lambda c^\lambda_{\mu,\nu} s_\lambda(\mathbf{z})\right)\end{aligned}$$

since, regarding $x_1, x_2, \ldots, y_1, y_2, \ldots$ as lying in a single variable set $(\mathbf{x},\mathbf{y})$, separate from the variables $\mathbf{z}$, the Cauchy identity (2.5.1) expands the same product as

$$\begin{aligned}\prod_{i,j=1}^{\infty}(1-x_iz_j)^{-1}\prod_{i,j=1}^{\infty}(1-y_iz_j)^{-1} &= \sum_\lambda s_\lambda(\mathbf{x},\mathbf{y})s_\lambda(\mathbf{z})\\ &= \sum_\lambda\left(\sum_{\mu,\nu}\hat{c}^\lambda_{\mu,\nu}s_\mu(\mathbf{x})s_\nu(\mathbf{y})\right)s_\lambda(\mathbf{z}).\end{aligned}$$

□

Definition 2.5.8. The coefficients $c^\lambda_{\mu,\nu} = \hat{c}^\lambda_{\mu,\nu}$ appearing in the expansions (2.5.6) and (2.5.7) are called *Littlewood-Richardson coefficients*.

Remark 2.5.9. We will interpret $c^\lambda_{\mu,\nu}$ combinatorially in Section 2.6. By now, however, we can already prove some properties of these coefficients:

We have

$$c^\lambda_{\mu,\nu} = c^\lambda_{\nu,\mu} \qquad \text{for all } \lambda, \mu, \nu \in \mathrm{Par} \tag{2.5.8}$$

[123] Let us explain why speaking of coefficients makes sense here:

We want to use the fact that if a power series $f \in \mathbf{k}[[\mathbf{x},\mathbf{y},\mathbf{z}]]$ is written in the form $f = \sum_{(\mu,\nu,\lambda)\in\mathrm{Par}^3} a_{\lambda,\mu,\nu} s_\mu(\mathbf{x}) s_\nu(\mathbf{y}) s_\lambda(\mathbf{z})$ for some coefficients $a_{\lambda,\mu,\nu} \in \mathbf{k}$, then these coefficients $a_{\lambda,\mu,\nu}$ are uniquely determined by f. But this fact is precisely the claim of Exercise 2.5.5(c) above (applied to $q_\lambda = s_\lambda$).

(by comparing coefficients in $\sum_\lambda c^\lambda_{\mu,\nu} s_\lambda = s_\mu s_\nu = s_\nu s_\mu = \sum_\lambda c^\lambda_{\nu,\mu} s_\lambda$). Furthermore, let λ and μ be two partitions (not necessarily satisfying $\mu \subseteq \lambda$). Comparing the expansion

$$s_\lambda(\mathbf{x},\mathbf{y}) = \Delta(s_\lambda) = \sum_{\mu,\nu} c^\lambda_{\mu,\nu} s_\mu(\mathbf{x}) s_\nu(\mathbf{y}) = \sum_{\mu\in\mathrm{Par}} \left(\sum_{\nu\in\mathrm{Par}} c^\lambda_{\mu,\nu} s_\nu(\mathbf{y}) \right) s_\mu(\mathbf{x})$$

with

$$s_\lambda(\mathbf{x},\mathbf{y}) = \sum_{\mu\subseteq\lambda} s_\mu(\mathbf{x}) s_{\lambda/\mu}(\mathbf{y}) = \sum_{\mu\in\mathrm{Par}} s_\mu(\mathbf{x}) s_{\lambda/\mu}(\mathbf{y})$$

[124], one concludes that

$$\sum_{\mu\in\mathrm{Par}} \left(\sum_{\nu\in\mathrm{Par}} c^\lambda_{\mu,\nu} s_\nu(\mathbf{y}) \right) s_\mu(\mathbf{x}) = \sum_{\mu\in\mathrm{Par}} s_\mu(\mathbf{x}) s_{\lambda/\mu}(\mathbf{y}) = \sum_{\mu\in\mathrm{Par}} s_{\lambda/\mu}(\mathbf{y}) s_\mu(\mathbf{x}).$$

Treating the indeterminates $\mathbf{y}$ as constants, and comparing coefficients before $s_\mu(\mathbf{x})$ on both sides of this equality[125], we arrive at another standard interpretation for $c^\lambda_{\mu,\nu}$:

$$s_{\lambda/\mu} = \sum_\nu c^\lambda_{\mu,\nu} s_\nu.$$

In particular, $c^\lambda_{\mu,\nu}$ vanishes unless $\mu \subseteq \lambda$. Consequently, $c^\lambda_{\mu,\nu}$ vanishes unless $\nu \subseteq \lambda$ as well (since $c^\lambda_{\mu,\nu} = c^\lambda_{\nu,\mu}$) and furthermore vanishes unless the equality $|\mu| + |\nu| = |\lambda|$ holds[126]. Altogether, we conclude that $c^\lambda_{\mu,\nu}$ vanishes unless $\mu, \nu \subseteq \lambda$ and $|\mu| + |\nu| = |\lambda|$.

Exercise 2.5.10. Show that any four partitions κ, λ, φ and ψ satisfy

$$\sum_{\rho\in\mathrm{Par}} c^\rho_{\kappa,\lambda} c^\rho_{\varphi,\psi} = \sum_{(\alpha,\beta,\gamma,\delta)\in\mathrm{Par}^4} c^\kappa_{\alpha,\gamma} c^\lambda_{\beta,\delta} c^\varphi_{\alpha,\beta} c^\psi_{\gamma,\delta}.$$

Exercise 2.5.11. (a) For any partition μ, prove that

$$\sum_{\lambda\in\mathrm{Par}} s_\lambda(\mathbf{x}) s_{\lambda/\mu}(\mathbf{y}) = s_\mu(\mathbf{x}) \cdot \prod_{i,j=1}^{\infty} (1 - x_i y_j)^{-1}$$

in the power series ring $\mathbf{k}[[\mathbf{x},\mathbf{y}]] = \mathbf{k}[[x_1, x_2, x_3, \ldots, y_1, y_2, y_3, \ldots]]$.

(b) Let α and β be two partitions. Show that

$$\sum_{\lambda\in\mathrm{Par}} s_{\lambda/\alpha}(\mathbf{x}) s_{\lambda/\beta}(\mathbf{y}) = \left(\sum_{\rho\in\mathrm{Par}} s_{\beta/\rho}(\mathbf{x}) s_{\alpha/\rho}(\mathbf{y}) \right) \cdot \prod_{i,j=1}^{\infty} (1 - x_i y_j)^{-1}$$

in the power series ring $\mathbf{k}[[\mathbf{x},\mathbf{y}]] = \mathbf{k}[[x_1, x_2, x_3, \ldots, y_1, y_2, y_3, \ldots]]$.

[**Hint:** For (b), expand the product

$$\prod_{i,j=1}^{\infty} (1 - x_i y_j)^{-1} \prod_{i,j=1}^{\infty} (1 - x_i w_j)^{-1} \prod_{i,j=1}^{\infty} (1 - z_i y_j)^{-1} \prod_{i,j=1}^{\infty} (1 - z_i w_j)^{-1}$$

in the power series ring $\mathbf{k}[[x_1, x_2, x_3, \ldots, y_1, y_2, y_3, \ldots, z_1, z_2, z_3, \ldots, w_1, w_2, w_3, \ldots]]$ in two ways: once by applying Theorem 2.5.1 to the two variable sets $(\mathbf{z}, \mathbf{x})$ and $(\mathbf{w}, \mathbf{y})$ and then using (2.3.3); once again by applying (2.5.1) to the two variable sets $\mathbf{z}$ and $\mathbf{w}$ and then applying Exercise 2.5.11(a) twice.]

The statement of Exercise 2.5.11(b) is known as the *skew Cauchy identity*, and appears in Sagan-Stanley [188, Cor. 6.12], Stanley [206, exercise 7.27(c)] and Macdonald [142, §I.5, example 26]; it seems to be due to Zelevinsky. It generalizes the statement of Exercise 2.5.11(a), which in turn is a generalization of Theorem 2.5.1.

[124] In the last equality, we removed the condition $\mu \subseteq \lambda$ on the addends of the sum; this does not change the value of the sum (because we have $s_{\lambda/\mu} = 0$ whenever we don't have $\mu \subseteq \lambda$).

[125] "Comparing coefficients" means applying Exercise 2.5.5(a) to $q_\lambda = s_\lambda$ in this case (although the base ring $\mathbf{k}$ is now replaced by $\mathbf{k}[[\mathbf{y}]]$, and the index μ is used instead of λ, since λ is already taken).

[126] In fact, this is clear when we don't have $\mu \subseteq \lambda$. When we do have $\mu \subseteq \lambda$, this follows from observing that $s_{\lambda/\mu} \in \Lambda_{|\lambda/\mu|}$ has zero coefficient before s_ν whenever $|\mu| + |\nu| \neq |\lambda|$.

Definition 2.5.12. Define the *Hall inner product* on Λ to be the $\mathbf{k}$-bilinear form $(\cdot,\cdot)$ which makes $\{s_\lambda\}$ an orthonormal basis, that is, $(s_\lambda, s_\nu) = \delta_{\lambda,\nu}$.

Exercise 2.5.13. (a) If n and m are two distinct nonnegative integers, and if $f \in \Lambda_n$ and $g \in \Lambda_m$, then show that $(f, g) = 0$.

(b) If $n \in \mathbb{N}$ and $f \in \Lambda_n$, then prove that $(h_n, f) = f(1)$ (where $f(1)$ is defined as in Exercise 2.1.2).

(c) Show that $(f, g) = (g, f)$ for all $f \in \Lambda$ and $g \in \Lambda$. (In other words, the Hall inner product is symmetric.)

The Hall inner product induces a $\mathbf{k}$-module homomorphism $\Lambda \to \Lambda^o$ (sending every $f \in \Lambda$ to the $\mathbf{k}$-linear map $\Lambda \to \mathbf{k},\ g \mapsto (f, g)$). This homomorphism is invertible (since the Hall inner product has an orthonormal basis), so that $\Lambda^o \cong \Lambda$ as $\mathbf{k}$-modules. But in fact, more can be said:

Corollary 2.5.14. *The isomorphism $\Lambda^o \cong \Lambda$ induced by the Hall inner product is an isomorphism of Hopf algebras.*

Proof. We have seen that the orthonormal basis $\{s_\lambda\}$ of Schur functions is *self-dual*, in the sense that its multiplication and comultiplication structure constants are the same. Thus the isomorphism $\Lambda^o \cong \Lambda$ induced by the Hall inner product is an isomorphism of bialgebras[127], and hence also a Hopf algebra isomorphism by Corollary 1.4.27. □

We next identify two other dual pairs of bases, by expanding the Cauchy product in two other ways.

Proposition 2.5.15. *One can also expand*

$$\prod_{i,j=1}^{\infty} (1 - x_i y_j)^{-1} = \sum_{\lambda \in \mathrm{Par}} h_\lambda(\mathbf{x}) m_\lambda(\mathbf{y}) = \sum_{\lambda \in \mathrm{Par}} z_\lambda^{-1} p_\lambda(\mathbf{x}) p_\lambda(\mathbf{y}) \tag{2.5.11}$$

[127] Here are some details on the proof:

Let $\gamma : \Lambda \to \Lambda^o$ be the $\mathbf{k}$-module isomorphism $\Lambda \to \Lambda^o$ induced by the Hall inner product. We want to show that γ is an isomorphism of bialgebras.

Let $\{s_\lambda^*\}$ be the basis of Λ^o dual to the basis $\{s_\lambda\}$ of Λ. Thus, for any partition λ, we have

$$\gamma(s_\lambda) = s_\lambda^* \tag{2.5.9}$$

(since any partition μ satisfies $(\gamma(s_\lambda))(s_\mu) = (s_\lambda, s_\mu) = \delta_{\lambda,\mu} = s_\lambda^*(s_\mu)$, and thus the two $\mathbf{k}$-linear maps $\gamma(s_\lambda) : \Lambda \to \mathbf{k}$ and $s_\lambda^* : \Lambda \to \mathbf{k}$ are equal to each other on the basis $\{s_\mu\}$ of Λ, which forces them to be identical).

The coproduct structure constants of the basis $\{s_\lambda^*\}$ of Λ^o equal the product structure constants of the basis $\{s_\lambda\}$ of Λ (according to our discussion of duals in Section 1.6). Since the latter are the Littlewood-Richardson numbers $c^\lambda_{\mu,\nu}$ (because of (2.5.6)), we thus conclude that the former are $c^\lambda_{\mu,\nu}$ as well. In other words, every $\lambda \in \mathrm{Par}$ satisfies

$$\Delta_{\Lambda^o} s_\lambda^* = \sum_{\mu,\nu} c^\lambda_{\mu,\nu} s_\mu^* \otimes s_\nu^* \tag{2.5.10}$$

(where the sum is over all pairs (μ,ν) of partitions). On the other hand, applying the map $\gamma \otimes \gamma : \Lambda \otimes \Lambda \to \Lambda^o \otimes \Lambda^o$ to the equality (2.5.7) yields

$$\begin{aligned}(\gamma \otimes \gamma)(\Delta(s_\lambda)) &= (\gamma \otimes \gamma)\left(\sum_{\mu,\nu} \hat{c}^\lambda_{\mu,\nu} s_\mu \otimes s_\nu\right) = \sum_{\mu,\nu} \underbrace{\hat{c}^\lambda_{\mu,\nu}}_{=c^\lambda_{\mu,\nu}} \underbrace{\gamma(s_\mu)}_{\substack{=s_\mu^* \\ \text{(by (2.5.9))}}} \otimes \underbrace{\gamma(s_\nu)}_{\substack{=s_\nu^* \\ \text{(by (2.5.9))}}} = \sum_{\mu,\nu} c^\lambda_{\mu,\nu} s_\mu^* \otimes s_\nu^* \\ &= \Delta_{\Lambda^o} \underbrace{s_\lambda^*}_{\substack{=\gamma(s_\lambda) \\ \text{(by (2.5.9))}}} \qquad \text{(by (2.5.10))} \\ &= \Delta_{\Lambda^o}(\gamma(s_\lambda))\end{aligned}$$

for each $\lambda \in \mathrm{Par}$. In other words, the two $\mathbf{k}$-linear maps $(\gamma \otimes \gamma) \circ \Delta$ and $\Delta_{\Lambda^o} \circ \gamma$ are equal to each other on each s_λ with $\lambda \in \mathrm{Par}$. Hence, these two maps must be identical (since the s_λ form a basis of Λ). Hence, $\Delta_{\Lambda^o} \circ \gamma = (\gamma \otimes \gamma) \circ \Delta$.

Our next goal is to show that $\epsilon_{\Lambda^o} \circ \gamma = \epsilon$. Indeed, each $\lambda \in \mathrm{Par}$ satisfies

$$\begin{aligned}(\epsilon_{\Lambda^o} \circ \gamma)(s_\lambda) &= \epsilon_{\Lambda^o}(\gamma(s_\lambda)) = (\gamma(s_\lambda))(1) \qquad \text{(by the definition of } \epsilon_{\Lambda^o}) \\ &= \left(s_\lambda, \underbrace{1}_{=s_\varnothing}\right) = (s_\lambda, s_\varnothing) = \delta_{\lambda,\varnothing} = \epsilon(s_\lambda).\end{aligned}$$

Hence, $\epsilon_{\Lambda^o} \circ \gamma = \epsilon$. Combined with $\Delta_{\Lambda^o} \circ \gamma = (\gamma \otimes \gamma) \circ \Delta$, this shows that γ is a $\mathbf{k}$-coalgebra homomorphism. Similar reasoning can be used to prove that γ is a $\mathbf{k}$-algebra homomorphism. Altogether, we thus conclude that γ is a bialgebra homomorphism. Since γ is a $\mathbf{k}$-module isomorphism, this yields that γ is an isomorphism of bialgebras. Qed.

where $z_\lambda := m_1! \cdot 1^{m_1} \cdot m_2! \cdot 2^{m_2} \cdots$ *if* λ *is written in multiplicative notation as* $\lambda = (1^{m_1}, 2^{m_2}, \ldots)$ *with multiplicity* m_i *for the part* i*. (Here, we assume that* $\mathbb{Q}$ *is a subring of* $\mathbf{k}$ *for the last equality.)*

Remark 2.5.16. It is relevant later (and explains the notation) that z_λ is the size of the $\mathfrak{S}_n$-centralizer subgroup for a permutation having cycle type[128] λ with $|\lambda| = n$. This is a classical (and fairly easy) result (see, e.g., [186, Prop. 1.1.1] or [206, Prop. 7.7.3] for a proof).

Proof of Proposition 2.5.15. For the first expansion, note that (2.2.18) shows

$$\begin{aligned}
\prod_{i,j=1}^{\infty} (1 - x_i y_j)^{-1} &= \prod_{j=1}^{\infty} \sum_{n \geq 0} h_n(\mathbf{x}) y_j^n \\
&= \sum_{\substack{\text{weak} \\ \text{compositions} \\ (n_1, n_2, \ldots)}} (h_{n_1}(\mathbf{x}) h_{n_2}(\mathbf{x}) \cdots)(y_1^{n_1} y_2^{n_2} \cdots) \\
&= \sum_{\lambda \in \mathrm{Par}} \sum_{\substack{\text{weak} \\ \text{compositions} \\ (n_1, n_2, \ldots) \\ \text{satisfying} \\ (n_1, n_2, \ldots) \in \mathfrak{S}_{(\infty)} \lambda}} \underbrace{(h_{n_1}(\mathbf{x}) h_{n_2}(\mathbf{x}) \cdots)}_{\substack{= h_\lambda(\mathbf{x}) \\ (\text{since } (n_1, n_2, \ldots) \in \mathfrak{S}_{(\infty)} \lambda)}} \underbrace{(y_1^{n_1} y_2^{n_2} \cdots)}_{= \mathbf{y}^{(n_1, n_2, \ldots)}} \\
&= \sum_{\lambda \in \mathrm{Par}} h_\lambda(\mathbf{x}) \underbrace{\sum_{\substack{\text{weak} \\ \text{compositions} \\ (n_1, n_2, \ldots) \\ \text{satisfying} \\ (n_1, n_2, \ldots) \in \mathfrak{S}_{(\infty)} \lambda}} \mathbf{y}^{(n_1, n_2, \ldots)}}_{= m_\lambda(\mathbf{y})} \\
&= \sum_{\lambda \in \mathrm{Par}} h_\lambda(\mathbf{x}) m_\lambda(\mathbf{y}).
\end{aligned}$$

For the second expansion (and for later use in the proof of Theorem 4.9.5) note that

$$\log H(t) = \log \prod_{i=1}^{\infty} (1 - x_i t)^{-1} = \sum_{i=1}^{\infty} - \log(1 - x_i t) = \sum_{i=1}^{\infty} \sum_{m=1}^{\infty} \frac{(x_i t)^m}{m} = \sum_{m=1}^{\infty} \frac{1}{m} p_m(\mathbf{x}) t^m, \tag{2.5.12}$$

so that taking $\frac{d}{dt}$ then shows that

$$P(t) := \sum_{m \geq 0} p_{m+1} t^m = \frac{H'(t)}{H(t)} = H'(t) E(-t). \tag{2.5.13}$$

A similar calculation shows that

$$\log \prod_{i,j=1}^{\infty} (1 - x_i y_j)^{-1} = \sum_{m=1}^{\infty} \frac{1}{m} p_m(\mathbf{x}) p_m(\mathbf{y}) \tag{2.5.14}$$

[128] If σ is a permutation of a finite set X, then the *cycle type* of σ is defined as the list of the lengths of all cycles of σ (that is, of all orbits of σ acting on X) written in decreasing order. This is clearly a partition of $|X|$. (Some other authors write it in increasing order instead, or treat it as a multiset.)

For instance, the permutation of the set $\{0, 3, 6, 9, 12\}$ that sends 0 to 3, 3 to 9, 6 to 6, 9 to 0, and 12 to 12 has cycle type $(3, 1, 1)$, since the cycles of this permutation have lengths 3, 1 and 1.

It is known that two permutations in $\mathfrak{S}_n$ have the same cycle type if and only if they are conjugate. Thus, for a given partition λ with $|\lambda| = n$, any two permutations in $\mathfrak{S}_n$ having cycle type λ are conjugate and therefore their $\mathfrak{S}_n$-centralizer subgroups have the same size.

and hence

$$\begin{aligned}
\prod_{i,j=1}^{\infty}(1-x_iy_j)^{-1} &= \exp\left(\sum_{m=1}^{\infty}\frac{1}{m}p_m(\mathbf{x})p_m(\mathbf{y})\right) = \prod_{m=1}^{\infty}\exp\left(\frac{1}{m}p_m(\mathbf{x})p_m(\mathbf{y})\right)\\
&= \prod_{m=1}^{\infty}\sum_{k=0}^{\infty}\frac{1}{k!}\left(\frac{1}{m}p_m(\mathbf{x})p_m(\mathbf{y})\right)^k = \sum_{\substack{\text{weak compositions}\\(k_1,k_2,k_3,\ldots)}}\prod_{m=1}^{\infty}\left(\frac{1}{k_m!}\left(\frac{1}{m}p_m(\mathbf{x})p_m(\mathbf{y})\right)^{k_m}\right)\\
&\qquad\text{(by the product rule)}\\
&= \sum_{\substack{\text{weak compositions}\\(k_1,k_2,k_3,\ldots)}}\prod_{m=1}^{\infty}\frac{(p_m(\mathbf{x})p_m(\mathbf{y}))^{k_m}}{k_m!m^{k_m}} = \sum_{\substack{\text{weak compositions}\\(k_1,k_2,k_3,\ldots)}}\frac{\prod_{m=1}^{\infty}(p_m(\mathbf{x}))^{k_m}\prod_{m=1}^{\infty}(p_m(\mathbf{y}))^{k_m}}{\prod_{m=1}^{\infty}(k_m!m^{k_m})}\\
&= \sum_{\substack{\text{weak compositions}\\(k_1,k_2,k_3,\ldots)}}\frac{p_{\left(1^{k_1}2^{k_2}3^{k_3}\cdots\right)}(\mathbf{x})\,p_{\left(1^{k_1}2^{k_2}3^{k_3}\cdots\right)}(\mathbf{y})}{z_{\left(1^{k_1}2^{k_2}3^{k_3}\cdots\right)}} = \sum_{\lambda\in\mathrm{Par}}\frac{p_\lambda(\mathbf{x})p_\lambda(\mathbf{y})}{z_\lambda}
\end{aligned}$$

due to the fact that every partition can be uniquely written in the form $\left(1^{k_1}2^{k_2}3^{k_3}\cdots\right)$ with $(k_1,k_2,k_3,\ldots)$ a weak composition. □

Corollary 2.5.17. (a) *With respect to the Hall inner product on Λ, one also has dual bases $\{h_\lambda\}$ and $\{m_\lambda\}$.*

(b) *If $\mathbb{Q}$ is a subring of $\mathbf{k}$, then $\{p_\lambda\}$ and $\left\{z_\lambda^{-1}p_\lambda\right\}$ are also dual bases with respect to the Hall inner product on Λ.*

(c) *If $\mathbb{R}$ is a subring of $\mathbf{k}$, then $\left\{\dfrac{p_\lambda}{\sqrt{z_\lambda}}\right\}$ is an orthonormal basis of Λ with respect to the Hall inner product.*

Proof. Since (2.5.1) and (2.5.11) showed

$$\prod_{i,j=1}^{\infty}(1-x_iy_j)^{-1} = \sum_{\lambda\in\mathrm{Par}}s_\lambda(\mathbf{x})s_\lambda(\mathbf{y}) = \sum_{\lambda\in\mathrm{Par}}h_\lambda(\mathbf{x})m_\lambda(\mathbf{y}) = \sum_{\lambda\in\mathrm{Par}}p_\lambda(\mathbf{x})z_\lambda^{-1}p_\lambda(\mathbf{y}) = \sum_{\lambda\in\mathrm{Par}}\frac{p_\lambda(\mathbf{x})}{\sqrt{z_\lambda}}\frac{p_\lambda(\mathbf{y})}{\sqrt{z_\lambda}},$$

it suffices to show that any pair of graded bases[129] $\{u_\lambda\},\{v_\lambda\}$ of Λ having

$$\sum_{\lambda\in\mathrm{Par}}s_\lambda(\mathbf{x})s_\lambda(\mathbf{y}) = \sum_{\lambda\in\mathrm{Par}}u_\lambda(\mathbf{x})v_\lambda(\mathbf{y})$$

will be dual with respect to $(\cdot,\cdot)$. To show this, consider such a pair of graded bases. Write transition matrices $A=(a_{\nu,\lambda})_{(\nu,\lambda)\in\mathrm{Par}\times\mathrm{Par}}$ and $B=(b_{\nu,\lambda})_{(\nu,\lambda)\in\mathrm{Par}\times\mathrm{Par}}$ uniquely expressing

$$u_\lambda = \sum_\nu a_{\nu,\lambda}s_\nu, \tag{2.5.15}$$

$$v_\lambda = \sum_\nu b_{\nu,\lambda}s_\nu. \tag{2.5.16}$$

Recall that $\mathrm{Par}=\bigsqcup_{r\in\mathbb{N}}\mathrm{Par}_r$. Hence, we can view A as a block matrix, where the blocks are indexed by pairs of nonnegative integers, and the (r,s)-th block is $(a_{\nu,\lambda})_{(\nu,\lambda)\in\mathrm{Par}_r\times\mathrm{Par}_s}$. For reasons of homogeneity[130], we have $a_{\nu,\lambda}=0$ for any $(\nu,\lambda)\in\mathrm{Par}^2$ satisfying $|\nu|\neq|\lambda|$. Therefore, the (r,s)-th block of A is zero whenever $r\neq s$. In other words, the block matrix A is block-diagonal. Similarly, B can be viewed as a block-diagonal matrix. The diagonal blocks of A and B are finite square matrices (since Par_r is a finite set for each $r\in\mathbb{N}$); therefore, products such as A^tB, B^tA and AB^t are well-defined (since all sums involved in their definition have only finitely many nonzero addends) and subject to the law of associativity. Moreover, the matrix A is

[129] See Definition 1.3.21 for the concept of a "graded basis", and recall our convention that a graded basis of Λ is tacitly assumed to have its indexing set Par partitioned into $\mathrm{Par}_0,\mathrm{Par}_1,\mathrm{Par}_2,\ldots$. Thus, a graded basis of Λ means a basis $\{w_\lambda\}_{\lambda\in\mathrm{Par}}$ of the $\mathbf{k}$-module Λ (indexed by the partitions $\lambda\in\mathrm{Par}$) with the property that, for every $n\in\mathbb{N}$, the subfamily $\{w_\lambda\}_{\lambda\in\mathrm{Par}_n}$ is a basis of the $\mathbf{k}$-module Λ_n.

[130] More precisely: The power series u_λ is homogeneous of degree $|\lambda|$, and the power series s_ν is homogeneous of degree $|\nu|$.

invertible (being a transition matrix between two bases), and its inverse is again block-diagonal (because A is block-diagonal).

The equalities (2.5.15) and (2.5.16) show that $(u_\alpha, v_\beta) = \sum_\nu a_{\nu,\alpha} b_{\nu,\beta}$ (by the orthonormality of the s_λ). Hence, we want to prove that $\sum_\nu a_{\nu,\alpha} b_{\nu,\beta} = \delta_{\alpha,\beta}$. In other words, we want to prove that $A^t B = I$, that is, $B^{-1} = A^t$. On the other hand, one has

$$\sum_\lambda s_\lambda(\mathbf{x}) s_\lambda(\mathbf{y}) = \sum_\lambda u_\lambda(\mathbf{x}) v_\lambda(\mathbf{y}) = \sum_\lambda \sum_\nu a_{\nu,\lambda} s_\nu(\mathbf{x}) \sum_\rho b_{\rho,\lambda} s_\rho(\mathbf{y}).$$

Comparing coefficients[131] of $s_\nu(\mathbf{x}) s_\rho(\mathbf{y})$ forces $\sum_\lambda a_{\nu,\lambda} b_{\rho,\lambda} = \delta_{\nu,\rho}$, or in other words, $AB^t = I$. Since A is invertible, this yields $B^t A = I$, and hence $A^t B = I$, as desired.[132] □

Corollary 2.5.17 is a known and fundamental fact[133]. However, our definition of the Hall inner product is unusual; most authors (e.g., Macdonald in [142, §I.4, (4.5)], Hazewinkel/Gubareni/Kirichenko in [93, Def. 4.1.21], and Stanley in [206, (7.30)]) *define* the Hall inner product as the bilinear form satisfying $(h_\lambda, m_\mu) = \delta_{\lambda,\mu}$ (or, alternatively, $(m_\lambda, h_\mu) = \delta_{\lambda,\mu}$), and only later prove that the basis $\{s_\lambda\}$ is orthonormal with respect to this scalar product. (Of course, the fact that this definition is equivalent to our Definition 2.5.12 follows either from this orthonormality, or from our Corollary 2.5.17(a).)

The tactic applied in the proof of Corollary 2.5.17 can not only be used to show that certain bases of Λ are dual, but also, with a little help from linear algebra over rings (Exercise 2.5.18), it can be strengthened to show that certain families of symmetric functions are bases to begin with, as we will see in Exercise 2.5.19 and Exercise 2.5.20.

Exercise 2.5.18. (a) Prove that if an endomorphism of a finitely generated $\mathbf{k}$-module is surjective, then this endomorphism is a $\mathbf{k}$-module isomorphism.

(b) Let A be a finite free $\mathbf{k}$-module with finite basis $(\gamma_i)_{i\in I}$. Let $(\beta_i)_{i\in I}$ be a family of elements of A which spans the $\mathbf{k}$-module A. Prove that $(\beta_i)_{i\in I}$ is a $\mathbf{k}$-basis of A.

Exercise 2.5.19. For each partition λ, let v_λ be an element of $\Lambda_{|\lambda|}$. Assume that the family $(v_\lambda)_{\lambda\in\mathrm{Par}}$ spans the $\mathbf{k}$-module Λ. Prove that the family $(v_\lambda)_{\lambda\in\mathrm{Par}}$ is a graded basis of the graded $\mathbf{k}$-module Λ.

Exercise 2.5.20. (a) Assume that for every partition λ, two homogeneous elements u_λ and v_λ of Λ, both having degree $|\lambda|$, are given. Assume further that

$$\sum_{\lambda\in\mathrm{Par}} s_\lambda(\mathbf{x})\, s_\lambda(\mathbf{y}) = \sum_{\lambda\in\mathrm{Par}} u_\lambda(\mathbf{x})\, v_\lambda(\mathbf{y})$$

in $\mathbf{k}[[\mathbf{x},\mathbf{y}]] = \mathbf{k}[[x_1,x_2,x_3,\ldots,y_1,y_2,y_3,\ldots]]$. Show that $(u_\lambda)_{\lambda\in\mathrm{Par}}$ and $(v_\lambda)_{\lambda\in\mathrm{Par}}$ are $\mathbf{k}$-bases of Λ, and actually are dual bases with respect to the Hall inner product on Λ.

(b) Use this to give a new proof of the fact that $(h_\lambda)_{\lambda\in\mathrm{Par}}$ is a $\mathbf{k}$-basis of Λ.

Exercise 2.5.21. Prove that $\sum_{m\geq 0} p_{m+1} t^m = \dfrac{H'(t)}{H(t)}$. (This was proven in (2.5.13) in the case when $\mathbb{Q}$ is a subring of $\mathbf{k}$, but here we make no requirements on $\mathbf{k}$.)

The following exercises give some useful criteria for algebraic independence of families of symmetric functions:

[131] Comparing coefficients is legitimate because if a power series $f \in \mathbf{k}[[\mathbf{x},\mathbf{y}]]$ is written in the form $f = \sum_{(\nu,\rho)\in\mathrm{Par}^2} a_{\rho,\nu} s_\nu(\mathbf{x})\, s_\rho(\mathbf{y})$ for some coefficients $a_{\rho,\nu} \in \mathbf{k}$, then these coefficients $a_{\rho,\nu}$ are uniquely determined by f. This is just a restatement of Exercise 2.5.5(b).

[132] In our argument above, we have obtained the invertibility of A from the fact that A is a transition matrix between two bases. Here is an alternative way to prove that A is invertible:

Recall that A and B^t are block-diagonal matrices. Hence, the equality $AB^t = I$ rewrites as $A_{r,r}\left(B^t\right)_{r,r} = I$ for all $r \in \mathbb{N}$, where we are using the notation $C_{r,s}$ for the (r,s)-th block of a block matrix C. But this shows that each diagonal block $A_{r,r}$ of A is right-invertible. Therefore, each diagonal block $A_{r,r}$ of A is invertible (because $A_{r,r}$ is a square matrix of finite size, and such matrices are always invertible when they are right-invertible). Consequently, the block-diagonal matrix A is invertible, and its inverse is again a block-diagonal matrix (whose diagonal blocks are the inverses of the $A_{r,r}$).

[133] For example, Corollary 2.5.17(a) appears in [126, Corollary 3.3] (though the definition of Schur functions in [126] is different from ours; we will meet this alternative definition later on), and parts (b) and (c) of Corollary 2.5.17 are equivalent to [142, §I.4, (4.7)] (though Macdonald defines the Hall inner product using Corollary 2.5.17(a)).

Exercise 2.5.22. Let $v_1, v_2, v_3, \ldots$ be elements of Λ. Assume that $v_n \in \Lambda_n$ for each positive integer n. Assume further that $v_1, v_2, v_3, \ldots$ generate the $\mathbf{k}$-algebra Λ. Then:

(a) Prove that $v_1, v_2, v_3, \ldots$ are algebraically independent over $\mathbf{k}$.

(b) For every partition λ, define an element $v_\lambda \in \Lambda$ by $v_\lambda = v_{\lambda_1} v_{\lambda_2} \cdots v_{\lambda_{\ell(\lambda)}}$. Prove that the family $(v_\lambda)_{\lambda \in \operatorname{Par}}$ is a graded basis of the graded $\mathbf{k}$-module Λ.

Exercise 2.5.23. For each partition λ, let $a_\lambda \in \mathbf{k}$. Assume that the element $a_{(n)} \in \mathbf{k}$ is invertible for each positive integer n. Let $v_1, v_2, v_3, \ldots$ be elements of Λ such that each positive integer n satisfies $v_n = \sum_{\lambda \in \operatorname{Par}_n} a_\lambda h_\lambda$. Prove that the elements $v_1, v_2, v_3, \ldots$ generate the $\mathbf{k}$-algebra Λ and are algebraically independent over $\mathbf{k}$.

Exercise 2.5.24. Let $v_1, v_2, v_3, \ldots$ be elements of Λ. Assume that $v_n \in \Lambda_n$ for each positive integer n. Assume further that $(p_n, v_n) \in \mathbf{k}$ is invertible for each positive integer n. Prove that the elements $v_1, v_2, v_3, \ldots$ generate the $\mathbf{k}$-algebra Λ and are algebraically independent over $\mathbf{k}$.

Exercise 2.5.25. Let $f \in \Lambda$, and let β be a weak composition. Let $\mu \in \operatorname{Par}$ be the partition consisting of the nonzero entries of β (sorted in decreasing order).[134] Prove that

$$(f, h_\mu) = (h_\mu, f) = \left(\text{the coefficient of } \mathbf{x}^\beta \text{ in } f\right).$$

Exercise 2.5.26. Assume that $\mathbb{Q}$ is a subring of $\mathbf{k}$. Define a positive integer z_λ for each $\lambda \in \operatorname{Par}$ as in Proposition 2.5.15. Prove that every $n \in \mathbb{N}$ satisfies the two equalities

$$h_n = \sum_{\lambda \in \operatorname{Par}_n} z_\lambda^{-1} p_\lambda \tag{2.5.17}$$

and

$$e_n = \sum_{\lambda \in \operatorname{Par}_n} (-1)^{|\lambda| - \ell(\lambda)} z_\lambda^{-1} p_\lambda. \tag{2.5.18}$$

2.6. **Bialternants, Littlewood-Richardson: Stembridge's concise proof.** There is a more natural way in which Schur functions arise as a $\mathbf{k}$-basis for Λ, coming from consideration of polynomials in a finite variable set, and the relation between those which are symmetric and those which are *alternating*.

For the remainder of this section, fix a nonnegative integer n, and let $\mathbf{x} = (x_1, \ldots, x_n)$ be a finite variable set. This means that $s_{\lambda/\mu} = s_{\lambda/\mu}(\mathbf{x}) = \sum_T \mathbf{x}^{\operatorname{cont}(T)}$ is a generating function for column-strict tableaux T as in Definition 2.3.1, but with the extra condition that T have entries in $\{1, 2, \ldots, n\}$.[135] As a consequence, $s_{\lambda/\mu}$ is a polynomial in $\mathbf{k}[x_1, x_2, \ldots, x_n]$ (not just a power series), since there are only finitely many column-strict tableaux T of shape λ/μ having all their entries in $\{1, 2, \ldots, n\}$. We will assume without further mention that all partitions appearing in the section have at most n parts.

Definition 2.6.1. Let $\mathbf{k}$ be the ring $\mathbb{Z}$ or a field of characteristic not equal to 2. (We require this to avoid certain annoyances in the discussion of alternating polynomials in characteristic 2.)

Say that a polynomial $f(\mathbf{x}) = f(x_1, \ldots, x_n)$ is *alternating* if for every permutation w in $\mathfrak{S}_n$ one has that

$$(wf)(\mathbf{x}) = f(x_{w(1)}, \ldots, x_{w(n)}) = \operatorname{sgn}(w) f(\mathbf{x}).$$

Let $\Lambda^{\operatorname{sgn}} \subset \mathbf{k}[x_1, \ldots, x_n]$ denote the subset of alternating polynomials[136].

As with Λ and its monomial basis $\{m_\lambda\}$, there is an obvious $\mathbf{k}$-basis for $\Lambda^{\operatorname{sgn}}$, coming from the fact that a polynomial $f = \sum_\alpha c_\alpha \mathbf{x}^\alpha$ is alternating if and only if $c_{w(\alpha)} = \operatorname{sgn}(w) c_\alpha$ for every w in $\mathfrak{S}_n$ and every $\alpha \in \mathbb{N}^n$. This means that every alternating f is a $\mathbf{k}$-linear combination of the following elements.

Definition 2.6.2. For $\alpha = (\alpha_1, \ldots, \alpha_n)$ in $\mathbb{N}^n$, define the *alternant*

$$a_\alpha := \sum_{w \in \mathfrak{S}_n} \operatorname{sgn}(w) w(\mathbf{x}^\alpha) = \det \begin{bmatrix} x_1^{\alpha_1} & \cdots & x_1^{\alpha_n} \\ x_2^{\alpha_1} & \cdots & x_2^{\alpha_n} \\ \vdots & \ddots & \vdots \\ x_n^{\alpha_1} & \cdots & x_n^{\alpha_n} \end{bmatrix}.$$

[134] For example, if $\beta = (1, 0, 3, 1, 2, 3, 0, 0, 0, \ldots)$, then $\mu = (3, 3, 2, 1, 1)$.

[135] See Exercise 2.3.8(a) for this.

[136] When $\mathbf{k}$ has characteristic 2 (or, more generally, is an arbitrary commutative ring), it is probably best to define the alternating polynomials $\Lambda_{\mathbf{k}}^{\operatorname{sgn}}$ as the $\mathbf{k}$-submodule $\Lambda^{\operatorname{sgn}} \otimes_{\mathbb{Z}} \mathbf{k}$ of $\mathbb{Z}[x_1, \ldots, x_n] \otimes_{\mathbb{Z}} \mathbf{k} \cong \mathbf{k}[x_1, \ldots, x_n]$.

Example 2.6.3. One has

$$a_{(1,5,0)} = x_1^1x_2^5x_3^0 - x_1^5x_2^1x_3^0 - x_1^1x_2^0x_3^5 - x_1^0x_2^5x_3^1 + x_1^0x_2^1x_3^5 + x_1^5x_2^0x_3^1 = -a_{(5,1,0)}.$$

Similarly, $a_{w(\alpha)} = \operatorname{sgn}(w)a_\alpha$ for every $w \in \mathfrak{S}_n$ and every $\alpha \in \mathbb{N}^n$.

Meanwhile, $a_{(5,2,2)} = 0$ since the transposition $t = \binom{123}{132}$ fixes $(5,2,2)$ and hence

$$a_{(5,2,2)} = t(a_{(5,2,2)}) = \operatorname{sgn}(t)a_{(5,2,2)} = -a_{(5,2,2)}.$$

[137] Alternatively, $a_{(5,2,2)} = 0$ as it is a determinant of a matrix with two equal columns. Similarly, $a_\alpha = 0$ for every n-tuple $\alpha \in \mathbb{N}^n$ having two equal entries.

This example illustrates that, for a $\mathbf{k}$-basis for Λ^{sgn}, one can restrict attention to alternants a_α in which α is a *strict partition*, i.e., in which α satisfies $\alpha_1 > \alpha_2 > \cdots > \alpha_n$. One can therefore uniquely express $\alpha = \lambda + \rho$, where λ is a (weak) partition $\lambda_1 \geq \cdots \geq \lambda_n \geq 0$ and where $\rho := (n-1, n-2, \ldots, 2, 1, 0)$ is sometimes called the *staircase partition*[138]. For example $\alpha = (5,1,0) = (3,0,0) + (2,1,0) = \lambda + \rho$.

Proposition 2.6.4. *Let $\mathbf{k}$ be the ring $\mathbb{Z}$ or a field of characteristic not equal to 2.*

The alternants $\{a_{\lambda+\rho}\}$ as λ runs through the partitions with at most n parts form a $\mathbf{k}$-basis for Λ^{sgn}. In addition, the bialternants $\{\frac{a_{\lambda+\rho}}{a_\rho}\}$ *as λ runs through the same set form a $\mathbf{k}$-basis for $\Lambda(x_1, \ldots, x_n) = \mathbf{k}[x_1, \ldots, x_n]^{\mathfrak{S}_n}$.*

Proof. The first assertion should be clear from our previous discussion: the alternants $\{a_{\lambda+\rho}\}$ span Λ^{sgn} by definition, and they are $\mathbf{k}$-linearly independent because they are supported on disjoint sets of monomials $\mathbf{x}^\alpha$.

The second assertion follows from the first, after proving the following **Claim**: $f(\mathbf{x})$ lies in Λ^{sgn} if and only if $f(\mathbf{x}) = a_\rho \cdot g(\mathbf{x})$ where $g(\mathbf{x})$ lies in $\mathbf{k}[\mathbf{x}]^{\mathfrak{S}_n}$ and where

$$a_\rho = \det(x_i^{n-j})_{i,j=1,2,\ldots,n} = \prod_{1 \leq i < j \leq n} (x_i - x_j)$$

is the *Vandermonde determinant/product*. In other words,

$$\Lambda^{\mathrm{sgn}} = a_\rho \cdot \mathbf{k}[\mathbf{x}]^{\mathfrak{S}_n}$$

is a free $\mathbf{k}[\mathbf{x}]^{\mathfrak{S}_n}$-module of rank 1, with a_ρ as its $\mathbf{k}[\mathbf{x}]^{\mathfrak{S}_n}$-basis element.

To see the Claim, first note the inclusion

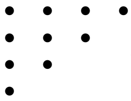
$$\Lambda^{\mathrm{sgn}} \supset a_\rho \cdot \mathbf{k}[\mathbf{x}]^{\mathfrak{S}_n}$$

since the product of a symmetric polynomial and an alternating polynomial is an alternating polynomial. For the reverse inclusion, note that since an alternating polynomial $f(\mathbf{x})$ changes sign whenever one exchanges two distinct variables x_i, x_j, it must vanish upon setting $x_i = x_j$, and therefore be divisible by $x_i - x_j$, so divisible by the entire product $\prod_{1 \leq i < j \leq n}(x_i - x_j) = a_\rho$. But then the quotient $g(\mathbf{x}) = \frac{f(\mathbf{x})}{a_\rho}$ is symmetric, as it is a quotient of two alternating polynomials. □

Let us now return to the general setting, where $\mathbf{k}$ is an arbitrary commutative ring. We are not requiring that the assumptions of Proposition 2.6.4 be valid; we can still study the a_α of Definition 2.6.2, but we cannot use Proposition 2.6.4 anymore. We will show that the fraction $\frac{a_{\lambda+\rho}}{a_\rho}$ is nevertheless a well-defined polynomial in $\Lambda(x_1, \ldots, x_n)$ whenever λ is a partition[139], and in fact equals the Schur function $s_\lambda(\mathbf{x})$. As a consequence, the mysterious bialternant basis $\{\frac{a_{\lambda+\rho}}{a_\rho}\}$ of $\Lambda(x_1, \ldots, x_n)$ defined in Proposition 2.6.4 still

[137] One subtlety should be addressed: We want to prove that $a_{(5,2,2)} = 0$ in $\mathbf{k}[x_1, \ldots, x_n]$ for every commutative ring $\mathbf{k}$. It is clearly enough to prove that $a_{(5,2,2)} = 0$ in $\mathbb{Z}[x_1, \ldots, x_n]$. Since 2 is not a zero-divisor in $\mathbb{Z}[x_1, \ldots, x_n]$, we can achieve this by showing that $a_{(5,2,2)} = -a_{(5,2,2)}$. We would not be able to make this argument directly over an arbitrary commutative ring $\mathbf{k}$.

[138] The name is owed to its Ferrers shape. For instance, if $n = 5$, then the Ferrers diagram of ρ (represented using dots) has the form



.

[139] This can also be deduced by base change from the $\mathbf{k} = \mathbb{Z}$ case of Proposition 2.6.4.

exists in the general setting, and is plainly the Schur functions $\{s_\lambda(\mathbf{x})\}$. Stembridge [210] noted that one could give a remarkably concise proof of an even stronger assertion, which simultaneously gives one of the standard combinatorial interpretations for the Littlewood-Richardson coefficients $c^\lambda_{\mu,\nu}$. For the purposes of stating it, we introduce for a tableau T the notation $T|_{\mathrm{cols}\geq j}$ (resp. $T|_{\mathrm{cols}\leq j}$) to indicate the subtableau which is the restriction of T to the union of its columns $j, j+1, j+2, \ldots$ (resp. columns $1, 2, \ldots, j$).

Example 2.6.5. If $T = \begin{array}{cccc} & & 1 & 2 \\ & 2 & 2 & 3 \\ 3 & 5 & & \end{array}$, then

$$T|_{\mathrm{cols}\geq 3} = \begin{array}{cc} 1 & 2 \\ 2 & 3 \end{array} \qquad \text{and} \qquad T|_{\mathrm{cols}\leq 2} = \begin{array}{cc} & \\ & 2 \\ 3 & 5 \end{array}$$

(note that $T|_{\mathrm{cols}\leq 2}$ has an empty first row).

Theorem 2.6.6. *For partitions λ, μ, ν with $\mu \subseteq \lambda$, one has*[140]

$$a_{\nu+\rho} s_{\lambda/\mu} = \sum_T a_{\nu+\mathrm{cont}(T)+\rho}$$

where T runs through all column-strict tableaux with entries in $\{1, 2, \ldots, n\}$ of shape λ/μ with the property that for each $j = 1, 2, 3, \ldots$, the weak composition $\nu + \mathrm{cont}(T|_{\mathrm{cols}\geq j})$ is a partition.

Before proving Theorem 2.6.6, let us see some of its consequences.

Corollary 2.6.7. *For any partition λ, we have*[141]

$$s_\lambda(\mathbf{x}) = \frac{a_{\lambda+\rho}}{a_\rho}.$$

Proof. Fix a partition λ. Take $\nu = \mu = \varnothing$ in Theorem 2.6.6. Note that there is only one column-strict tableau T of shape λ such that each $\mathrm{cont}(T|_{\mathrm{cols}\geq j})$ is a partition, namely the tableau having every entry in row i equal to i:

$$\begin{array}{ccccc} 1 & 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & & \\ 3 & 3 & 3 & & \\ 4 & 4 & & & \end{array}$$

[142]. Furthermore, this T has $\mathrm{cont}(T) = \lambda$, so the theorem says $a_\rho s_\lambda = a_{\lambda+\rho}$. □

[140] Again, we can drop the requirement that $\mu \subseteq \lambda$, provided that we understand that there are no column-strict tableaux of shape λ/μ unless $\mu \subseteq \lambda$.

[141] Notice that division by a_ρ is unambiguous in the ring $\mathbf{k}[x_1, \ldots, x_n]$, since a_ρ is not a zero-divisor (in fact, $a_\rho = \prod_{1\leq i<j\leq n}(x_i - x_j)$ is the product of the binomials $x_i - x_j$, none of which is a zero-divisor).

[142] *Proof.* It is clear that the tableau having every entry in row i equal to i indeed satisfies the condition that each $\mathrm{cont}(T|_{\mathrm{cols}\geq j})$ is a partition. It remains to show that it is the only column-strict tableau (of shape λ) satisfying this condition.

Let T be a column-strict tableau of shape λ satisfying the condition that each $\mathrm{cont}(T|_{\mathrm{cols}\geq j})$ is a partition. We must show that for each i, every entry in row i of T is equal to i. Assume the contrary. Thus, there exists some i such that row i of T contains an entry distinct from i. Consider the smallest such i. Hence, rows $1, 2, \ldots, i-1$ of T are filled with entries $1, 2, \ldots, i-1$, whereas row i has some entry distinct from i. Choose some j such that the j-th entry of row i of T is distinct from i. This entry cannot be smaller than i (since it has $i-1$ entries above it in its column, and the entries of T increase strictly down columns); thus, it has to be larger than i. Therefore, all entries in rows $i, i+1, i+2, \ldots$ of $T|_{\mathrm{cols}\geq j}$ are larger than i as well (since they lie southeast of this entry). Hence, each entry of $T|_{\mathrm{cols}\geq j}$ is either smaller than i (if it is in one of rows $1, 2, \ldots, i-1$) or larger than i (if it is in row i or further down). Thus, i is not an entry of $T|_{\mathrm{cols}\geq j}$. In other words, $\mathrm{cont}_i(T|_{\mathrm{cols}\geq j}) = 0$. Since $\mathrm{cont}(T|_{\mathrm{cols}\geq j})$ is a partition, we thus conclude that $\mathrm{cont}_k(T|_{\mathrm{cols}\geq j}) = 0$ for all $k > i$. In other words, $T|_{\mathrm{cols}\geq j}$ has no entries larger than i. But this contradicts the fact that the j-th entry of row i of T is larger than i. This contradiction completes our proof.

Example 2.6.8. For $n = 2$, so that $\rho = (1, 0)$, if we take $\lambda = (4, 2)$, then one has

$$\begin{aligned}
\frac{a_{\lambda+\rho}}{a_\rho} &= \frac{a_{(4,2)+(1,0)}}{a_{(1,0)}} = \frac{a_{(5,2)}}{a_{(1,0)}} \\
&= \frac{x_1^5 x_2^2 - x_1^2 x_2^5}{x_1 - x_2} \\
&= x_1^4 x_2^2 + x_1^3 x_2^3 + x_1^2 x_2^4 \\
&= \mathbf{x}^{\operatorname{cont}\begin{pmatrix}1111\\22\end{pmatrix}} + \mathbf{x}^{\operatorname{cont}\begin{pmatrix}1112\\22\end{pmatrix}} + \mathbf{x}^{\operatorname{cont}\begin{pmatrix}1122\\22\end{pmatrix}} \\
&= s_{(4,2)} = s_\lambda .
\end{aligned}$$

Some authors use the equality in Corollary 2.6.7 to *define* the Schur polynomial $s_\lambda (x_1, x_2, \ldots, x_n)$ in n variables; this definition, however, has the drawback of not generalizing easily to infinitely many variables or to skew Schur functions[143].

Next divide through by a_ρ on both sides of Theorem 2.6.6 (and use Corollary 2.6.7) to give the following.

Corollary 2.6.9. *For partitions λ, μ, ν having at most n parts, one has*

$$s_\nu s_{\lambda/\mu} = \sum_T s_{\nu+\operatorname{cont}(T)} \tag{2.6.1}$$

where T runs through the same set as in Theorem 2.6.6. In particular, taking $\nu = \varnothing$, we obtain

$$s_{\lambda/\mu} = \sum_T s_{\operatorname{cont}(T)} \tag{2.6.2}$$

where in the sum T runs through all column-strict tableaux of shape λ/μ for which each $\operatorname{cont}(T|_{\text{cols}\geq j})$ *is a partition.*

Proof of Theorem 2.6.6. Start by rewriting the left side of the theorem:

$$\begin{aligned}
a_{\nu+\rho} s_{\lambda/\mu} &= \sum_{w\in\mathfrak{S}_n} \operatorname{sgn}(w) \mathbf{x}^{w(\nu+\rho)} s_{\lambda/\mu} = \sum_{w\in\mathfrak{S}_n} \operatorname{sgn}(w) \mathbf{x}^{w(\nu+\rho)} w(s_{\lambda/\mu}) \\
&\qquad \big(\text{since } w(s_{\lambda/\mu}) = s_{\lambda/\mu} \text{ for any } w \in \mathfrak{S}_n\big) \\
&= \sum_{w\in\mathfrak{S}_n} \operatorname{sgn}(w) \mathbf{x}^{w(\nu+\rho)} \sum_{\substack{\text{column-strict } T \\ \text{of shape } \lambda/\mu}} \mathbf{x}^{w(\operatorname{cont}(T))} \\
&= \sum_{\substack{\text{column-strict } T \\ \text{of shape } \lambda/\mu}} \sum_{w\in\mathfrak{S}_n} \operatorname{sgn}(w) \mathbf{x}^{w(\nu+\operatorname{cont}(T)+\rho)} \\
&= \sum_{\substack{\text{column-strict } T \\ \text{of shape } \lambda/\mu}} a_{\nu+\operatorname{cont}(T)+\rho} .
\end{aligned}$$

We wish to cancel out all the summands indexed by column-strict tableaux T which fail any of the conditions that $\nu + \operatorname{cont}(T|_{\text{cols}\geq j})$ be a partition. Given such a T, find the maximal j for which it fails this condition[144], and then find the minimal k for which

$$\nu_k + \operatorname{cont}_k(T|_{\text{cols}\geq j}) < \nu_{k+1} + \operatorname{cont}_{k+1}(T|_{\text{cols}\geq j}).$$

Maximality of j forces

$$\nu_k + \operatorname{cont}_k(T|_{\text{cols}\geq j+1}) \geq \nu_{k+1} + \operatorname{cont}_{k+1}(T|_{\text{cols}\geq j+1}).$$

[143] With some effort, it is possible to use Corollary 2.6.7 in order to define the Schur function s_λ in infinitely many variables. Indeed, one can define this Schur function as the unique element of Λ whose evaluation at $(x_1, x_2, \ldots, x_n)$ equals $\frac{a_{\lambda+\rho}}{a_\rho}$ for every $n \in \mathbb{N}$. If one wants to use such a definition, however, one needs to check that such an element exists. This is the approach to defining s_λ taken in [126, Definition 1.4.2] and in [142, §I.3].

[144] Such a j exists because $\nu + \operatorname{cont}(T|_{\text{cols}\geq j})$ is a partition for all sufficiently high j (in fact, ν itself is a partition).

Since column-strictness implies that column j of T can contain at most one occurrence of k or of $k+1$ (or neither or both), the previous two inequalities imply that column j must contain an occurrence of $k+1$ and no occurrence of k, so that

$$\nu_k + \operatorname{cont}_k(T|_{\text{cols}\geq j}) + 1 = \nu_{k+1} + \operatorname{cont}_{k+1}(T|_{\text{cols}\geq j}).$$

This implies that the adjacent transposition $t_{k,k+1}$ swapping k and $k+1$ fixes the vector $\nu+\operatorname{cont}(T|_{\text{cols}\geq j})+\rho$.

Now create a new tableau T^* from T by applying the Bender-Knuth involution (from the proof of Proposition 2.2.4) on letters $k, k+1$, but *only to columns* $1, 2, \ldots, j-1$ of T, leaving columns $j, j+1, j+2, \ldots$ unchanged.[145] One should check that T^* is still column-strict, but this holds because column j of T has no occurrences of letter k. Note that

$$t_{k,k+1} \operatorname{cont}(T|_{\text{cols}\leq j-1}) = \operatorname{cont}(T^*|_{\text{cols}\leq j-1})$$

and hence

$$t_{k,k+1}(\nu + \operatorname{cont}(T) + \rho) = \nu + \operatorname{cont}(T^*) + \rho,$$

so that $a_{\nu+\operatorname{cont}(T)+\rho} = -a_{\nu+\operatorname{cont}(T^*)+\rho}$.

Because T and T^* have exactly the same columns $j, j+1, j+2, \ldots$, the tableau T^* is also a violator of at least one of the conditions that $\nu + \operatorname{cont}(T^*|_{\text{cols}\geq j})$ be a partition, and has the same choice of maximal j and minimal k as did T. Hence the map $T \mapsto T^*$ is an involution on the violators that lets one cancel their summands $a_{\nu+\operatorname{cont}(T)+\rho}$ and $a_{\nu+\operatorname{cont}(T^*)+\rho}$ in pairs.[146] □

Example 2.6.10. Here is an example of the construction of T^* in the above proof. Let $n = 6$ and $\lambda = (5, 4, 4)$ and $\mu = (2, 2)$ and $\nu = (1)$. Let T be the column-strict tableau

$$\begin{array}{ccccc} & & 1 & 2 & 2 \\ & & 2 & 3 & \\ 2 & 2 & 3 & 4 & \end{array} \qquad \text{of shape } \lambda/\mu.$$

Then,

$$\begin{aligned} \operatorname{cont}(T|_{\text{cols}\geq 5}) &= (0,1,0,0,0,\ldots) \quad (\text{since } T|_{\text{cols}\geq 5} \text{ has a single entry, which is } 2), \\ \text{so that} \quad \nu + \operatorname{cont}(T|_{\text{cols}\geq 5}) &= (1,1,0,0,0,\ldots) \text{ is a partition.} \end{aligned}$$

But

$$\begin{aligned} \operatorname{cont}(T|_{\text{cols}\geq 4}) &= (0,2,1,1,0,0,0,\ldots), \\ \text{and thus} \quad \nu + \operatorname{cont}(T|_{\text{cols}\geq 4}) &= (1,2,1,1,0,0,0,\ldots) \text{ is not a partition.} \end{aligned}$$

Thus, the j in the above proof of Theorem 2.6.6 is 4. Furthermore, the k in the proof is 1, since $\nu_1 + \operatorname{cont}_1(T|_{\text{cols}\geq 4}) = 1+0 = 1 < 2 = 0+2 = \nu_2 + \operatorname{cont}_2(T|_{\text{cols}\geq 4})$. Thus, T^* is obtained from T by applying the Bender-Knuth involution on letters $1, 2$ to columns $1, 2, 3$ only, leaving columns $4, 5$ unchanged. The result is

$$T^* = \begin{array}{ccccc} & & 1 & 2 & 2 \\ & & 2 & 3 & \\ 1 & 1 & 3 & 4 & \end{array} .$$

So far (in this section) we have worked with a finite set of variables $x_1, x_2, \ldots, x_n$ (where n is a fixed nonnegative integer) and with partitions having at most n parts. We now drop these conventions and restrictions; thus, partitions again mean arbitrary partitions, and $\mathbf{x}$ again means the infinite family $(x_1, x_2, x_3, \ldots)$ of variables. In this setting, we have the following analogue of Corollary 2.6.9:

[145] See Example 2.6.10 below for an example of this construction.

[146] One remark is in order: The tableaux T and T^* may be equal. In this case, the summands $a_{\nu+\operatorname{cont}(T)+\rho}$ and $a_{\nu+\operatorname{cont}(T^*)+\rho}$ do not cancel, as they are the same summand. However, this summand is zero (because $t_{k,k+1}(\nu+\operatorname{cont}(T)+\rho) = \nu+\operatorname{cont}\left(\underbrace{T^*}_{=T}\right)+\rho = \nu+\operatorname{cont}(T)+\rho$ shows that the n-tuple $\nu+\operatorname{cont}(T)+\rho$ has two equal entries, and thus $a_{\nu+\operatorname{cont}(T)+\rho} = 0$), and thus does not affect the sum.

Corollary 2.6.11. *For partitions λ, μ, ν (of any lengths), one has*

$$s_\nu s_{\lambda/\mu} = \sum_T s_{\nu+\operatorname{cont}(T)} \tag{2.6.3}$$

where T runs through all column-strict tableaux of shape λ/μ with the property that for each $j = 1, 2, 3, \ldots$, the weak composition $\nu + \operatorname{cont}(T|_{\operatorname{cols}\geq j})$ is a partition. In particular, taking $\nu = \varnothing$, we obtain

$$s_{\lambda/\mu} = \sum_T s_{\operatorname{cont}(T)} \tag{2.6.4}$$

where in the sum T runs through all column-strict tableaux of shape λ/μ for which each $\operatorname{cont}(T|_{\operatorname{cols}\geq j})$ is a partition.

Proof of Corollary 2.6.11. Essentially, Corollary 2.6.11 is obtained from Corollary 2.6.9 by "letting n (that is, the number of variables) tend to ∞". This can be formalized in different ways: One way is to endow the ring of power series $\mathbf{k}[[\mathbf{x}]] = \mathbf{k}[[x_1, x_2, x_3, \ldots]]$ with the coefficientwise topology[147], and to show that the left hand side of (2.6.1) tends to the left hand side of (2.6.3) when $n \to \infty$, and the same holds for the right hand sides. A different approach proceeds by regarding Λ as the inverse limit of the $\Lambda(x_1, x_2, \ldots, x_n)$. □

Comparing coefficients of a given Schur function s_ν in (2.6.4), we obtain the following version of the Littlewood-Richardson rule.

Corollary 2.6.12. *For partitions λ, μ, ν (of any lengths), the Littlewood-Richardson coefficient $c^\lambda_{\mu,\nu}$ counts column-strict tableaux T of shape λ/μ with $\operatorname{cont}(T) = \nu$ having the property that each $\operatorname{cont}(T|_{\operatorname{cols}\geq j})$ is a partition.*

2.7. **The Pieri and Assaf-McNamara skew Pieri rule.** The classical *Pieri rule* refers to two special cases of the Littlewood-Richardson rule. To state them, recall that a skew shape is called a *horizontal (resp. vertical) strip* if no two of its cells lie in the same column (resp. row). A *horizontal (resp. vertical) n-strip* (for $n \in \mathbb{N}$) shall mean a horizontal (resp. vertical) strip of size n (that is, having exactly n cells).

Theorem 2.7.1. *For every partition λ and any $n \in \mathbb{N}$, we have*

$$s_\lambda h_n = \sum_{\substack{\lambda^+ : \lambda^+/\lambda \text{ is a} \\ \text{horizontal } n\text{-strip}}} s_{\lambda^+}; \tag{2.7.1}$$

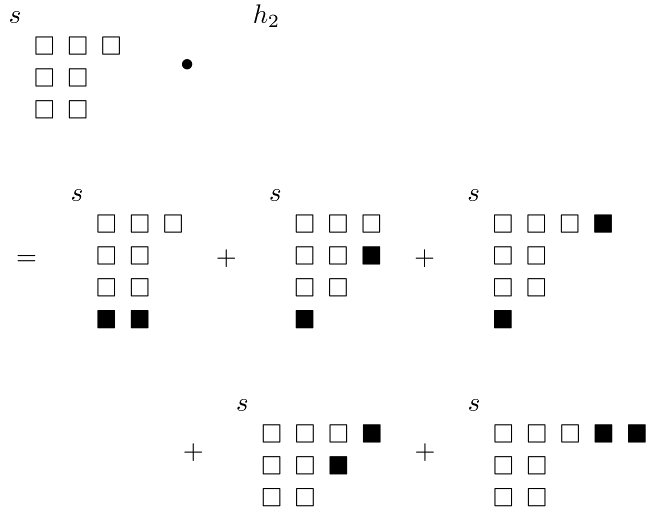
$$s_\lambda e_n = \sum_{\substack{\lambda^+ : \lambda^+/\lambda \text{ is a} \\ \text{vertical } n\text{-strip}}} s_{\lambda^+}. \tag{2.7.2}$$

[147] This topology is defined as follows:

We endow the ring $\mathbf{k}$ with the discrete topology. Then, we can regard the $\mathbf{k}$-module $\mathbf{k}[[\mathbf{x}]]$ as a direct product of infinitely many copies of $\mathbf{k}$ (by identifying every power series in $\mathbf{k}[[\mathbf{x}]]$ with the family of its coefficients). Hence, the product topology is a well-defined topology on $\mathbf{k}[[\mathbf{x}]]$; this topology is denoted as the *coefficientwise topology*. Its name is due to the fact that a sequence $(a_n)_{n\in\mathbb{N}}$ of power series converges to a power series a with respect to this topology if and only if for every monomial $\mathfrak{m}$, all sufficiently high $n \in \mathbb{N}$ satisfy

$$(\text{the coefficient of } \mathfrak{m} \text{ in } a_n) = (\text{the coefficient of } \mathfrak{m} \text{ in } a).$$

Example 2.7.2. In the following equality, we are representing each partition by its Ferrers diagram[148].



If λ is the partition $(3, 2, 2)$ on the left hand side, then all partitions λ^+ on the right hand side visibly have the property that λ^+/λ is a horizontal 2-strip[149], as (2.7.1) predicts.

Proof of Theorem 2.7.1. For the first Pieri formula involving h_n, as $h_n = s_{(n)}$ one has

$$s_\lambda h_n = \sum_{\lambda^+} c^{\lambda^+}_{\lambda,(n)} s_{\lambda^+}.$$

Corollary 2.6.12 says $c^{\lambda^+}_{\lambda,(n)}$ counts column-strict tableaux T of shape λ^+/λ having $\mathrm{cont}(T) = (n)$ (i.e. all entries of T are 1's), with an extra condition. Since its entries are all equal, such a T must certainly have shape being a horizontal strip, and more precisely a horizontal n-strip (since it has n cells). Conversely, for any horizontal n-strip, there is a unique such filling, and it will trivially satisfy the extra condition that $\mathrm{cont}(T|_{\mathrm{cols}\geq j})$ is a partition for each j. Hence $c^{\lambda^+}_{\lambda,(n)}$ is 1 if λ^+/λ is a horizontal n-strip, and 0 else.

For the second Pieri formula involving e_n, using $e_n = s_{(1^n)}$ one has

$$s_\lambda e_n = \sum_{\lambda^+} c^{\lambda^+}_{\lambda,(1^n)} s_{\lambda^+}.$$

Corollary 2.6.12 says $c^{\lambda^+}_{\lambda,(1^n)}$ counts column-strict tableaux T of shape λ^+/λ having $\mathrm{cont}(T) = (1^n)$, so its entries are $1, 2, \ldots, n$ each occurring once, with the extra condition that $1, 2, \ldots, n$ appear from right to left. Together with the tableau condition, this forces at most one entry in each row, that is λ^+/λ is a vertical strip, and then there is a unique way to fill it (maintaining column-strictness and the extra condition that $1, 2, \ldots, n$ appear from right to left). Thus $c^{\lambda^+}_{\lambda,(1^n)}$ is 1 if λ^+/λ is a vertical n-strip, and 0 else. □

In 2009, Assaf and McNamara [9] proved an elegant generalization.

[148] And we are drawing each Ferrers diagram with its boxes spaced out, in order to facilitate counting the boxes.

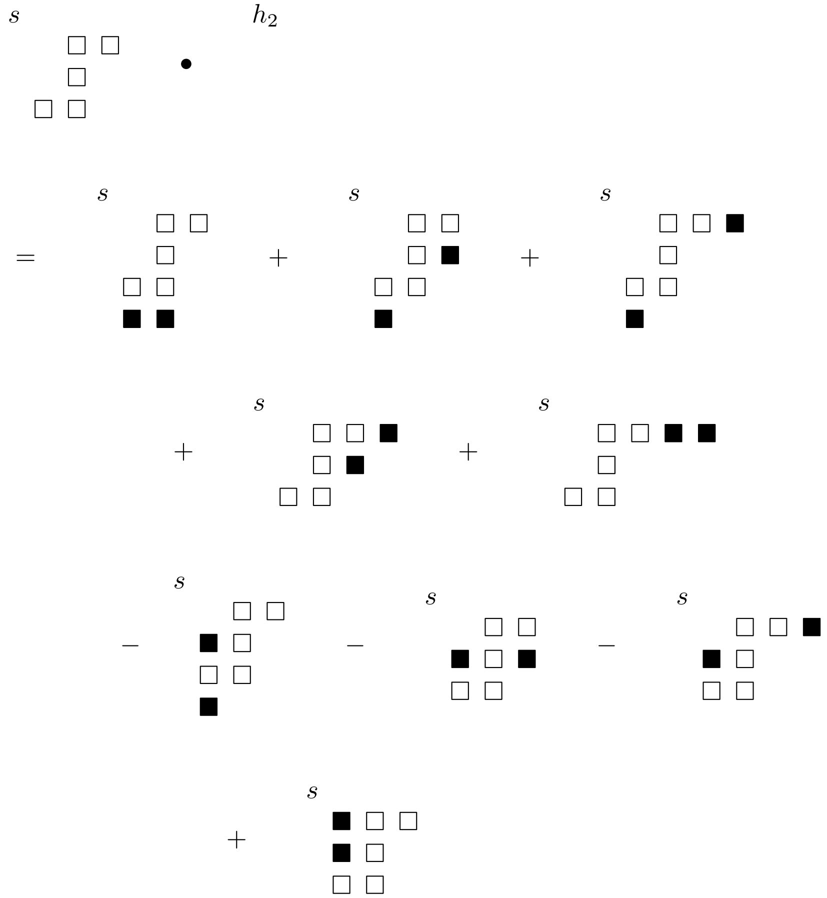
[149] We have colored the boxes of λ^+/λ black.

Theorem 2.7.3. *For any partitions λ and μ and any $n \in \mathbb{N}$, we have*[150]

$$s_{\lambda/\mu} h_n = \sum_{\substack{\lambda^+,\mu^-: \\ \lambda^+/\lambda \text{ a horizontal strip;} \\ \mu/\mu^- \text{ a vertical strip;} \\ |\lambda^+/\lambda|+|\mu/\mu^-|=n}} (-1)^{|\mu/\mu^-|} s_{\lambda^+/\mu^-}; \tag{2.7.3}$$

$$s_{\lambda/\mu} e_n = \sum_{\substack{\lambda^+,\mu^-: \\ \lambda^+/\lambda \text{ a vertical strip;} \\ \mu/\mu^- \text{ a horizontal strip;} \\ |\lambda^+/\lambda|+|\mu/\mu^-|=n}} (-1)^{|\mu/\mu^-|} s_{\lambda^+/\mu^-}. \tag{2.7.4}$$

Example 2.7.4. With the same conventions as in Example 2.7.2[151], we have



which illustrates the first equality of Theorem 2.7.3.

Theorem 2.7.3 is proven in the next section, using an important Hopf algebra tool.

Exercise 2.7.5. Let $\lambda = (\lambda_1, \lambda_2, \lambda_3, \ldots)$ and $\mu = (\mu_1, \mu_2, \mu_3, \ldots)$ be two partitions such that $\mu \subseteq \lambda$.

(a) Show that λ/μ is a horizontal strip if and only if every $i \in \{1, 2, 3, \ldots\}$ satisfies $\mu_i \geq \lambda_{i+1}$. [152]

(b) Show that λ/μ is a vertical strip if and only if every $i \in \{1, 2, 3, \ldots\}$ satisfies $\lambda_i \leq \mu_i + 1$.

Exercise 2.7.6. (a) Let λ and μ be two partitions such that $\mu \subseteq \lambda$. Let $n \in \mathbb{N}$. Show that $\left(h_n, s_{\lambda/\mu}\right)$ equals 1 if λ/μ is a horizontal n-strip, and equals 0 otherwise.

[150] Note that $\mu \subseteq \lambda$ is not required. (The left hand sides are 0 otherwise, but this does not trivialize the equalities.)

[151] but this time coloring both the boxes in λ^+/λ and the boxes in μ/μ^- black

[152] In other words, λ/μ is a horizontal strip if and only if $(\lambda_2, \lambda_3, \lambda_4, \ldots) \subseteq \mu$. This simple observation has been used by Pak and Postnikov [165, §10] for a new approach to RSK-type algorithms.

(b) Use part (a) to give a new proof of (2.7.1).

Exercise 2.7.7. Prove Theorem 2.7.1 again using the ideas of the proof of Theorem 2.5.1.

Exercise 2.7.8. Let A be a commutative ring, and $n \in \mathbb{N}$.

(a) Let $a_1, a_2, \ldots, a_n$ be n elements of A. Let $b_1, b_2, \ldots, b_n$ be n further elements of A. If $a_i - b_j$ is an invertible element of A for every $i \in \{1, 2, \ldots, n\}$ and $j \in \{1, 2, \ldots, n\}$, then prove that

$$\det\left(\left(\frac{1}{a_i - b_j}\right)_{i,j=1,2,\ldots,n}\right) = \frac{\prod_{1 \le j < i \le n}\left(\left(a_i - a_j\right)\left(b_j - b_i\right)\right)}{\prod_{(i,j) \in \{1,2,\ldots,n\}^2}\left(a_i - b_j\right)}.$$

(b) Let $a_1, a_2, \ldots, a_n$ be n elements of A. Let $b_1, b_2, \ldots, b_n$ be n further elements of A. If $1 - a_i b_j$ is an invertible element of A for every $i \in \{1, 2, \ldots, n\}$ and $j \in \{1, 2, \ldots, n\}$, then prove that

$$\det\left(\left(\frac{1}{1 - a_i b_j}\right)_{i,j=1,2,\ldots,n}\right) = \frac{\prod_{1 \le j < i \le n}\left(\left(a_i - a_j\right)\left(b_i - b_j\right)\right)}{\prod_{(i,j) \in \{1,2,\ldots,n\}^2}\left(1 - a_i b_j\right)}.$$

(c) Use the result of part (b) to give a new proof for Theorem 2.5.1.[153]

The determinant on the left hand side of Exercise 2.7.8(a) is known as the *Cauchy determinant*.

Exercise 2.7.9. Prove that $s_{(a,b)} = h_a h_b - h_{a+1} h_{b-1}$ for any two integers $a \ge b \ge 0$ (where we set $h_{-1} = 0$ as usual).

(Note that this is precisely the Jacobi-Trudi formula (2.4.16) in the case when $\lambda = (a, b)$ is a partition with at most two entries and $\mu = \varnothing$.)

Exercise 2.7.10. If λ is a partition and μ is a weak composition, let $K_{\lambda,\mu}$ denote the number of column-strict tableaux T of shape λ having $\operatorname{cont}(T) = \mu$. (This $K_{\lambda,\mu}$ is called the (λ, μ)-*Kostka number*.)

(a) Use Theorem 2.7.1 to show that every partition μ satisfies $h_\mu = \sum_\lambda K_{\lambda,\mu} s_\lambda$, where the sum ranges over all partitions λ.

(b) Use this to give a new proof for Theorem 2.5.1.[154]

(c) Give a new proof of the fact (previously shown as Proposition 2.4.3(j)) that $(h_\lambda)_{\lambda \in \mathrm{Par}}$ is a graded basis of the graded $\mathbf{k}$-module Λ.

Exercise 2.7.11. (a) Define a $\mathbf{k}$-linear map $\mathfrak{Z} : \Lambda \to \Lambda$ by having it send s_λ to s_{λ^t} for every partition λ. (This is clearly well-defined, since $(s_\lambda)_{\lambda \in \mathrm{Par}}$ is a $\mathbf{k}$-basis of Λ.) Show that

$$\mathfrak{Z}(f h_n) = \mathfrak{Z}(f) \cdot \mathfrak{Z}(h_n) \qquad \text{for every } f \in \Lambda \text{ and every } n \in \mathbb{N}.$$

(b) Show that $\mathfrak{Z} = \omega$.

(c) Show that $c^{\lambda}_{\mu,\nu} = c^{\lambda^t}_{\mu^t,\nu^t}$ for any three partitions λ, μ and ν.

(d) Use this to prove (2.4.15).[155]

Exercise 2.7.12. (a) Show that

$$\prod_{i,j=1}^{\infty}\left(1 + x_i y_j\right) = \sum_{\lambda \in \mathrm{Par}} s_\lambda(\mathbf{x})\, s_{\lambda^t}(\mathbf{y}) = \sum_{\lambda \in \mathrm{Par}} e_\lambda(\mathbf{x})\, m_\lambda(\mathbf{y})$$

in the power series ring $\mathbf{k}[[\mathbf{x}, \mathbf{y}]] = \mathbf{k}[[x_1, x_2, x_3, \ldots, y_1, y_2, y_3, \ldots]]$.

(b) Assume that $\mathbb{Q}$ is a subring of $\mathbf{k}$. Show that

$$\prod_{i,j=1}^{\infty}\left(1 + x_i y_j\right) = \sum_{\lambda \in \mathrm{Par}} (-1)^{|\lambda| - \ell(\lambda)} z_\lambda^{-1} p_\lambda(\mathbf{x})\, p_\lambda(\mathbf{y})$$

in the power series ring $\mathbf{k}[[\mathbf{x}, \mathbf{y}]] = \mathbf{k}[[x_1, x_2, x_3, \ldots, y_1, y_2, y_3, \ldots]]$, where z_λ is defined as in Proposition 2.5.15.

[153] This approach to Theorem 2.5.1 is taken in [44, §4] (except that [44] only works with finitely many variables).

[154] Of course, this gives a new proof of Theorem 2.5.1 only when coupled with a proof of Theorem 2.7.1 which does not rely on Theorem 2.5.1. The proof of Theorem 2.7.1 we gave in the text above did not rely on Theorem 2.5.1, whereas the proof of (2.7.1) given in Exercise 2.7.6(b) did.

[155] The first author learned this approach to (2.4.15) from Alexander Postnikov.

The first equality of Exercise 2.7.12(a) appears in [206, Thm. 7.14.3], [186, Thm. 4.8.6] and several other references under the name of the *dual Cauchy identity*, and is commonly proven using a "dual" analogue of the Robinson-Schensted-Knuth algorithm.

Exercise 2.7.13. Prove Theorem 2.4.6.

[**Hint:**[156] Switch $\mathbf{x}$ and $\mathbf{y}$ in the formula of Exercise 2.5.11(a), and specialize the resulting equality by replacing $\mathbf{y}$ by a finite set of variables $(y_1, y_2, \ldots, y_\ell)$; then, set $n = \ell$ and $\rho = (n-1, n-2, \ldots, 0)$, and multiply with the alternant $a_\rho(y_1, y_2, \ldots, y_\ell)$, using Corollary 2.6.7 to simplify the result; finally, extract the coefficient of $\mathbf{y}^{\lambda+\rho}$.]

Exercise 2.7.14. Prove the following:

(a) We have $(S(f), S(g)) = (f, g)$ for all $f \in \Lambda$ and $g \in \Lambda$.

(b) We have $(e_n, f) = (-1)^n \cdot (S(f))(1)$ for any $n \in \mathbb{N}$ and $f \in \Lambda_n$. (See Exercise 2.1.2 for the meaning of $(S(f))(1)$.)

2.8. **Skewing and Lam's proof of the skew Pieri rule.** We codify here the operation $s_\mu^\perp$ of *skewing by* s_μ, acting on Schur functions via

$$s_\mu^\perp(s_\lambda) = s_{\lambda/\mu}$$

(where, as before, one defines $s_{\lambda/\mu} = 0$ if $\mu \not\subseteq \lambda$). These operations play a crucial role

- in Lam's proof of the skew Pieri rule,
- in Lam, Lauve, and Sottile's proof [120] of a more general skew Littlewood-Richardson rule that had been conjectured by Assaf and McNamara, and
- in Zelevinsky's structure theory of PSH's to be developed in the next chapter.

We are going to define them in the general setting of any graded Hopf algebra.

Definition 2.8.1. Given a graded Hopf algebra A, and its (graded) dual A^o, let $(\cdot,\cdot) = (\cdot,\cdot)_A : A^o \times A \to \mathbf{k}$ be the pairing defined by $(f, a) := f(a)$ for f in A^o and a in A. Then define for each f in A^o an operator $A \overset{f^\perp}{\to} A$ as follows[157]: for a in A with $\Delta(a) = \sum a_1 \otimes a_2$, let

$$f^\perp(a) = \sum (f, a_1) a_2.$$

In other words, $f^\perp$ is the composition

$$A \xrightarrow{\Delta} A \otimes A \xrightarrow{f \otimes \mathrm{id}} \mathbf{k} \otimes A \xrightarrow{\cong} A,$$

where the rightmost arrow is the canonical isomorphism $\mathbf{k} \otimes A \to A$. This operator $f^\perp$ is called *skewing by* f.

Now, recall that the Hall inner product induces an isomorphism $\Lambda^o \cong \Lambda$ (by Corollary 2.5.14). Hence, we can regard any element $f \in \Lambda$ as an element of Λ^o; this allows us to define an operator $f^\perp : \Lambda \to \Lambda$ for each $f \in \Lambda$ (by regarding f as an element of Λ^o, and applying Definition 2.8.1 to $A = \Lambda$). Explicitly, this operator is given by

$$f^\perp(a) = \sum (f, a_1) a_2 \qquad \text{whenever} \qquad \Delta(a) = \sum a_1 \otimes a_2, \tag{2.8.1}$$

where the inner product (f, a_1) is now understood as a Hall inner product.

Recall that each partition λ satisfies

$$\Delta s_\lambda = \sum_{\mu \subseteq \lambda} s_\mu \otimes s_{\lambda/\mu} = \sum_{\nu \subseteq \lambda} s_\nu \otimes s_{\lambda/\nu} = \sum_\nu s_\nu \otimes s_{\lambda/\nu}$$

(since $s_{\lambda/\nu} = 0$ unless $\nu \subseteq \lambda$). Hence, for any two partitions λ and μ, we have

$$\begin{aligned} s_\mu^\perp(s_\lambda) &= \sum_\nu \underbrace{(s_\mu, s_\nu)}_{=\delta_{\mu,\nu}} s_{\lambda/\nu} \qquad \text{(by (2.8.1), applied to } f = s_\mu \text{ and } a = s_\lambda) \\ &= \sum_\nu \delta_{\mu,\nu} s_{\lambda/\nu} = s_{\lambda/\mu}. \qquad (2.8.2) \end{aligned}$$

[156] This is the proof given in Stanley [206, §7.16, Second Proof of Thm. 7.16.1] and Macdonald [142, proof of (5.4)].

[157] This $f^\perp(a)$ is called $a \leftharpoonup f$ in Montgomery [157, Example 1.6.5].

Thus, skewing acts on the Schur functions exactly as desired.

Proposition 2.8.2. *Let A be a graded Hopf algebra. The $f^\perp$ operators $A \to A$ have the following properties.*

(i) *For every $f \in A^o$, the map $f^\perp$ is adjoint to left multiplication $A^o \xrightarrow{f\cdot} A^o$ in the sense that*

$$(g, f^\perp(a)) = (fg, a).$$

(ii) *For every $f, g \in A^o$, we have $(fg)^\perp(a) = g^\perp(f^\perp(a))$, that is, A becomes a right A^o-module via the $f^\perp$ action.*[158]

(iii) *The unity 1_{A^o} of the* **k***-algebra A^o satisfies $(1_{A^o})^\perp = \mathrm{id}_A$.*

(iv) *Assume that A is of finite type (so A^o becomes a Hopf algebra, not just an algebra). If an $f \in A^o$ satisfies $\Delta(f) = \sum f_1 \otimes f_2$, then*

$$f^\perp(ab) = \sum f_1^\perp(a) f_2^\perp(b).$$

In particular, if f is primitive in A^o, so that $\Delta(f) = f \otimes 1 + 1 \otimes f$, then $f^\perp$ is a derivation*:*

$$f^\perp(ab) = f^\perp(a) \cdot b + a \cdot f^\perp(b).$$

Proof. For (i), note that

$$(g, f^\perp(a)) = \sum (f, a_1)(g, a_2) = (f \otimes g, \Delta_A(a)) = (m_{A^o}(f \otimes g), a) = (fg, a).$$

For (ii), using (i) and considering any h in A^o, one has that

$$(h, (fg)^\perp(a)) = (fgh, a) = (gh, f^\perp(a)) = (h, g^\perp(f^\perp(a))).$$

For (iii), we recall that the unity 1_{A^o} of A^o is the counit ϵ of A, and thus every $a \in A$ satisfies

$$\begin{aligned}(1_{A^o})^\perp(a) = \epsilon^\perp(a) &= \sum_{(a)} \underbrace{(\epsilon, a_1)}_{=\epsilon(a_1)} a_2 \qquad \left(\text{by the definition of } \epsilon^\perp\right) \\ &= \sum_{(a)} \epsilon(a_1) a_2 = a \qquad (\text{by the axioms of a coalgebra}),\end{aligned}$$

so that $(1_{A^o})^\perp = \mathrm{id}_A$.

For (iv), noting that

$$\Delta(ab) = \Delta(a)\Delta(b) = \left(\sum_{(a)} a_1 \otimes a_2\right)\left(\sum_{(b)} b_1 \otimes b_2\right) = \sum_{(a),(b)} a_1 b_1 \otimes a_2 b_2,$$

one has that

$$\begin{aligned} f^\perp(ab) &= \sum_{(a),(b)} (f, a_1 b_1)_A \ a_2 b_2 = \sum_{(a),(b)} (\Delta(f), a_1 \otimes b_1)_{A \otimes A} \ a_2 b_2 \\ &= \sum_{(f),(a),(b)} (f_1, a_1)_A (f_2, b_1)_A \ a_2 b_2 \\ &= \sum_{(f)} \left(\sum_{(a)} (f_1, a_1)_A a_2\right)\left(\sum_{(b)} (f_2, b_1)_A b_2\right) = \sum_{(f)} f_1^\perp(a) f_2^\perp(b). \end{aligned}$$

□

The Pieri rules (Theorem 2.7.1) expressed multiplication by h_n or by e_n in the basis $(s_\lambda)_{\lambda \in \mathrm{Par}}$ of Λ. We can similarly express skewing by h_n or by e_n:

[158] This makes sense, since A^o is a **k**-algebra (by Exercise 1.6.1(c), applied to $C = A$).

Proposition 2.8.3. *For every partition λ and any $n \in \mathbb{N}$, we have*

$$h_n^\perp s_\lambda = \sum_{\substack{\lambda^- : \lambda/\lambda^- \text{ is a} \\ \text{horizontal } n\text{-strip}}} s_{\lambda^-}; \tag{2.8.3}$$

$$e_n^\perp s_\lambda = \sum_{\substack{\lambda^- : \lambda/\lambda^- \text{ is a} \\ \text{vertical } n\text{-strip}}} s_{\lambda^-}. \tag{2.8.4}$$

Exercise 2.8.4. Prove Proposition 2.8.3.
[**Hint:** Use Theorem 2.7.1 and $(s_{\mu^-}, e_n^\perp s_\mu) = (e_n s_{\mu^-}, s_\mu)$.]

The following interaction between multiplication and $h^\perp$ is the key to deducing the skew Pieri formula from the usual Pieri formulas.

Lemma 2.8.5. *For any f, g in Λ and any $n \in \mathbb{N}$, one has*

$$f \cdot h_n^\perp(g) = \sum_{k=0}^n (-1)^k h_{n-k}^\perp (e_k^\perp(f) \cdot g).$$

Proof. Starting with the right side, first apply Proposition 2.8.2(iv):

$$\begin{aligned}
&\sum_{k=0}^n (-1)^k \underbrace{h_{n-k}^\perp(e_k^\perp(f) \cdot g)}_{\substack{= \sum_{j=0}^{n-k} h_j^\perp(e_k^\perp(f)) \cdot h_{n-k-j}^\perp(g) \\ \text{(by Proposition 2.8.2(iv), applied} \\ \text{to } h_{n-k},\ e_k^\perp(f) \text{ and } g \text{ instead of } f,\ a \text{ and } b)}} \\
&= \sum_{k=0}^n (-1)^k \sum_{j=0}^{n-k} h_j^\perp(e_k^\perp(f)) \cdot h_{n-k-j}^\perp(g) \\
&= \sum_{j=0}^n \sum_{k=0}^{n-j} (-1)^k h_j^\perp(e_k^\perp(f)) \cdot h_{n-k-j}^\perp(g) \\
&= \sum_{j=0}^n \sum_{i=0}^{n-j} (-1)^{n-i-j} h_j^\perp(e_{n-i-j}^\perp(f)) \cdot h_i^\perp(g) \qquad \text{(reindexing } i := n-k-j \text{ in the inner sum)} \\
&= \sum_{i=0}^n (-1)^{n-i} \left(\sum_{j=0}^{n-i} (-1)^j h_j^\perp(e_{n-i-j}^\perp(f)) \right) \cdot h_i^\perp(g) \\
&= \sum_{i=0}^n (-1)^{n-i} \left(\sum_{j=0}^{n-i} (-1)^j e_{n-i-j} h_j \right)^\perp (f) \cdot h_i^\perp(g) \qquad \text{(by Proposition 2.8.2(ii))} \\
&= 1^\perp(f) \cdot h_n^\perp(g) = f \cdot h_n^\perp(g)
\end{aligned}$$

where the second-to-last equality used (2.4.4). □

Proof of Theorem 2.7.3. We prove (2.7.3); the equality (2.7.4) is analogous, swapping $h_i \leftrightarrow e_i$ and swapping the words "vertical" ↔ "horizontal". For any $f \in \Lambda$, we have

$$\begin{aligned}
\left(s_{\lambda/\mu}, f\right) &= \left(s_\mu^\perp(s_\lambda), f\right) && \text{(by (2.8.2))} \\
&= \left(f, s_\mu^\perp(s_\lambda)\right) && \text{(by symmetry of } (\cdot,\cdot)_\Lambda) \\
&= (s_\mu f, s_\lambda) && \text{(by Proposition 2.8.2(i))} \\
&= (s_\lambda, s_\mu f) && \text{(by symmetry of } (\cdot,\cdot)_\Lambda) .
\end{aligned} \tag{2.8.5}$$

Hence for any g in Λ, one can compute that

$$\begin{aligned}(h_n s_{\lambda/\mu}\ ,\ g) &\underset{2.8.2(i)}{\overset{Prop.}{=}} (s_{\lambda/\mu}\ ,\ h_n^\perp g) \overset{(2.8.5)}{=} (s_\lambda\ ,\ s_\mu \cdot h_n^\perp g)\\ &\underset{2.8.5}{\overset{Lemma}{=}} \sum_{k=0}^{n} (-1)^k (s_\lambda\ ,\ h_{n-k}^\perp (e_k^\perp (s_\mu) \cdot g))\\ &\underset{2.8.2(i)}{\overset{Prop.}{=}} \sum_{k=0}^{n} (-1)^k (h_{n-k} s_\lambda\ , e_k^\perp (s_\mu) \cdot g). \qquad (2.8.6)\end{aligned}$$

The first Pieri rule in Theorem 2.7.1 lets one rewrite $h_{n-k}s_\lambda = \sum_{\lambda^+} s_{\lambda^+}$, with the sum running through λ^+ for which λ^+/λ is a horizontal $(n-k)$-strip. Meanwhile, (2.8.4) lets one rewrite $e_k^\perp s_\mu = \sum_{\mu^-} s_{\mu^-}$, with the sum running through μ^- for which μ/μ^- is a vertical k-strip. Thus the right hand side of (2.8.6) becomes

$$\sum_{k=0}^{n} (-1)^k \left(\sum_{\lambda^+} s_{\lambda^+}\ ,\ \sum_{\mu^-} s_{\mu^-} \cdot g \right) \overset{(2.8.5)}{=} \left(\sum_{k=0}^{n} (-1)^k \sum_{(\lambda^+,\mu^-)} s_{\lambda^+/\mu^-}, g \right)$$

where the sum is over the pairs (λ^+, μ^-) for which λ^+/λ is a horizontal $(n-k)$-strip and μ/μ^- is a vertical k-strip. This proves (2.7.3). □

Exercise 2.8.6. Let $n \in \mathbb{N}$.

(a) For every $k \in \mathbb{N}$, let $p(n,k)$ denote the number of partitions of n of length k. Let $c(n)$ denote the number of *self-conjugate* partitions of n (that is, partitions λ of n satisfying $\lambda^t = \lambda$). Show that

$$(-1)^n c(n) = \sum_{k=0}^{n} (-1)^k p(n,k).$$

(This application of Hopf algebras was found by Aguiar and Lauve, [5, §5.1]. See also [206, Chapter 1, Exercise 22(b)] for an elementary proof.)

(b) For every partition λ, let $C(\lambda)$ denote the number of corner cells of the Ferrers diagram of λ (these are the cells of the Ferrers diagram whose neighbors to the east and to the south both lie outside of the Ferrers diagram). For every partition λ, let $\mu_1(\lambda)$ denote the number of parts of λ equal to 1. Show that

$$\sum_{\lambda \in \operatorname{Par}_n} C(\lambda) = \sum_{\lambda \in \operatorname{Par}_n} \mu_1(\lambda).$$

(This is also due to Stanley.)

Exercise 2.8.7. The goal of this exercise is to prove (2.4.15) using the skewing operators that we have developed.[159] Recall the involution $\omega : \Lambda \to \Lambda$ defined in (2.4.10).

(a) Show that $\omega(p_\lambda) = (-1)^{|\lambda|-\ell(\lambda)} p_\lambda$ for any $\lambda \in \operatorname{Par}$, where $\ell(\lambda)$ denotes the length of the partition λ.
(b) Show that ω is an isometry.
(c) Show that this same map $\omega : \Lambda \to \Lambda$ is a Hopf automorphism.
(d) Prove that $\omega\left(a^\perp b\right) = (\omega(a))^\perp (\omega(b))$ for every $a \in \Lambda$ and $b \in \Lambda$.
(e) For any partition $\lambda = (\lambda_1, \ldots, \lambda_\ell)$ with length $\ell(\lambda) = \ell$, prove that

$$e_\ell^\perp s_\lambda = s_{(\lambda_1-1,\lambda_2-1,\ldots,\lambda_\ell-1)}.$$

(f) For any partition $\lambda = (\lambda_1, \lambda_2, \ldots)$, prove that

$$h_{\lambda_1}^\perp s_\lambda = s_{(\lambda_2,\lambda_3,\lambda_4,\ldots)}.$$

(g) Prove (2.4.15).

Exercise 2.8.8. Let n be a positive integer. Prove the following:

(a) We have $(e_n, p_n) = (-1)^{n-1}$.
(b) We have $(e_m, p_n) = 0$ for each $m \in \mathbb{N}$ satisfying $m \neq n$.

[159] Make sure not to use the results of Exercise 2.7.11 or Exercise 2.7.12 or Exercise 2.7.14 here, or anything else that relied on (2.4.15), in order to avoid circular reasoning.

(c) We have $e_n^\perp p_n = (-1)^{n-1}$.
(d) We have $e_m^\perp p_n = 0$ for each positive integer m satisfying $m \neq n$.

2.9. Assorted exercises on symmetric functions.

Over a hundred exercises on symmetric functions are collected in Stanley's [206, chapter 7], and even more (but without any hints or references) on his website[160]. Further sources for results related to symmetric functions are Macdonald's work, including his monograph [142] and his expository [143]. In this section, we gather a few exercises that are not too difficult to handle with the material given above.

Exercise 2.9.1. (a) Let $m \in \mathbb{Z}$. Prove that, for every $f \in \Lambda$, the infinite sum $\sum_{i \in \mathbb{N}} (-1)^i h_{m+i} e_i^\perp f$ is convergent in the discrete topology (i.e., all but finitely many addends of this sum are zero). Hence, we can define a map $\mathbf{B}_m : \Lambda \to \Lambda$ by setting

$$\mathbf{B}_m (f) = \sum_{i \in \mathbb{N}} (-1)^i h_{m+i} e_i^\perp f \qquad \text{for all } f \in \Lambda.$$

Show that this map $\mathbf{B}_m$ is $\mathbf{k}$-linear.

(b) Let $\lambda = (\lambda_1, \lambda_2, \lambda_3, \ldots)$ be a partition, and let $m \in \mathbb{Z}$ be such that $m \geq \lambda_1$. Show that

$$\sum_{i \in \mathbb{N}} (-1)^i h_{m+i} e_i^\perp s_\lambda = s_{(m,\lambda_1,\lambda_2,\lambda_3,\ldots)}.$$

(c) Let $n \in \mathbb{N}$. For every n-tuple $(\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{Z}^n$, we define an element $\overline{s}_{(\alpha_1,\alpha_2,\ldots,\alpha_n)} \in \Lambda$ by

$$\overline{s}_{(\alpha_1,\alpha_2,\ldots,\alpha_n)} = \det\left(\left(h_{\alpha_i - i + j} \right)_{i,j=1,2,\ldots,n} \right).$$

Show that

$$s_\lambda = \overline{s}_{(\lambda_1,\lambda_2,\ldots,\lambda_n)} \tag{2.9.1}$$

for every partition $\lambda = (\lambda_1, \lambda_2, \lambda_3, \ldots)$ having at most n parts[161].

Furthermore, show that for every n-tuple $(\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{Z}^n$, the symmetric function $\overline{s}_{(\alpha_1,\alpha_2,\ldots,\alpha_n)}$ either is 0 or equals $\pm s_\nu$ for some partition ν having at most n parts.

Finally, show that for any n-tuples $(\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{Z}^n$ and $(\beta_1, \beta_2, \ldots, \beta_n) \in \mathbb{N}^n$, we have

$$\overline{s}^\perp_{(\beta_1,\beta_2,\ldots,\beta_n)} \overline{s}_{(\alpha_1,\alpha_2,\ldots,\alpha_n)} = \det\left(\left(h_{\alpha_i - \beta_j - i + j} \right)_{i,j=1,2,\ldots,n} \right). \tag{2.9.2}$$

(d) For every $n \in \mathbb{N}$, every $m \in \mathbb{Z}$ and every n-tuple $(\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{Z}^n$, prove that

$$\sum_{i \in \mathbb{N}} (-1)^i h_{m+i} e_i^\perp \overline{s}_{(\alpha_1,\alpha_2,\ldots,\alpha_n)} = \overline{s}_{(m,\alpha_1,\alpha_2,\ldots,\alpha_n)}, \tag{2.9.3}$$

where we are using the notations of Exercise 2.9.1(c).

(e) For every $n \in \mathbb{N}$ and every n-tuple $(\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{Z}^n$, prove that

$$\overline{s}_{(\alpha_1,\alpha_2,\ldots,\alpha_n)} = \left(\mathbf{B}_{\alpha_1} \circ \mathbf{B}_{\alpha_2} \circ \cdots \circ \mathbf{B}_{\alpha_n} \right)(1),$$

where we are using the notations of Exercise 2.9.1(c) and Exercise 2.9.1(a).

(f) For every $m \in \mathbb{Z}$ and every positive integer n, prove that $\mathbf{B}_m (p_n) = h_m p_n - h_{m+n}$. Here, we are using the notations of Exercise 2.9.1(a).

Remark 2.9.2. The map $\mathbf{B}_m$ defined in Exercise 2.9.1(a) is the so-called *m-th Bernstein creation operator*; it appears in Zelevinsky [227, §4.20(a)] and has been introduced by J.N. Bernstein, who found the result of Exercise 2.9.1(b). It is called a "Schur row adder" in [74]. Exercise 2.9.1(e) appears in Berg/Bergeron/Saliola/Serrano/Zabrocki [17, Theorem 2.3], where it is used as a prototype for defining *noncommutative* analogues of Schur functions, the so-called *immaculate functions*. The particular case of Exercise 2.9.1(e) for $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ a partition of length n (a restatement of Exercise 2.9.1(b)) is proven in [142, §I.5, example 29].

[160] http://math.mit.edu/~rstan/ec/ch7supp.pdf

[161] Recall that a *part* of a partition means a nonzero entry of the partition.

Exercise 2.9.3. (a) Prove that there exists a unique family $(x_n)_{n\geq 1}$ of elements of Λ such that

$$H(t) = \prod_{n=1}^{\infty} (1 - x_n t^n)^{-1}.$$

Denote this family $(x_n)_{n\geq 1}$ by $(w_n)_{n\geq 1}$. For instance,

$$w_1 = s_{(1)}, \qquad w_2 = -s_{(1,1)}, \qquad w_3 = -s_{(2,1)},$$
$$w_4 = -s_{(1,1,1,1)} - s_{(2,1,1)} - s_{(2,2)} - s_{(3,1)}, \qquad w_5 = -s_{(2,1,1,1)} - s_{(2,2,1)} - s_{(3,1,1)} - s_{(3,2)} - s_{(4,1)}.$$

(b) Show that w_n is homogeneous of degree n for every positive integer n.

(c) For every partition λ, define $w_\lambda \in \Lambda$ by $w_\lambda = w_{\lambda_1} w_{\lambda_2} \cdots w_{\lambda_\ell}$ (where $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_\ell)$ with $\ell = \ell(\lambda)$). Notice that w_λ is homogeneous of degree $|\lambda|$. Prove that $\sum_{\lambda \in \operatorname{Par}_n} w_\lambda = h_n$ for every $n \in \mathbb{N}$.

(d) Show that $\{w_\lambda\}_{\lambda \in \operatorname{Par}}$ is a $\mathbf{k}$-basis of Λ. (This basis is called the *Witt basis*[162]; it is studied in [90, §9-§10].[163])

(e) Prove that $p_n = \sum_{d|n} d w_d^{n/d}$ for every positive integer n. (Here, the summation sign $\sum_{d|n}$ means a sum over all positive divisors d of n.)

(f) We are going to show that $-w_n$ is a sum of Schur functions (possibly with repetitions, but without signs!) for every $n \geq 2$. (For $n = 1$, the opposite is true: w_1 is a single Schur function.) This proof goes back to Doran [55][164].

For any positive integers n and k, define $f_{n,k} \in \Lambda$ by $f_{n,k} = \sum_{\substack{\lambda \in \operatorname{Par}_n, \\ \min \lambda \geq k}} w_\lambda$, where $\min \lambda$ denotes the smallest part[165] of λ. Show that

$$-f_{n,k} = s_{(n-1,1)} + \sum_{i=2}^{k-1} f_{i,i} f_{n-i,i} \qquad \text{for every } n \geq k \geq 2.$$

Conclude that $-f_{n,k}$ is a sum of Schur functions for every $n \in \mathbb{N}$ and $k \geq 2$. Conclude that $-w_n$ is a sum of Schur functions for every $n \geq 2$.

(g) For every partition λ, define $r_\lambda \in \Lambda$ by $r_\lambda = \prod_{i\geq 1} h_{v_i}(x_1^i, x_2^i, x_3^i, \ldots)$, where v_i is the number of occurrences of i in λ. Show that $\sum_{\lambda \in \operatorname{Par}} w_\lambda(\mathbf{x}) r_\lambda(\mathbf{y}) = \prod_{i,j=1}^{\infty} (1 - x_i y_j)^{-1}$.

(h) Show that $\{r_\lambda\}_{\lambda \in \operatorname{Par}}$ and $\{w_\lambda\}_{\lambda \in \operatorname{Par}}$ are dual bases of Λ.

Exercise 2.9.4. For this exercise, set $\mathbf{k} = \mathbb{Z}$, and consider $\Lambda = \Lambda_{\mathbb{Z}}$ as a subring of $\Lambda_{\mathbb{Q}}$. Also, consider $\Lambda \otimes_{\mathbb{Z}} \Lambda$ as a subring of $\Lambda_{\mathbb{Q}} \otimes_{\mathbb{Q}} \Lambda_{\mathbb{Q}}$. [166] Recall that the family $(p_n)_{n\geq 1}$ generates the $\mathbb{Q}$-algebra $\Lambda_{\mathbb{Q}}$, but does not generate the $\mathbb{Z}$-algebra Λ.

(a) Define a $\mathbb{Q}$-linear map $Z : \Lambda_{\mathbb{Q}} \to \Lambda_{\mathbb{Q}}$ by setting

$$Z(p_\lambda) = z_\lambda p_\lambda \qquad \text{for every partition } \lambda,$$

where z_λ is defined as in Proposition 2.5.15.[167] Show that $Z(\Lambda) \subset \Lambda$.

[162] This is due to its relation with Witt vectors in the appropriate sense. Most of the work on this basis has been done by Reutenauer and Hazewinkel.

[163] It also implicitly appears in [12, §5]. Indeed, the q_n of [12] are our w_n (for $\mathbf{k} = R$).

[164] See also Stanley [206, Exercise 7.46].

[165] Recall that a *part* of a partition means a nonzero entry of the partition.

[166] Here is how this works: We have $\Lambda_{\mathbb{Q}} \cong \mathbb{Q} \otimes_{\mathbb{Z}} \Lambda$. But fundamental properties of tensor products yield

$$\mathbb{Q} \otimes_{\mathbb{Z}} (\Lambda \otimes_{\mathbb{Z}} \Lambda) \cong \underbrace{(\mathbb{Q} \otimes_{\mathbb{Z}} \Lambda)}_{\cong \Lambda_{\mathbb{Q}}} \otimes_{\mathbb{Q}} \underbrace{(\mathbb{Q} \otimes_{\mathbb{Z}} \Lambda)}_{\cong \Lambda_{\mathbb{Q}}} \cong \Lambda_{\mathbb{Q}} \otimes_{\mathbb{Q}} \Lambda_{\mathbb{Q}} \tag{2.9.4}$$

as $\mathbb{Q}$-algebras. But $\Lambda \otimes_{\mathbb{Z}} \Lambda$ is a free $\mathbb{Z}$-module (since Λ is a free $\mathbb{Z}$-module), and so the canonical ring homomorphism $\Lambda \otimes_{\mathbb{Z}} \Lambda \to \mathbb{Q} \otimes_{\mathbb{Z}} (\Lambda \otimes_{\mathbb{Z}} \Lambda)$ sending every u to $1_{\mathbb{Q}} \otimes_{\mathbb{Z}} u$ is injective. Composing this ring homomorphism with the $\mathbb{Q}$-algebra isomorphism of (2.9.4) gives an injective ring homomorphism $\Lambda \otimes_{\mathbb{Z}} \Lambda \to \Lambda_{\mathbb{Q}} \otimes_{\mathbb{Q}} \Lambda_{\mathbb{Q}}$. We use this latter homomorphism to identify $\Lambda \otimes_{\mathbb{Z}} \Lambda$ with a subring of $\Lambda_{\mathbb{Q}} \otimes_{\mathbb{Q}} \Lambda_{\mathbb{Q}}$.

[167] This is well-defined, since $(p_\lambda)_{\lambda \in \operatorname{Par}}$ is a $\mathbb{Q}$-module basis of $\Lambda_{\mathbb{Q}}$.

(b) Define a $\mathbb{Q}$-algebra homomorphism $\Delta_\times : \Lambda_\mathbb{Q} \to \Lambda_\mathbb{Q} \otimes_\mathbb{Q} \Lambda_\mathbb{Q}$ by setting

$$\Delta_\times (p_n) = p_n \otimes p_n \qquad \text{for every positive integer } n.$$

[168] Show that $\Delta_\times (\Lambda) \subset \Lambda \otimes_\mathbb{Z} \Lambda$.

(c) Let $r \in \mathbb{Z}$. Define a $\mathbb{Q}$-algebra homomorphism $\epsilon_r : \Lambda_\mathbb{Q} \to \mathbb{Q}$ by setting

$$\epsilon_r (p_n) = r \qquad \text{for every positive integer } n.$$

[169] Show that $\epsilon_r (\Lambda) \subset \mathbb{Z}$.

(d) Let $r \in \mathbb{Z}$. Define a $\mathbb{Q}$-algebra homomorphism $\mathbf{i}_r : \Lambda_\mathbb{Q} \to \Lambda_\mathbb{Q}$ by setting

$$\mathbf{i}_r (p_n) = r p_n \qquad \text{for every positive integer } n.$$

[170] Show that $\mathbf{i}_r (\Lambda) \subset \Lambda$.

(e) Define a $\mathbb{Q}$-linear map $\operatorname{Sq} : \Lambda_\mathbb{Q} \to \Lambda_\mathbb{Q}$ by setting

$$\operatorname{Sq} (p_\lambda) = p_\lambda^2 \qquad \text{for every partition } \lambda.$$

[171] Show that $\operatorname{Sq} (\Lambda) \subset \Lambda$.

(f) Let $r \in \mathbb{Z}$. Define a $\mathbb{Q}$-algebra homomorphism $\Delta_r : \Lambda_\mathbb{Q} \to \Lambda_\mathbb{Q} \otimes_\mathbb{Q} \Lambda_\mathbb{Q}$ by setting

$$\Delta_r (p_n) = \sum_{i=1}^{n-1} \binom{n}{i} p_i \otimes p_{n-i} + r \otimes p_n + p_n \otimes r \qquad \text{for every positive integer } n.$$

[172] Show that $\Delta_r (\Lambda) \subset \Lambda \otimes_\mathbb{Z} \Lambda$.

(g) Consider the map $\Delta_\times$ introduced in Exercise 2.9.4(b) and the map ϵ_1 introduced in Exercise 2.9.4(c). Show that the $\mathbb{Q}$-algebra $\Lambda_\mathbb{Q}$, endowed with the comultiplication $\Delta_\times$ and the counit ϵ_1, becomes a cocommutative $\mathbb{Q}$-bialgebra.[173]

(h) Define a $\mathbb{Q}$-bilinear map $* : \Lambda_\mathbb{Q} \times \Lambda_\mathbb{Q} \to \Lambda_\mathbb{Q}$, which will be written in infix notation (that is, we will write $a * b$ instead of $* (a, b)$), by setting

$$p_\lambda * p_\mu = \delta_{\lambda,\mu} z_\lambda p_\lambda \qquad \text{for any partitions } \lambda \text{ and } \mu$$

(where z_λ is defined as in Proposition 2.5.15). [174] Show that $f * g \in \Lambda$ for any $f \in \Lambda$ and $g \in \Lambda$.

(i) Show that $\epsilon_1 (f) = f (1)$ for every $f \in \Lambda_\mathbb{Q}$ (where we are using the notation ϵ_r defined in Exercise 2.9.4(c)).

[**Hint:**

- For (b), show that, for every $f \in \Lambda_\mathbb{Q}$, the tensor $\Delta_\times (f)$ is the preimage of $f\left((x_i y_j)_{(i,j) \in \{1,2,3,\ldots\}^2} \right) = f (x_1 y_1, x_1 y_2, x_1 y_3, \ldots, x_2 y_1, x_2 y_2, x_2 y_3, \ldots, \ldots) \in \mathbb{Q} [[\mathbf{x}, \mathbf{y}]]$ under the canonical injection $\Lambda_\mathbb{Q} \otimes_\mathbb{Q} \Lambda_\mathbb{Q} \to \mathbb{Q} [[\mathbf{x}, \mathbf{y}]]$ which maps every $f \otimes g$ to $f (\mathbf{x}) g (\mathbf{y})$. (This requires making sure that the evaluation $f\left((x_i y_j)_{(i,j) \in \{1,2,3,\ldots\}^2} \right)$ is well-defined to begin with, i.e., converges as a formal power series.)

  For an alternative solution to (b), compute $\Delta_\times (h_n)$ or $\Delta_\times (e_n)$.
- For (c), compute $\epsilon_r (e_n)$ or $\epsilon_r (h_n)$.
- Reduce (d) to (b) and (c) using Exercise 1.3.6.
- Reduce (e) to (b).
- (f) is the hardest part. It is tempting to try and interpret the definition of Δ_r as a convoluted way of saying that $\Delta_r (f)$ is the preimage of $f\left((x_i + y_j)_{(i,j) \in \{1,2,3,\ldots\}^2} \right)$ under the canonical injection $\Lambda_\mathbb{Q} \otimes_\mathbb{Q} \Lambda_\mathbb{Q} \to \mathbb{Q} [[\mathbf{x}, \mathbf{y}]]$ which maps every $f \otimes g$ to $f (\mathbf{x}) g (\mathbf{y})$. However, this does not make sense since the evaluation $f\left((x_i + y_j)_{(i,j) \in \{1,2,3,\ldots\}^2} \right)$ is (in general) not well-defined[175] (and even if it was, it would fail to explain the r). So we need to get down to finitely many variables. For every $N \in \mathbb{N}$, define a

[168] This is well-defined, since the family $(p_n)_{n \geq 1}$ generates the $\mathbb{Q}$-algebra $\Lambda_\mathbb{Q}$ and is algebraically independent.

[169] This is well-defined, since the family $(p_n)_{n \geq 1}$ generates the $\mathbb{Q}$-algebra $\Lambda_\mathbb{Q}$ and is algebraically independent.

[170] This is well-defined, since the family $(p_n)_{n \geq 1}$ generates the $\mathbb{Q}$-algebra $\Lambda_\mathbb{Q}$ and is algebraically independent.

[171] This is well-defined, since $(p_\lambda)_{\lambda \in \operatorname{Par}}$ is a $\mathbb{Q}$-module basis of $\Lambda_\mathbb{Q}$.

[172] This is well-defined, since the family $(p_n)_{n \geq 1}$ generates the $\mathbb{Q}$-algebra $\Lambda_\mathbb{Q}$ and is algebraically independent.

[173] But unlike $\Lambda_\mathbb{Q}$ with the usual coalgebra structure, it is neither graded nor a Hopf algebra.

[174] This is well-defined, since $(p_\lambda)_{\lambda \in \operatorname{Par}}$ is a $\mathbb{Q}$-module basis of $\Lambda_\mathbb{Q}$.

[175] e.g., it involves summing infinitely many x_1's if $f = e_1$

$\mathbb{Q}$-algebra homomorphism $\mathcal{E}_N : \Lambda_\mathbb{Q} \otimes_\mathbb{Q} \Lambda_\mathbb{Q} \to \mathbb{Q}[x_1, x_2, \ldots, x_N, y_1, y_2, \ldots, y_N]$ by sending each $f \otimes g$ to $f(x_1, x_2, \ldots, x_N) g(y_1, y_2, \ldots, y_N)$. Show that $\Delta_N(\Lambda) \subset \mathcal{E}_N^{-1}(\mathbb{Z}[x_1, x_2, \ldots, x_N, y_1, y_2, \ldots, y_N])$. This shows that, at least, the coefficients of $\Delta_r(f)$ in front of the $m_\lambda \otimes m_\mu$ with $\ell(\lambda) \leq r$ and $\ell(\mu) \leq r$ (in the $\mathbb{Q}$-basis $(m_\lambda \otimes m_\mu)_{\lambda,\mu \in \mathrm{Par}}$ of $\Lambda_\mathbb{Q} \otimes_\mathbb{Q} \Lambda_\mathbb{Q}$) are integral for $f \in \Lambda$. Of course, we want all coefficients. Show that $\Delta_a = \Delta_b \star \left(\Delta_{\Lambda_\mathbb{Q}} \circ \mathbf{i}_{a-b}\right)$ in $\operatorname{Hom}(\Lambda_\mathbb{Q}, \Lambda_\mathbb{Q} \otimes_\mathbb{Q} \Lambda_\mathbb{Q})$ for any integers a and b. This allows "moving" the r. This approach to (f) was partly suggested to the first author by Richard Stanley.

- For (h), notice that Definition 3.1.1(b) (below) allows us to construct a bilinear form $(\cdot,\cdot)_{\Lambda_\mathbb{Q} \otimes_\mathbb{Q} \Lambda_\mathbb{Q}} : (\Lambda_\mathbb{Q} \otimes_\mathbb{Q} \Lambda_\mathbb{Q}) \times (\Lambda_\mathbb{Q} \otimes_\mathbb{Q} \Lambda_\mathbb{Q}) \to \mathbb{Q}$ from the Hall inner product $(\cdot,\cdot) : \Lambda_\mathbb{Q} \times \Lambda_\mathbb{Q} \to \mathbb{Q}$. Show that

$$(a * b, c) = (a \otimes b, \Delta_\times(c))_{\Lambda_\mathbb{Q} \otimes_\mathbb{Q} \Lambda_\mathbb{Q}} \qquad \text{for all } a, b, c \in \Lambda_\mathbb{Q}, \tag{2.9.5}$$

and then use (b).

]

Remark 2.9.5. The map $\Delta_\times$ defined in Exercise 2.9.4(b) is known as the *internal comultiplication* (or *Kronecker comultiplication*) on $\Lambda_\mathbb{Q}$. Unlike the standard comultiplication $\Delta_{\Lambda_\mathbb{Q}}$, it is not a graded map, but rather sends every homogeneous component $(\Lambda_\mathbb{Q})_n$ into $(\Lambda_\mathbb{Q})_n \otimes (\Lambda_\mathbb{Q})_n$. The bilinear map $*$ from Exercise 2.9.4(h) is the so-called *internal multiplication* (or *Kronecker multiplication*), and is similarly not graded but rather takes $(\Lambda_\mathbb{Q})_n \times (\Lambda_\mathbb{Q})_m$ to $(\Lambda_\mathbb{Q})_n$ if $n = m$ and to 0 otherwise.

The analogy between the two internal structures is not perfect: While we saw in Exercise 2.9.4(g) how the internal comultiplication yields another bialgebra structure on $\Lambda_\mathbb{Q}$, it is not true that the internal multiplication (combined with the usual coalgebra structure of $\Lambda_\mathbb{Q}$) forms a bialgebra structure as well. What is missing is a multiplicative unity; if we would take the closure of $\Lambda_\mathbb{Q}$ with respect to the grading, then $1 + h_1 + h_2 + h_3 + \cdots$ would be such a unity.

The structure constants of the internal comultiplication on the Schur basis $(s_\lambda)_{\lambda \in \mathrm{Par}}$ are equal to the structure constants of the internal multiplication on the Schur basis[176], and are commonly referred to as the *Kronecker coefficients*. They are known to be nonnegative integers (this follows from Exercise 4.4.8(c)[177]), but no combinatorial proof is known for their nonnegativity. Combinatorial interpretations for these coefficients akin to the Littlewood-Richardson rule have been found only in special cases (cf., e.g., [183] and [23] and [132]).

The map Δ_r of Exercise 2.9.4(f) also has some classical theory behind it, relating to Chern classes of tensor products ([151], [142, §I.4, example 5]).

Parts (b), (c), (d), (e) and (f) of Exercise 2.9.4 are instances of a general phenomenon: Many $\mathbb{Z}$-algebra homomorphisms $\Lambda \to A$ (with A a commutative ring, usually torsionfree) are easiest to define by first defining a $\mathbb{Q}$-algebra homomorphism $\Lambda_\mathbb{Q} \to A \otimes \mathbb{Q}$ and then showing that this homomorphism restricts to a $\mathbb{Z}$-algebra homomorphism $\Lambda \to A$. One might ask for general criteria when this is possible; specifically, for what choices of $(b_n)_{n \geq 1} \in A^{\{1,2,3,\ldots\}}$ does there exist a $\mathbb{Z}$-algebra homomorphism $\Lambda \to A$ sending the p_n to b_n ? Such choices are called *ghost-Witt vectors* in Hazewinkel [90], and we can give various equivalent conditions for a family $(b_n)_{n \geq 1}$ to be a ghost-Witt vector:

Exercise 2.9.6. Let A be a commutative ring.

For every $n \in \{1, 2, 3, \ldots\}$, let $\varphi_n : A \to A$ be a ring endomorphism of A. Assume that the following properties hold:

- We have $\varphi_n \circ \varphi_m = \varphi_{nm}$ for any two positive integers n and m.
- We have $\varphi_1 = \mathrm{id}$.
- We have $\varphi_p(a) \equiv a^p \bmod pA$ for every $a \in A$ and every prime number p.

(For example, when $A = \mathbb{Z}$, one can set $\varphi_n = \mathrm{id}$ for all n; this simplifies the exercise somewhat. More generally, setting $\varphi_n = \mathrm{id}$ works whenever A is a binomial ring[178]. However, the results of this exercise are at their most useful when A is a multivariate polynomial ring $\mathbb{Z}[x_1, x_2, x_3, \ldots]$ over $\mathbb{Z}$ and the homomorphism φ_n sends every $P \in A$ to $P(x_1^n, x_2^n, x_3^n, \ldots)$.)

[176] This can be obtained, e.g., from (2.9.5).

[177] Their integrality can also be easily deduced from Exercise 2.9.4(b).

[178] A *binomial ring* is defined to be a torsionfree (as an additive group) commutative ring A which has one of the following equivalent properties:

Let μ denote the *number-theoretic Möbius function*; this is the function $\{1,2,3,\ldots\} \to \mathbb{Z}$ defined by

$$\mu(m) = \begin{cases} 0, & \text{if } m \text{ is not squarefree;} \\ (-1)^{(\text{number of prime factors of } m)}, & \text{if } m \text{ is squarefree} \end{cases} \qquad \text{for every positive integer } m.$$

Let ϕ denote the *Euler totient function*; this is the function $\{1,2,3,\ldots\} \to \mathbb{N}$ which sends every positive integer m to the number of elements of $\{1,2,\ldots,m\}$ coprime to m.

Let $(b_n)_{n\geq 1} \in A^{\{1,2,3,\ldots\}}$ be a family of elements of A. Prove that the following seven assertions are equivalent:

- *Assertion $\mathcal{C}$:* For every positive integer n and every prime factor p of n, we have
$$\varphi_p\left(b_{n/p}\right) \equiv b_n \mod p^{v_p(n)} A.$$
Here, $v_p(n)$ denotes the exponent of p in the prime factorization of n.
- *Assertion $\mathcal{D}$:* There exists a family $(\alpha_n)_{n\geq 1} \in A^{\{1,2,3,\ldots\}}$ of elements of A such that every positive integer n satisfies
$$b_n = \sum_{d\mid n} d\alpha_d^{n/d}.$$
[179]
- *Assertion $\mathcal{E}$:* There exists a family $(\beta_n)_{n\geq 1} \in A^{\{1,2,3,\ldots\}}$ of elements of A such that every positive integer n satisfies
$$b_n = \sum_{d\mid n} d\varphi_{n/d}(\beta_d).$$
- *Assertion $\mathcal{F}$:* Every positive integer n satisfies
$$\sum_{d\mid n} \mu(d)\varphi_d\left(b_{n/d}\right) \in nA.$$
- *Assertion $\mathcal{G}$:* Every positive integer n satisfies
$$\sum_{d\mid n} \phi(d)\varphi_d\left(b_{n/d}\right) \in nA.$$
- *Assertion $\mathcal{H}$:* Every positive integer n satisfies
$$\sum_{i=1}^{n} \varphi_{n/\gcd(i,n)}\left(b_{\gcd(i,n)}\right) \in nA.$$
- *Assertion $\mathcal{J}$:* There exists a ring homomorphism $\Lambda_{\mathbb{Z}} \to A$ which, for every positive integer n, sends p_n to b_n.

- For every $n \in \mathbb{N}$ and $a \in A$, we have $a(a-1)\cdots(a-n+1) \in n! \cdot A$. (That is, binomial coefficients $\binom{a}{n}$ with $a \in A$ and $n \in \mathbb{N}$ are defined in A.)
- We have $a^p \equiv a \mod pA$ for every $a \in A$ and every prime number p.

See [226] and the references therein for studies of these rings. It is not hard to check that $\mathbb{Z}$ and every localization of $\mathbb{Z}$ are binomial rings, and so is any commutative $\mathbb{Q}$-algebra as well as the ring

$$\{P \in \mathbb{Q}[X] \mid P(n) \in \mathbb{Z} \text{ for every } n \in \mathbb{Z}\}$$

(but not the ring $\mathbb{Z}[X]$ itself).

[179] Here and in the following, summations of the form $\sum_{d\mid n}$ range over all **positive** divisors of n.

[**Hint:** The following identities hold for every positive integer n:

$$\sum_{d \mid n} \phi\left(d\right) = n; \tag{2.9.6}$$

$$\sum_{d \mid n} \mu\left(d\right) = \delta_{n,1}; \tag{2.9.7}$$

$$\sum_{d \mid n} \mu\left(d\right) \frac{n}{d} = \phi\left(n\right); \tag{2.9.8}$$

$$\sum_{d \mid n} d\mu\left(d\right) \phi\left(\frac{n}{d}\right) = \mu\left(n\right). \tag{2.9.9}$$

Furthermore, the following simple lemma is useful: If k is a positive integer, and if $p \in \mathbb{N}$, $a \in A$ and $b \in A$ are such that $a \equiv b \mod p^k A$, then $a^{p^\ell} \equiv b^{p^\ell} \mod p^{k+\ell} A$ for every $\ell \in \mathbb{N}$.]

Remark 2.9.7. Much of Exercise 2.9.6 is folklore, but it is hard to pinpoint concrete appearances in literature. The equivalence $\mathcal{C} \iff \mathcal{D}$ appears in Hesselholt [95, Lemma 1] and [96, Lemma 1.1] (in slightly greater generality), where it is referred to as Dwork's lemma and used in the construction of the Witt vector functor. This equivalence is also [90, Lemma 9.93]. The equivalence $\mathcal{D} \iff \mathcal{F} \iff \mathcal{G} \iff \mathcal{H}$ in the case $A = \mathbb{Z}$ is [57, Corollary on p. 10], where it is put into the context of Burnside rings and necklace counting. The equivalence $\mathcal{C} \iff \mathcal{F}$ for finite families $\left(b_n\right)_{n \in \{1,2,\ldots,m\}}$ in lieu of $\left(b_n\right)_{n \geq 1}$ is [206, Exercise 5.2 **a**]. One of the likely oldest relevant sources is Schur's [195], which proves the equivalence $\mathcal{C} \iff \mathcal{D} \iff \mathcal{F}$ for finite families $\left(b_n\right)_{n \in \{1,2,\ldots,m\}}$, as well as a "finite version" of $\mathcal{C} \iff \mathcal{J}$ (Schur did not have Λ, but was working with actual power sums of roots of polynomials).

Exercise 2.9.8. Let A denote the ring $\mathbb{Z}$. For every $n \in \{1, 2, 3, \ldots\}$, let φ_n denote the identity endomorphism id of A. Prove that the seven equivalent assertions $\mathcal{C}$, $\mathcal{D}$, $\mathcal{E}$, $\mathcal{F}$, $\mathcal{G}$, $\mathcal{H}$ and $\mathcal{J}$ of Exercise 2.9.6 are satisfied for each of the following families $\left(b_n\right)_{n \geq 1} \in \mathbb{Z}^{\{1,2,3,\ldots\}}$:

- the family $\left(b_n\right)_{n \geq 1} = \left(q^n\right)_{n \geq 1}$, where q is a given integer.
- the family $\left(b_n\right)_{n \geq 1} = \left(q\right)_{n \geq 1}$, where q is a given integer.
- the family $\left(b_n\right)_{n \geq 1} = \left(\dbinom{qn}{rn}\right)_{n \geq 1}$, where $r \in \mathbb{Q}$ and $q \in \mathbb{Z}$ are given. (Here, a binomial coefficient $\dbinom{a}{b}$ has to be interpreted as 0 when $b \notin \mathbb{N}$.)
- the family $\left(b_n\right)_{n \geq 1} = \left(\dbinom{qn-1}{rn-1}\right)_{n \geq 1}$, where $r \in \mathbb{Z}$ and $q \in \mathbb{Z}$ are given.

Exercise 2.9.9. For every $n \in \{1, 2, 3, \ldots\}$, define a map $\mathbf{f}_n : \Lambda \to \Lambda$ by setting

$$\mathbf{f}_n\left(a\right) = a\left(x_1^n, x_2^n, x_3^n, \ldots\right) \qquad \text{for every } a \in \Lambda.$$

(So what $\mathbf{f}_n$ does to a symmetric function is replacing all variables $x_1, x_2, x_3, \ldots$ by their n-th powers.)

(a) Show that $\mathbf{f}_n : \Lambda \to \Lambda$ is a $\mathbf{k}$-algebra homomorphism for every $n \in \{1, 2, 3, \ldots\}$.
(b) Show that $\mathbf{f}_n \circ \mathbf{f}_m = \mathbf{f}_{nm}$ for any two positive integers n and m.
(c) Show that $\mathbf{f}_1 = \mathrm{id}$.
(d) Prove that $\mathbf{f}_n : \Lambda \to \Lambda$ is a Hopf algebra homomorphism for every $n \in \{1, 2, 3, \ldots\}$.
(e) Prove that $\mathbf{f}_2\left(h_m\right) = \sum_{i=0}^{2m} \left(-1\right)^i h_i h_{2m-i}$ for every $m \in \mathbb{N}$.
(f) Assume that $\mathbf{k} = \mathbb{Z}$. Prove that $\mathbf{f}_p\left(a\right) \equiv a^p \mod p\Lambda$ for every $a \in \Lambda$ and every prime number p.
(g) Use Exercise 2.9.6 to obtain new solutions to parts (b), (c), (d), (e) and (f) of Exercise 2.9.4.

The maps $\mathbf{f}_n$ constructed in Exercise 2.9.9 are known as the *Frobenius endomorphisms* of Λ. They are a (deceptively) simple particular case of the notion of *plethysm* ([206, Chapter 7, Appendix 2] and [142, Section I.8]), and are often used as intermediate steps in computing more complicated plethysms[180].

[180] In the notations of [206, (A2.160)], the value $\mathbf{f}_n\left(a\right)$ for an $a \in \Lambda$ can be written as $a\left[p_n\right]$ or (when $\mathbf{k} = \mathbb{Z}$) as $p_n\left[a\right]$.

Exercise 2.9.10. For every $n \in \{1, 2, 3, \ldots\}$, define a $\mathbf{k}$-algebra homomorphism $\mathbf{v}_n : \Lambda \to \Lambda$ by

$$\mathbf{v}_n \left(h_m\right) = \begin{cases} h_{m/n}, & \text{if } n \mid m; \\ 0, & \text{if } n \nmid m \end{cases} \qquad \text{for every positive integer } m$$

[181].

(a) Show that any positive integers n and m satisfy

$$\mathbf{v}_n \left(p_m\right) = \begin{cases} n p_{m/n}, & \text{if } n \mid m; \\ 0, & \text{if } n \nmid m \end{cases} .$$

(b) Show that any positive integers n and m satisfy

$$\mathbf{v}_n \left(e_m\right) = \begin{cases} \left(-1\right)^{m-m/n} e_{m/n}, & \text{if } n \mid m; \\ 0, & \text{if } n \nmid m \end{cases} .$$

(c) Prove that $\mathbf{v}_n \circ \mathbf{v}_m = \mathbf{v}_{nm}$ for any two positive integers n and m.
(d) Prove that $\mathbf{v}_1 = \mathrm{id}$.
(e) Prove that $\mathbf{v}_n : \Lambda \to \Lambda$ is a Hopf algebra homomorphism for every $n \in \{1, 2, 3, \ldots\}$.

Now, consider also the maps $\mathbf{f}_n : \Lambda \to \Lambda$ defined in Exercise 2.9.9. Fix a positive integer n.

(f) Prove that the maps $\mathbf{f}_n : \Lambda \to \Lambda$ and $\mathbf{v}_n : \Lambda \to \Lambda$ are adjoint with respect to the Hall inner product on Λ.
(g) Show that $\mathbf{v}_n \circ \mathbf{f}_n = \mathrm{id}_\Lambda^{\star n}$.
(h) Prove that $\mathbf{f}_n \circ \mathbf{v}_m = \mathbf{v}_m \circ \mathbf{f}_n$ whenever m is a positive integer coprime to n.

Finally, recall the $w_m \in \Lambda$ defined in Exercise 2.9.3.

(i) Show that any positive integer m satisfies

$$\mathbf{v}_n \left(w_m\right) = \begin{cases} w_{m/n}, & \text{if } n \mid m; \\ 0, & \text{if } n \nmid m \end{cases} .$$

The homomorphisms $\mathbf{v}_n : \Lambda \to \Lambda$ defined in Exercise 2.9.10 are called the *Verschiebung endomorphisms* of Λ; this name comes from German, where "Verschiebung" means "shift". This terminology, as well as that of Frobenius endomorphisms, originates in the theory of Witt vectors, and the connection between the Frobenius and Verschiebung endomorphisms of Λ and the identically named operators on Witt vectors is elucidated in [90, Chapter 13][182].

Exercise 2.9.11. Fix $n \in \mathbb{N}$. For any n-tuple $w = (w_1, w_2, \ldots, w_n)$ of integers, define the *descent set* $\mathrm{Des}(w)$ of w to be the set $\{i \in \{1, 2, \ldots, n-1\} : w_i > w_{i+1}\}$.

(a) We say that an n-tuple $(w_1, w_2, \ldots, w_n)$ is *Smirnov* if every $i \in \{1, 2, \ldots, n-1\}$ satisfies $w_i \neq w_{i+1}$.
Fix $k \in \mathbb{N}$, and let $X_{n,k} \in \mathbf{k}[[\mathbf{x}]]$ denote the sum of the monomials $x_{w_1} x_{w_2} \cdots x_{w_n}$ over all Smirnov n-tuples $w = (w_1, w_2, \ldots, w_n) \in \{1, 2, 3, \ldots\}^n$ satisfying $|\mathrm{Des}(w)| = k$. Prove that $X_{n,k} \in \Lambda$.
(b) For any n-tuple $w = (w_1, w_2, \ldots, w_n)$, define the *stagnation set* $\mathrm{Stag}(w)$ of w to be the set $\{i \in \{1, 2, \ldots, n-1\} : w_i = w_{i+1}\}$. (Thus, an n-tuple is Smirnov if and only if its stagnation set is empty.)
For any $d \in \mathbb{N}$ and $s \in \mathbb{N}$, define a power series $X_{n,d,s} \in \mathbf{k}[[\mathbf{x}]]$ as the sum of the monomials $x_{w_1} x_{w_2} \cdots x_{w_n}$ over all n-tuples $w = (w_1, w_2, \ldots, w_n) \in \{1, 2, 3, \ldots\}^n$ satisfying $|\mathrm{Des}(w)| = d$ and $|\mathrm{Stag}(w)| = s$. Prove that $X_{n,d,s} \in \Lambda$ for any nonnegative integers d and s.

[181] This is well-defined, since the family $(h_m)_{m \geq 1}$ generates the $\mathbf{k}$-algebra Λ and is algebraically independent.

[182] which is also where most of the statements of Exercises 2.9.9 and 2.9.10 come from

(c) Assume that n is positive. For any $d \in \mathbb{N}$ and $s \in \mathbb{N}$, define three further power series $U_{n,d,s}$, $V_{n,d,s}$ and $W_{n,d,s}$ in $\mathbf{k}[[\mathbf{x}]]$ by the following formulas:

$$U_{n,d,s} = \sum_{\substack{w=(w_1,w_2,\ldots,w_n)\in\{1,2,3,\ldots\}^n;\\ |\operatorname{Des}(w)|=d;\ |\operatorname{Stag}(w)|=s;\\ w_1<w_n}} x_{w_1}x_{w_2}\cdots x_{w_n}; \tag{2.9.10}$$

$$V_{n,d,s} = \sum_{\substack{w=(w_1,w_2,\ldots,w_n)\in\{1,2,3,\ldots\}^n;\\ |\operatorname{Des}(w)|=d;\ |\operatorname{Stag}(w)|=s;\\ w_1=w_n}} x_{w_1}x_{w_2}\cdots x_{w_n}; \tag{2.9.11}$$

$$W_{n,d,s} = \sum_{\substack{w=(w_1,w_2,\ldots,w_n)\in\{1,2,3,\ldots\}^n;\\ |\operatorname{Des}(w)|=d;\ |\operatorname{Stag}(w)|=s;\\ w_1>w_n}} x_{w_1}x_{w_2}\cdots x_{w_n}. \tag{2.9.12}$$

Prove that these three power series $U_{n,d,s}$, $V_{n,d,s}$ and $W_{n,d,s}$ belong to Λ.

Remark 2.9.12. The function $X_{n,k}$ in Exercise 2.9.11(a) is a simple example ([199, Example 2.5, Theorem C.3]) of a chromatic quasisymmetric function that happens to be symmetric. See Shareshian/Wachs [199] for more general criteria for such functions to be symmetric, as well as deeper results. For example, [199, Theorem 6.3] gives an expansion for a wide class of chromatic quasisymmetric functions in the Schur basis of Λ, which, in particular, shows that our $X_{n,k}$ satisfies

$$X_{n,k} = \sum_{\lambda\in\operatorname{Par}_n} a_{\lambda,k} s_\lambda,$$

where $a_{\lambda,k}$ is the number of all assignments T of entries in $\{1,2,\ldots,n\}$ to the cells of the Ferrers diagram of λ such that the following four conditions are satisfied:

- Every element of $\{1,2,\ldots,n\}$ is used precisely once in the assignment (i.e., we have $\operatorname{cont}(T) = (1^n)$).
- Whenever a cell y of the Ferrers diagram lies immediately to the right of a cell x, we have $T(y) - T(x) \geq 2$.
- Whenever a cell y of the Ferrers diagram lies immediately below a cell x, we have $T(y) - T(x) \geq -1$.
- There exist precisely k elements $i \in \{1,2,\ldots,n-1\}$ such that the cell $T^{-1}(i)$ lies in a row below $T^{-1}(i+1)$.

Are there any such rules for the $X_{n,d,s}$ of part (b)?

Smirnov n-tuples are more usually called Smirnov words, or (occasionally) Carlitz words.

See [68, Chapter 6] for further properties of the symmetric functions $U_{n,d,0}$, $V_{n,d,0}$ and $W_{n,d,0}$ from Exercise 2.9.11(c) (or, more precisely, of their generating functions $\sum_d U_{n,d,0}t^d$ etc.).

Exercise 2.9.13. (a) Let $n \in \mathbb{N}$. Define a matrix $A_n = (a_{i,j})_{i,j=1,2,\ldots,n} \in \Lambda^{n\times n}$ by

$$a_{i,j} = \begin{cases} p_{i-j+1}, & \text{if } i \geq j; \\ i, & \text{if } i = j-1; \\ 0, & \text{if } i < j-1 \end{cases} \qquad \text{for all } (i,j) \in \{1,2,\ldots,n\}^2.$$

This matrix A_n looks as follows:

$$A_n = \begin{pmatrix} p_1 & 1 & 0 & \cdots & 0 & 0 \\ p_2 & p_1 & 2 & \cdots & 0 & 0 \\ p_3 & p_2 & p_1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ p_{n-1} & p_{n-2} & p_{n-3} & \cdots & p_1 & n-1 \\ p_n & p_{n-1} & p_{n-2} & \cdots & p_2 & p_1 \end{pmatrix}.$$

Show that $\det(A_n) = n!e_n$.

(b) Let n be a positive integer. Define a matrix $B_n = (b_{i,j})_{i,j=1,2,\ldots,n} \in \Lambda^{n\times n}$ by

$$b_{i,j} = \begin{cases} ie_i, & \text{if } j = 1; \\ e_{i-j+1}, & \text{if } j > 1 \end{cases} \qquad \text{for all } (i,j) \in \{1,2,\ldots,n\}^2.$$

The matrix B_n looks as follows:

$$B_n = \begin{pmatrix} e_1 & e_0 & e_{-1} & \cdots & e_{-n+3} & e_{-n+2} \\ 2e_2 & e_1 & e_0 & \cdots & e_{-n+4} & e_{-n+3} \\ 3e_3 & e_2 & e_1 & \cdots & e_{-n+5} & e_{-n+4} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ (n-1)\,e_{n-1} & e_{n-2} & e_{n-3} & \cdots & e_1 & e_0 \\ ne_n & e_{n-1} & e_{n-2} & \cdots & e_2 & e_1 \end{pmatrix}$$
$$= \begin{pmatrix} e_1 & 1 & 0 & \cdots & 0 & 0 \\ 2e_2 & e_1 & 1 & \cdots & 0 & 0 \\ 3e_3 & e_2 & e_1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ (n-1)\,e_{n-1} & e_{n-2} & e_{n-3} & \cdots & e_1 & 1 \\ ne_n & e_{n-1} & e_{n-2} & \cdots & e_2 & e_1 \end{pmatrix}.$$

Show that $\det(B_n) = p_n$.

The formulas of Exercise 2.9.13, for finitely many variables, appear in Prasolov's [171, §4.1][183]. In [171, §4.2], Prasolov gives four more formulas, which express e_n as a polynomial in the $h_1, h_2, h_3, \ldots$, or h_n as a polynomial in the $e_1, e_2, e_3, \ldots$, or p_n as a polynomial in the $h_1, h_2, h_3, \ldots$, or $n!h_n$ as a polynomial in the $p_1, p_2, p_3, \ldots$. These are not novel for us, since the first two of them are particular cases of Theorem 2.4.6, whereas the latter two can be derived from Exercise 2.9.13 by applying ω. (Note that ω is only well-defined on symmetric functions in infinitely many indeterminates, so we need to apply ω **before** evaluating at finitely many indeterminates; this explains why Prasolov has to prove the latter two identities separately.)

Exercise 2.9.14. In the following, if $k \in \mathbb{N}$, we shall use the notation 1^k for $\underbrace{1, 1, \ldots, 1}_{k \text{ times}}$ (in contexts such as $(n, 1^m)$). So, for example, $(3, 1^4)$ is the partition $(3, 1, 1, 1, 1)$.

(a) Show that $e_n h_m = s_{(m+1,1^{n-1})} + s_{(m,1^n)}$ for any two positive integers n and m.

(b) Show that
$$\sum_{i=0}^{b} (-1)^i \, h_{a+i+1} e_{b-i} = s_{(a+1,1^b)}$$
for any $a \in \mathbb{N}$ and $b \in \mathbb{N}$.

(c) Show that
$$\sum_{i=0}^{b} (-1)^i \, h_{a+i+1} e_{b-i} = (-1)^b \, \delta_{a+b,-1}$$
for any negative integer a and every $b \in \mathbb{N}$. (As usual, we set $h_j = 0$ for $j < 0$ here.)

(d) Show that
$$\Delta s_{(a+1,1^b)} = 1 \otimes s_{(a+1,1^b)} + s_{(a+1,1^b)} \otimes 1 + \sum_{\substack{(c,d,e,f)\in\mathbb{N}^4; \\ c+e=a-1; \\ d+f=b}} s_{(c+1,1^d)} \otimes s_{(e+1,1^f)} + \sum_{\substack{(c,d,e,f)\in\mathbb{N}^4; \\ c+e=a; \\ d+f=b-1}} s_{(c+1,1^d)} \otimes s_{(e+1,1^f)}$$
for any $a \in \mathbb{N}$ and $b \in \mathbb{N}$.

Our next few exercises survey some results on Littlewood-Richardson coefficients.

Exercise 2.9.15. Let $m \in \mathbb{N}$ and $k \in \mathbb{N}$. Let λ and μ be two partitions such that $\ell(\lambda) \le k$ and $\ell(\mu) \le k$. Assume that all parts of λ and all parts of μ are $\le m$. (It is easy to see that this assumption is equivalent to requiring $\lambda_i \le m$ and $\mu_i \le m$ for every positive integer i. [184]). Let $\lambda^\vee$ and $\mu^\vee$ denote the k-tuples $(m-\lambda_k, m-\lambda_{k-1}, \ldots, m-\lambda_1)$ and $(m-\mu_k, m-\mu_{k-1}, \ldots, m-\mu_1)$, respectively.

[183] where our symmetric functions e_k, h_k, p_k, evaluated in finitely many indeterminates, are denoted σ_k, p_k, s_k, respectively

[184] As usual, we are denoting by ν_i the i-th entry of a partition ν here.

(a) Show that $\lambda^\vee$ and $\mu^\vee$ are partitions, and that $s_{\lambda/\mu} = s_{\mu^\vee/\lambda^\vee}$.

(b) Show that $c^\lambda_{\mu,\nu} = c^{\mu^\vee}_{\lambda^\vee,\nu}$ for any partition ν.

(c) Let ν be a partition such that $\ell(\nu) \leq k$, and such that all parts of ν are $\leq m$. Let $\nu^\vee$ denote the k-tuple $(m - \nu_k, m - \nu_{k-1}, \ldots, m - \nu_1)$. Show that $\nu^\vee$ is a partition, and satisfies

$$c^\lambda_{\mu,\nu} = c^\lambda_{\nu,\mu} = c^{\mu^\vee}_{\lambda^\vee,\nu} = c^{\mu^\vee}_{\nu,\lambda^\vee} = c^{\nu^\vee}_{\mu,\lambda^\vee} = c^{\nu^\vee}_{\lambda^\vee,\mu}.$$

(d) Show that

$$s_{\lambda^\vee}(x_1, x_2, \ldots, x_k) = (x_1 x_2 \cdots x_k)^m \cdot s_\lambda\left(x_1^{-1}, x_2^{-1}, \ldots, x_k^{-1}\right)$$

in the Laurent polynomial ring $\mathbf{k}\left[x_1, x_2, \ldots, x_k, x_1^{-1}, x_2^{-1}, \ldots, x_k^{-1}\right]$.

(e) Let r be a nonnegative integer. Show that $(r + \lambda_1, r + \lambda_2, \ldots, r + \lambda_k)$ is a partition and satisfies

$$s_{(r+\lambda_1, r+\lambda_2, \ldots, r+\lambda_k)}(x_1, x_2, \ldots, x_k) = (x_1 x_2 \cdots x_k)^r \cdot s_\lambda(x_1, x_2, \ldots, x_k)$$

in the polynomial ring $\mathbf{k}[x_1, x_2, \ldots, x_k]$.

Exercise 2.9.16. Let $m \in \mathbb{N}$, $n \in \mathbb{N}$ and $k \in \mathbb{N}$. Let μ and ν be two partitions such that $\ell(\mu) \leq k$ and $\ell(\nu) \leq k$. Assume that all parts of μ are $\leq m$ (that is, $\mu_i \leq m$ for every positive integer i) [185], and that all parts of ν are $\leq n$ (that is, $\nu_i \leq n$ for every positive integer i). Let $\mu^{\vee\{m\}}$ denote the k-tuple $(m - \mu_k, m - \mu_{k-1}, \ldots, m - \mu_1)$, and let $\nu^{\vee\{n\}}$ denote the k-tuple $(n - \nu_k, n - \nu_{k-1}, \ldots, n - \nu_1)$.

(a) Show that $\mu^{\vee\{m\}}$ and $\nu^{\vee\{n\}}$ are partitions.

Now, let λ be a further partition such that $\ell(\lambda) \leq k$.

(b) If not all parts of λ are $\leq m + n$, then show that $c^\lambda_{\mu,\nu} = 0$.

(c) If all parts of λ are $\leq m+n$, then show that $c^\lambda_{\mu,\nu} = c^{\lambda^{\vee\{m+n\}}}_{\mu^{\vee\{m\}},\nu^{\vee\{n\}}}$, where $\lambda^{\vee\{m+n\}}$ denotes the k-tuple $(m + n - \lambda_k, m + n - \lambda_{k-1}, \ldots, m + n - \lambda_1)$.

The results of Exercise 2.7.11(c) and Exercise 2.9.15(c) are two *symmetries of Littlewood-Richardson coefficients*[186]; combining them yields further such symmetries. While these symmetries were relatively easy consequences of our algebraic definition of the Littlewood-Richardson coefficients, it is a much more challenging task to derive them bijectively from a combinatorial definition of these coefficients (such as the one given in Corollary 2.6.12). Some such derivations appear in [218], in [11], in [16, Example 3.6, Proposition 5.11 and references therein], [73, §5.1, §A.1, §A.4] and [109, (2.12)] (though a different combinatorial interpretation of $c^\lambda_{\mu,\nu}$ is used in the latter three).

Exercise 2.9.17. Recall our usual notations: For every partition λ and every positive integer i, the i-th entry of λ is denoted by λ_i. The sign $\triangleright$ stands for dominance order. We let λ^t denote the conjugate partition of a partition λ.

For any two partitions μ and ν, we define two new partitions $\mu + \nu$ and $\mu \sqcup \nu$ of $|\mu| + |\nu|$ as follows:

- The partition $\mu + \nu$ is defined as $(\mu_1 + \nu_1, \mu_2 + \nu_2, \mu_3 + \nu_3, \ldots)$.
- The partition $\mu \sqcup \nu$ is defined as the result of sorting the list $\left(\mu_1, \mu_2, \ldots, \mu_{\ell(\mu)}, \nu_1, \nu_2, \ldots, \nu_{\ell(\nu)}\right)$ in decreasing order.

(a) Show that any two partitions μ and ν satisfy $(\mu + \nu)^t = \mu^t \sqcup \nu^t$ and $(\mu \sqcup \nu)^t = \mu^t + \nu^t$.

(b) Show that any two partitions μ and ν satisfy $c^{\mu+\nu}_{\mu,\nu} = 1$ and $c^{\mu\sqcup\nu}_{\mu,\nu} = 1$.

(c) If $k \in \mathbb{N}$ and $n \in \mathbb{N}$ satisfy $k \leq n$, and if $\mu \in \operatorname{Par}_k$, $\nu \in \operatorname{Par}_{n-k}$ and $\lambda \in \operatorname{Par}_n$ are such that $c^\lambda_{\mu,\nu} \neq 0$, then prove that $\mu + \nu \triangleright \lambda \triangleright \mu \sqcup \nu$.

(d) If $n \in \mathbb{N}$ and $m \in \mathbb{N}$ and $\alpha, \beta \in \operatorname{Par}_n$ and $\gamma, \delta \in \operatorname{Par}_m$ are such that $\alpha \triangleright \beta$ and $\gamma \triangleright \delta$, then show that $\alpha + \gamma \triangleright \beta + \delta$ and $\alpha \sqcup \gamma \triangleright \beta \sqcup \delta$.

(e) Let $m \in \mathbb{N}$ and $k \in \mathbb{N}$, and let λ be the partition $\left(m^k\right) = \left(\underbrace{m, m, \ldots, m}_{k \text{ times}}\right)$. Show that any two partitions μ and ν satisfy $c^\lambda_{\mu,\nu} \in \{0, 1\}$.

(f) Let $a \in \mathbb{N}$ and $b \in \mathbb{N}$, and let λ be the partition $\left(a + 1, 1^b\right)$ (using the notation of Exercise 2.9.14). Show that any two partitions μ and ν satisfy $c^\lambda_{\mu,\nu} \in \{0, 1\}$.

[185] As usual, we are denoting by ν_i the i-th entry of a partition ν here.

[186] The result of Exercise 2.9.16(c) can also be regarded as a symmetry of Littlewood-Richardson coefficients; see [10, §3.3].

(g) If λ is any partition, and if μ and ν are two rectangular partitions[187], then show that $c^{\lambda}_{\mu,\nu} \in \{0,1\}$.

Exercise 2.9.17(g) is part of Stembridge's [211, Thm. 2.1]; we refer to that article for further results of its kind.

The Littlewood-Richardson rule comes in many different forms, whose equivalence is not always immediate. Our version (Corollary 2.6.12) has the advantage of being the simplest to prove and one of the simplest to state. Other versions can be found in [206, appendix 1 to Ch. 7], Fulton's [73, Ch. 5] and van Leeuwen's [129]. We restrict ourselves to proving some very basic equivalences that allow us to restate parts of Corollary 2.6.12:

Exercise 2.9.18. We shall use the following notations:

- If T is a column-strict tableau and j is a positive integer, then we use the notation $T|_{\mathrm{cols}\geq j}$ for the restriction of T to the union of its columns $j, j+1, j+2, \ldots$. (This notation has already been used in Section 2.6.)
- If T is a column-strict tableau and S is a set of cells of T, then we write $T|_S$ for the restriction of T to the set S of cells.[188]
- If T is a column-strict tableau, then an *NE-set* of T means a set S of cells of T such that whenever $s \in S$, every cell of T which lies northeast[189] of s must also belong to S.
- The *Semitic reading word*[190] of a column-strict tableau T is the concatenation[191] $r_1 r_2 r_3 \cdots$, where r_i is the word obtained by reading the i-th row of T from right to left.[192]
- If $w = (w_1, w_2, \ldots, w_n)$ is a word, then a *prefix* of w means a word of the form $(w_1, w_2, \ldots, w_i)$ for some $i \in \{0, 1, \ldots, n\}$. (In particular, both w and the empty word are prefixes of w.)

  A word w over the set of positive integers is said to be *Yamanouchi* if for any prefix v of w and any positive integer i, there are at least as many i's among the letters of v as there are $(i+1)$'s among them.[193]

Prove the following two statements:

(a) Let μ be a partition. Let $b_{i,j}$ be a nonnegative integer for every two positive integers i and j. Assume that $b_{i,j} = 0$ for all but finitely many pairs (i,j).

The following two assertions are equivalent:

– *Assertion $\mathcal{A}$:* There exist a partition λ and a column-strict tableau T of shape λ/μ such that all $(i,j) \in \{1,2,3,\ldots\}^2$ satisfy

$$b_{i,j} = (\text{the number of all entries } i \text{ in the } j\text{-th row of } T)\,. \tag{2.9.13}$$

– *Assertion $\mathcal{B}$:* The inequality

$$\mu_{j+1} + (b_{1,j+1} + b_{2,j+1} + \cdots + b_{i+1,j+1}) \leq \mu_j + (b_{1,j} + b_{2,j} + \cdots + b_{i,j}) \tag{2.9.14}$$

holds for all $(i,j) \in \mathbb{N} \times \{1,2,3,\ldots\}$.

[187] A partition is called *rectangular* if it has the form $(m^k) = \left(\underbrace{m, m, \ldots, m}_{k \text{ times}}\right)$ for some $m \in \mathbb{N}$ and $k \in \mathbb{N}$.

[188] This restriction $T|_S$ is not necessarily a tableau of skew shape; it is just a map from S to $\{1,2,3,\ldots\}$. The content $\operatorname{cont}(T|_S)$ is nevertheless well-defined (in the usual way: $(\operatorname{cont}(T|_S))_i = \left|(T|_S)^{-1}(i)\right|$).

[189] A cell (r,c) is said to lie *northeast* of a cell (r',c') if and only if we have $r \leq r'$ and $c \geq c'$.

[190] The notation comes from [129] and is a reference to the Arabic and Hebrew way of writing.

[191] If $s_1, s_2, s_3, \ldots$ are several words (finitely or infinitely many), then the *concatenation* $s_1 s_2 s_3 \cdots$ is defined as the word which is obtained by starting with the empty word, then appending s_1 to its end, then appending s_2 to the end of the result, then appending s_3 to the end of the result, etc.

[192] For example, the Semitic reading word of the tableau

$$\begin{array}{cccccc} & 3 & 4 & 4 & 5 \\ 1 & 4 & 6 & & \\ 3 & 5 & & & \end{array}$$

is 544364153.

The Semitic reading word of a tableau T is what is called the *reverse reading word* of T in [206, §A.1.3].

[193] For instance, the words 11213223132 and 1213 are Yamanouchi, while the words 132, 21 and 1121322332111 are not. The Dyck words (defined as in [206, Example 6.6.6], and written using 1's and 2's instead of x's and y's) are precisely the Yamanouchi words whose letters are 1's and 2's and in which the letter 1 appears as often as the letter 2.

Yamanouchi words are often called lattice permutations.

(b) Let λ and μ be two partitions, and let T be a column-strict tableau of shape λ/μ. Then, the following five assertions are equivalent:
  - *Assertion $\mathcal{C}$:* For every positive integer j, the weak composition $\operatorname{cont}\left(T|_{\text{cols}\geq j}\right)$ is a partition.
  - *Assertion $\mathcal{D}$:* For every positive integers j and i, the number of entries $i+1$ in the first j rows[194] of T is $\leq$ to the number of entries i in the first $j-1$ rows of T.
  - *Assertion $\mathcal{E}$:* For every NE-set S of T, the weak composition $\operatorname{cont}\left(T|_S\right)$ is a partition.
  - *Assertion $\mathcal{F}$:* The Semitic reading word of T is Yamanouchi.
  - *Assertion $\mathcal{G}$:* There exists a column-strict tableau S whose shape is a partition and which satisfies the following property: For any positive integers i and j, the number of entries i in the j-th row of T equals the number of entries j in the i-th row of S.

Remark 2.9.19. The equivalence of Assertions $\mathcal{C}$ and $\mathcal{F}$ in Exercise 2.9.18(b) is the "not-too-difficult exercise" mentioned in [210]. It yields the equivalence between our version of the Littlewood-Richardson rule (Corollary 2.6.12) and that in [206, A1.3.3].

In the next exercises, we shall restate Corollary 2.6.11 in a different form. While Corollary 2.6.11 provided a decomposition of the product of a skew Schur function with a Schur function into a sum of Schur functions, the different form that we will encounter in Exercise 2.9.21(b) will give a combinatorial interpretation for the Hall inner product between two skew Schur functions. Let us first generalize Exercise 2.9.18(b):

Exercise 2.9.20. Let us use the notations of Exercise 2.9.18. Let κ, λ and μ be three partitions, and let T be a column-strict tableau of shape λ/μ.

(a) Prove that the following five assertions are equivalent:
  - *Assertion $\mathcal{C}^{(\kappa)}$:* For every positive integer j, the weak composition $\kappa + \operatorname{cont}\left(T|_{\text{cols}\geq j}\right)$ is a partition.
  - *Assertion $\mathcal{D}^{(\kappa)}$:* For every positive integers j and i, we have
$$\kappa_{i+1} + (\text{the number of entries } i+1 \text{ in the first } j \text{ rows of } T)$$
$$\leq \kappa_i + (\text{the number of entries } i \text{ in the first } j-1 \text{ rows of } T).$$
  - *Assertion $\mathcal{E}^{(\kappa)}$:* For every NE-set S of T, the weak composition $\kappa + \operatorname{cont}\left(T|_S\right)$ is a partition.
  - *Assertion $\mathcal{F}^{(\kappa)}$:* For every prefix v of the Semitic reading word of T, and for every positive integer i, we have
$$\kappa_i + (\text{the number of } i\text{'s among the letters of } v)$$
$$\geq \kappa_{i+1} + (\text{the number of } (i+1)\text{'s among the letters of } v).$$
  - *Assertion $\mathcal{G}^{(\kappa)}$:* There exist a partition ζ and a column-strict tableau S of shape ζ/κ which satisfies the following property: For any positive integers i and j, the number of entries i in the j-th row of T equals the number of entries j in the i-th row of S.

(b) Let τ be a partition such that $\tau = \kappa + \operatorname{cont} T$. Consider the five assertions $\mathcal{C}^{(\kappa)}$, $\mathcal{D}^{(\kappa)}$, $\mathcal{E}^{(\kappa)}$, $\mathcal{F}^{(\kappa)}$ and $\mathcal{G}^{(\kappa)}$ introduced in Exercise 2.9.20(a). Let us also consider the following assertion:
  - *Assertion $\mathcal{H}^{(\kappa)}$:* There exists a column-strict tableau S of shape τ/κ which satisfies the following property: For any positive integers i and j, the number of entries i in the j-th row of T equals the number of entries j in the i-th row of S.

  Prove that the six assertions $\mathcal{C}^{(\kappa)}$, $\mathcal{D}^{(\kappa)}$, $\mathcal{E}^{(\kappa)}$, $\mathcal{F}^{(\kappa)}$, $\mathcal{G}^{(\kappa)}$ and $\mathcal{H}^{(\kappa)}$ are equivalent.

Clearly, Exercise 2.9.18(b) is the particular case of Exercise 2.9.20 when $\kappa = \varnothing$.

Using Exercise 2.9.20, we can restate Corollary 2.6.11 in several ways:

Exercise 2.9.21. Let λ, μ and κ be three partitions.

(a) Show that
$$s_\kappa s_{\lambda/\mu} = \sum_T s_{\kappa + \operatorname{cont} T},$$
where the sum ranges over all column-strict tableaux T of shape λ/μ satisfying the five equivalent assertions $\mathcal{C}^{(\kappa)}$, $\mathcal{D}^{(\kappa)}$, $\mathcal{E}^{(\kappa)}$, $\mathcal{F}^{(\kappa)}$ and $\mathcal{G}^{(\kappa)}$ introduced in Exercise 2.9.20(a).

[194] The "first j rows" mean the 1-st row, the 2-nd row, etc., the j-th row (even if some of these rows are empty).

(b) Let τ be a partition. Show that $\left(s_{\lambda/\mu}, s_{\tau/\kappa}\right)_\Lambda$ is the number of all column-strict tableaux T of shape λ/μ satisfying $\tau = \kappa + \operatorname{cont} T$ and also satisfying the six equivalent assertions $\mathcal{C}^{(\kappa)}$, $\mathcal{D}^{(\kappa)}$, $\mathcal{E}^{(\kappa)}$, $\mathcal{F}^{(\kappa)}$, $\mathcal{G}^{(\kappa)}$ and $\mathcal{H}^{(\kappa)}$ introduced in Exercise 2.9.20.

Exercise 2.9.21(a) is merely Corollary 2.6.11, rewritten in light of Exercise 2.9.20. Various parts of it appear in the literature. For instance, [126, (53)] easily reveals to be a restatement of the fact that $s_\kappa s_{\lambda/\mu} = \sum_T s_{\nu + \operatorname{cont} T}$, where the sum ranges over all column-strict tableaux T of shape λ/μ satisfying Assertion $\mathcal{D}^{(\kappa)}$.

Exercise 2.9.21(b) is one version of a "skew Littlewood-Richardson rule" that goes back to Zelevinsky [228] (although Zelevinsky's version uses both a different language and a combinatorial interpretation which is not obviously equivalent to ours). It appears in various sources; for instance, [126, Theorem 5.2, second formula] says that $\left(s_{\lambda/\mu}, s_{\tau/\kappa}\right)_\Lambda$ is the number of all column-strict tableaux T of shape λ/μ satisfying $\tau = \kappa + \operatorname{cont} T$ and the assertion $\mathcal{H}^{(\kappa)}$, whereas [75, Theorem 1.2] says that $\left(s_{\lambda/\mu}, s_{\tau/\kappa}\right)_\Lambda$ is the number of all all column-strict tableaux T of shape λ/μ satisfying $\tau = \kappa + \operatorname{cont} T$ and the assertion $\mathcal{F}^{(\kappa)}$. (Notice that Gasharov's proof of [75, Theorem 1.2] uses the same involutions as Stembridge's proof of Theorem 2.6.6; it can thus be regarded as a close precursor to Stembridge's proof. However, it uses the Jacobi-Trudi identities, while Stembridge's does not.)

Exercise 2.9.22. Let $\mathbb{K}$ be a field.[195] If $N \in \mathbb{K}^{n \times n}$ is a nilpotent matrix, then the *Jordan type* of N is defined to be the list of the sizes of the Jordan blocks in the Jordan normal form of N, sorted in decreasing order[196]. This Jordan type is a partition of n, and uniquely determines N up to similarity (i.e., two nilpotent $n \times n$-matrices N and N' are similar if and only if the Jordan types of N and N' are equal). If f is a nilpotent endomorphism of a finite-dimensional $\mathbb{K}$-vector space V, then we define the *Jordan type* of f as the Jordan type of any matrix representing f (the choice of the matrix does not matter, since the Jordan type of a matrix remains unchanged under conjugation).

(a) Let $n \in \mathbb{N}$. Let $N \in \mathbb{K}^{n \times n}$ be a nilpotent matrix. Let $\lambda \in \operatorname{Par}_n$. Show that the matrix N has Jordan type λ if and only if every $k \in \mathbb{N}$ satisfies

$$\dim\left(\ker\left(N^k\right)\right) = \left(\lambda^t\right)_1 + \left(\lambda^t\right)_2 + \ldots + \left(\lambda^t\right)_k .$$

(Here, we are using the notation λ^t for the transpose of a partition λ, and the notation ν_i for the i-th entry of a partition ν.)

(b) Let f be a nilpotent endomorphism of a finite-dimensional $\mathbb{K}$-vector space V. Let U be an f-stable $\mathbb{K}$-vector subspace of V (that is, a $\mathbb{K}$-vector subspace of V satisfying $f(U) \subset U$). Then, restricting f to U gives a nilpotent endomorphism $f \mid U$ of U, and the endomorphism f also induces a nilpotent endomorphism $\overline{f}$ of the quotient space V/U. Let λ, μ and ν be the Jordan types of f, $f \mid U$ and $\overline{f}$, respectively. Show that $c^\lambda_{\mu,\nu} \neq 0$ (if $\mathbb{Z}$ is a subring of $\mathbf{k}$).

[**Hint:** For (b), Exercise 2.7.11(c) shows that it is enough to prove that $c^{\lambda^t}_{\mu^t,\nu^t} \neq 0$. Due to Corollary 2.6.12, this only requires constructing a column-strict tableau T of shape λ^t/μ^t with $\operatorname{cont} T = \nu^t$ which has the property that each $\operatorname{cont}\left(T|_{\text{cols}\geq j}\right)$ is a partition. Construct this tableau by defining $a_{i,j} = \dim\left(\left(f^i\right)^{-1}(U) \cap \ker\left(f^j\right)\right)$ for all $(i,j) \in \mathbb{N}^2$, and requiring that the number of entries i in the j-th row of T be $a_{i,j} - a_{i,j-1} - a_{i-1,j} + a_{i-1,j-1}$ for all $(i,j) \in \{1,2,3,\ldots\}^2$. Use Exercise 2.9.18(a) to prove that this indeed defines a column-strict tableau, and Exercise 2.9.18(b) to verify that it satisfies the condition on $\operatorname{cont}\left(T|_{\text{cols}\geq j}\right)$.]

Remark 2.9.23. Exercise 2.9.22 is a taste of the connections between the combinatorics of partitions and the Jordan normal form. Much more can, and has, been said. Marc van Leeuwen's [127] is dedicated to some of these connections; in particular, our Exercise 2.9.22(a) is [127, Proposition 1.1], and a far stronger version of Exercise 2.9.22(b) appears in [127, Theorem 4.3 (2)], albeit only for the case of an infinite $\mathbb{K}$. One can prove a converse to Exercise 2.9.22(b) as well: If $c^\lambda_{\mu,\nu} \neq 0$, then there exist V, f and U satisfying the premises of Exercise 2.9.22(b). When $\mathbb{K}$ is a finite field, we can ask enumerative questions, such as how many U's are

[195] This field has no relation to the ring $\mathbf{k}$, over which our symmetric functions are defined.

[196] The Jordan normal form of N is well-defined even if $\mathbb{K}$ is not algebraically closed, because N is nilpotent (so the characteristic polynomial of N is X^n).

there for given V, f, λ, μ and ν; we will see a few answers in Section 4.9 (specifically, Proposition 4.9.4), and a more detailed treatment is given in [142, Ch. 2].

The relationship between partitions and Jordan normal forms can be exploited to provide linear-algebraic proofs of purely combinatorial facts. See [28, Sections 6 and 9] for some examples. Note that [28, Lemma 9.10] is the statement that, under the conditions of Exercise 2.9.22(b), we have $\nu \subseteq \lambda$. This is a direct consequence of Exercise 2.9.22(b) (since $c^{\lambda}_{\mu,\nu} \neq 0$ can happen only if $\nu \subseteq \lambda$).

Exercise 2.9.24. Let $a \in \Lambda$. Prove the following:

(a) The set $\left\{ g \in \Lambda \ \mid \ g^{\perp} a = (\omega(g))^{\perp} a \right\}$ is a **k**-subalgebra of Λ.

(b) Assume that $e_k^{\perp} a = h_k^{\perp} a$ for each positive integer k. Then, $g^{\perp} a = (\omega(g))^{\perp} a$ for each $g \in \Lambda$.

Exercise 2.9.25. Let $n \in \mathbb{N}$. Let ρ be the partition $(n-1, n-2, \ldots, 1)$. Prove that $s_{\rho/\mu} = s_{\rho/\mu^t}$ for every $\mu \in \operatorname{Par}$.

Remark 2.9.26. Exercise 2.9.25 appears in [180, Corollary 7.32], and is due to John Stembridge. Using Remark 2.5.9, we can rewrite it as yet another equality between Littlewood-Richardson coefficients: Namely, $c^{\rho}_{\mu,\nu} = c^{\rho}_{\mu^t,\nu}$ for any $\mu \in \operatorname{Par}$ and $\nu \in \operatorname{Par}$.

3. Zelevinsky's structure theory of positive self-dual Hopf algebras

Chapter 2 showed that, as a $\mathbb{Z}$-basis for the Hopf algebra $\Lambda = \Lambda_{\mathbb{Z}}$, the Schur functions $\{s_\lambda\}$ have two special properties: they have the *same* structure constants $c^\lambda_{\mu,\nu}$ for their multiplication as for their comultiplication (Corollary 2.5.7), and these structure constants are all *nonnegative* integers (Corollary 2.6.12). Zelevinsky [227, §2,3] isolated these two properties as crucial.

Definition 3.0.1. Say that a connected graded Hopf algebra A over $\mathbf{k} = \mathbb{Z}$ with a distinguished $\mathbb{Z}$-basis $\{\sigma_\lambda\}$ consisting of homogeneous elements[197] is a *positive self-dual Hopf algebra* (or *PSH*) if it satisfies the two further axioms

- **(self-duality)** The same structure constants $a^\lambda_{\mu,\nu}$ appear for the product $\sigma_\mu \sigma_\nu = \sum_\lambda a^\lambda_{\mu,\nu} \sigma_\lambda$ and the coproduct $\Delta \sigma_\lambda = \sum_{\mu,\nu} a^\lambda_{\mu,\nu} \sigma_\mu \otimes \sigma_\nu$.
- **(positivity)** The $a^\lambda_{\mu,\nu}$ are all nonnegative (integers).

Call $\{\sigma_\lambda\}$ the *PSH-basis* of A.

He then developed a beautiful structure theory for PSH's, explaining how they can be uniquely expressed as tensor products of copies of PSH's each isomorphic to Λ after rescaling their grading. The next few sections explain this, following his exposition closely.

3.1. Self-duality implies polynomiality.

We begin with a property that forces a Hopf algebra to have algebra structure which is a *polynomial* algebra, specifically the symmetric algebra $\mathrm{Sym}(\mathfrak{p})$, where $\mathfrak{p}$ is the $\mathbf{k}$-submodule of primitive elements.

Recall from Exercise 1.3.20(g) that for a connected graded Hopf algebra $A = \bigoplus_{n=0}^\infty A_n$, every x in the two-sided ideal $I := \ker \epsilon = \bigoplus_{n>0} A_n$ has the property that its comultiplication takes the form

$$\Delta(x) = 1 \otimes x + x \otimes 1 + \Delta_+(x)$$

where $\Delta_+(x)$ lies in $I \otimes I$. Recall also that the elements x for which $\Delta_+(x) = 0$ are called the *primitives*. Denote by $\mathfrak{p}$ the $\mathbf{k}$-submodule of primitive elements inside A.

Given a PSH A (over $\mathbf{k} = \mathbb{Z}$) with a PSH-basis $\{\sigma_\lambda\}$, we consider the bilinear form $(\cdot,\cdot)_A : A \times A \to \mathbb{Z}$ on A that makes this basis orthonormal. Similarly, the elements $\{\sigma_\lambda \otimes \sigma_\mu\}$ give an orthonormal basis for a form $(\cdot,\cdot)_{A\otimes A}$ on $A \otimes A$. The bilinear form $(\cdot,\cdot)_A$ on the PSH A gives rise to a $\mathbb{Z}$-linear map $A \to A^o$, which is easily seen to be injective and a $\mathbb{Z}$-algebra homomorphism. We thus identify A with a subalgebra of A^o. When A is of finite type, this map is a Hopf algebra isomorphism, thus allowing us to identify A with A^o. This is an instance of the following notion of self-duality.

Definition 3.1.1. (a) If $(\cdot,\cdot) : V \times W \to \mathbf{k}$ is a bilinear form on the product $V \times W$ of two graded $\mathbf{k}$-modules $V = \bigoplus_{n\geq 0} V_n$ and $W = \bigoplus_{n \geq 0} W_n$, then we say that this form $(\cdot,\cdot)$ is *graded* if every two distinct nonnegative integers n and m satisfy $(V_n, W_m) = 0$ (that is, if every two homogeneous elements $v \in V$ and $w \in W$ having distinct degrees satisfy $(v,w) = 0$).

(b) If $(\cdot,\cdot)_V : V \times V \to \mathbf{k}$ and $(\cdot,\cdot)_W : W \times W \to \mathbf{k}$ are two symmetric bilinear forms on some $\mathbf{k}$-modules V and W, then we can canonically define a symmetric bilinear form $(\cdot,\cdot)_{V\otimes W}$ on the $\mathbf{k}$-module $V \otimes W$ by letting

$$(v \otimes w, v' \otimes w')_{V\otimes W} = (v, v')_V \, (w, w')_W \qquad \text{for all } v, v' \in V \text{ and } w, w' \in W.$$

This new bilinear form is graded if the original two forms $(\cdot,\cdot)_V$ and $(\cdot,\cdot)_W$ were graded (presuming that V and W are graded).

(c) Say that a bialgebra A is *self-dual* with respect to a given symmetric bilinear form $(\cdot,\cdot) : A \times A \to \mathbf{k}$ if one has $(a, m(b \otimes c))_A = (\Delta(a), b \otimes c)_{A\otimes A}$ and $(1_A, a) = \epsilon(a)$ for a, b, c in A. If A is a graded Hopf algebra of finite type, and this form $(\cdot,\cdot)$ is graded, then this is equivalent to the $\mathbf{k}$-module map $A \to A^o$ induced by $(\cdot,\cdot)_A$ giving a Hopf algebra homomorphism.

Thus, any PSH A is self-dual with respect to the bilinear form $(\cdot,\cdot)_A$ that makes its PSH-basis orthonormal.

Notice also that the injective $\mathbb{Z}$-algebra homomorphism $A \to A^o$ obtained from the bilinear form $(\cdot,\cdot)_A$ on a PSH A allows us to regard each $f \in A$ as an element of A^o. Thus, for any PSH A and any $f \in A$, an operator $f^\perp : A \to A$ is well-defined (indeed, regard f as an element of A^o, and apply Definition 2.8.1).

[197] not necessarily indexed by partitions

Proposition 3.1.2. *Let A be a Hopf algebra over $\mathbf{k} = \mathbb{Z}$ or $\mathbf{k} = \mathbb{Q}$ which is graded, connected, and self-dual with respect to a positive definite graded*[198] *bilinear form. Then:*

(a) *Within the ideal I, the $\mathbf{k}$-submodule of primitives $\mathfrak{p}$ is the orthogonal complement to the $\mathbf{k}$-submodule I^2.*

(b) *In particular, $\mathfrak{p} \cap I^2 = 0$.*

(c) *When $\mathbf{k} = \mathbb{Q}$, one has $I = \mathfrak{p} \oplus I^2$.*

Proof. (a) Note that $I^2 = m(I \otimes I)$. Hence an element x in I lies in the perpendicular space to I^2 if and only if one has for all y in $I \otimes I$ that

$$0 = (x, m(y))_A = (\Delta(x), y)_{A \otimes A} = (\Delta_+(x), y)_{A \otimes A}$$

where the second equality uses self-duality, while the third equality uses the fact that y lies in $I \otimes I$ and the form $(\cdot, \cdot)_{A \otimes A}$ makes distinct homogeneous components orthogonal. Since y was arbitrary, this means x is perpendicular to I^2 if and only if $\Delta_+(x) = 0$, that is, x lies in $\mathfrak{p}$.

(b) This follows from (a), since the form $(\cdot, \cdot)_A$ is positive definite.

(c) This follows from (a) using some basic linear algebra[199] when A is of finite type (which is the only case we will ever encounter in practice). See Exercise 3.1.6 for the general proof. □

Remark 3.1.3. One might wonder why we didn't just say $I = \mathfrak{p} \oplus I^2$ even when $\mathbf{k} = \mathbb{Z}$ in Proposition 3.1.2(c). However, this is false even for $A = \Lambda_{\mathbb{Z}}$: the second homogeneous component $(\mathfrak{p} \oplus I^2)_2$ is the index 2 sublattice of Λ_2 which is $\mathbb{Z}$-spanned by $\{p_2, e_1^2\}$, containing $2e_2$, but not containing e_2 itself.

Already the fact that $\mathfrak{p} \cap I^2 = 0$ has a strong implication.

Lemma 3.1.4. *A connected graded Hopf algebra A over any ring $\mathbf{k}$ having $\mathfrak{p} \cap I^2 = 0$ must necessarily be commutative (as an algebra).*

Proof. The component $A_0 = \mathbf{k}$ commutes with all of A. This forms the base case for an induction on $i + j$ in which one shows that any elements x in A_i and y in A_j with $i, j > 0$ will have $[x, y] := xy - yx = 0$. Since $[x, y]$ lies in I^2, it suffices to show that $[x, y]$ also lies in $\mathfrak{p}$:

$$\begin{aligned}
\Delta[x, y] &= [\Delta(x), \Delta(y)] \\
&= [1 \otimes x + x \otimes 1 + \Delta_+(x), 1 \otimes y + y \otimes 1 + \Delta_+(y)] \\
&= [1 \otimes x + x \otimes 1, 1 \otimes y + y \otimes 1] \\
&\qquad + [1 \otimes x + x \otimes 1, \Delta_+(y)] + [\Delta_+(x), 1 \otimes y + y \otimes 1] + [\Delta_+(x), \Delta_+(y)] \\
&= [1 \otimes x + x \otimes 1, 1 \otimes y + y \otimes 1] \\
&= 1 \otimes [x, y] + [x, y] \otimes 1
\end{aligned}$$

showing that $[x, y]$ lies in $\mathfrak{p}$. Here the second-to-last equality used the inductive hypotheses: homogeneity implies that $\Delta_+(x)$ is a sum of homogeneous tensors of the form $z_1 \otimes z_2$ satisfying $\deg(z_1), \deg(z_2) < i$, so that by induction they will commute with $1 \otimes y, y \otimes 1$, thus proving that $[\Delta_+(x), 1 \otimes y + y \otimes 1] = 0$; a symmetric argument shows $[1 \otimes x + x \otimes 1, \Delta_+(y)] = 0$, and a similar argument shows $[\Delta_+(x), \Delta_+(y)] = 0$. The last equality is an easy calculation, and was done already in the process of proving (1.3.7). □

Remark 3.1.5. Zelevinsky actually shows [227, Proof of A.1.3, p. 150] that the assumption of $\mathfrak{p} \cap I^2 = 0$ (along with hypotheses of unit, counit, graded, connected, and Δ being a morphism for multiplication) already implies the *associativity* of the multiplication in A ! One shows by induction on $i + j + k$ that any x, y, z in A_i, A_j, A_k with $i, j, k > 0$ have vanishing *associator* $\operatorname{assoc}(x, y, z) := x(yz) - (xy)z$. In the inductive step, one first notes that $\operatorname{assoc}(x, y, z)$ lies in I^2, and then checks that $\operatorname{assoc}(x, y, z)$ also lies in $\mathfrak{p}$, by a calculation very similar to the one above, repeatedly using the fact that $\operatorname{assoc}(x, y, z)$ is multilinear in its three arguments.

Exercise 3.1.6. Prove Proposition 3.1.2(c) in the general case.

[198] That is, $(A_i, A_j) = 0$ for $i \neq j$.

[199] Specifically, either the existence of an orthogonal projection on a subspace of a finite-dimensional inner-product space over $\mathbb{Q}$, or the fact that $\dim\left(W^{\perp}\right) = \dim V - \dim W$ for a subspace W of a finite-dimensional inner-product space V over $\mathbb{Q}$ can be used.

This leads to a general structure theorem.

Theorem 3.1.7. *If a connected graded Hopf algebra A over a field $\mathbf{k}$ of characteristic zero has $I = \mathfrak{p} \oplus I^2$, then the inclusion $\mathfrak{p} \hookrightarrow A$ extends to a Hopf algebra isomorphism from the symmetric algebra $\operatorname{Sym}_{\mathbf{k}}(\mathfrak{p}) \to A$. In particular, A is both commutative and cocommutative.*

Note that the hypotheses of Theorem 3.1.7 are valid, using Proposition 3.1.2(c), whenever A is obtained from a PSH (over $\mathbb{Z}$) by tensoring with $\mathbb{Q}$.

Proof of Theorem 3.1.7. Since Lemma 3.1.4 implies that A is commutative, the universal property of $\operatorname{Sym}_{\mathbf{k}}(\mathfrak{p})$ as a free commutative algebra on generators $\mathfrak{p}$ shows that the inclusion $\mathfrak{p} \hookrightarrow A$ at least extends to an algebra morphism $\operatorname{Sym}_{\mathbf{k}}(\mathfrak{p}) \xrightarrow{\varphi} A$. Since the Hopf structure on $\operatorname{Sym}_{\mathbf{k}}(\mathfrak{p})$ makes the elements of $\mathfrak{p}$ primitive (see Example 1.3.14), this φ is actually a coalgebra morphism (since $\Delta \circ \varphi = (\varphi \otimes \varphi) \circ \Delta$ and $\epsilon \circ \varphi = \epsilon$ need only to be checked on algebra generators), hence a bialgebra morphism, hence a Hopf algebra morphism (by Corollary 1.4.27). It remains to show that φ is surjective, and injective.

For the surjectivity of φ, note that the hypothesis $I = \mathfrak{p} \oplus I^2$ implies that the composite $\mathfrak{p} \hookrightarrow I \to I/I^2$ gives a $\mathbf{k}$-vector space isomorphism. What follows is a standard argument to deduce that $\mathfrak{p}$ generates A as a commutative graded $\mathbf{k}$-algebra. One shows by induction on n that any homogeneous element a in A_n lies in the $\mathbf{k}$-subalgebra generated by $\mathfrak{p}$. The base case $n = 0$ is trivial as a lies in $A_0 = \mathbf{k} \cdot 1_A$. In the inductive step where a lies in I, write $a \equiv p \bmod I^2$ for some p in $\mathfrak{p}$. Thus $a = p + \sum_i b_i c_i$, where b_i, c_i lie in I but have strictly smaller degree, so that by induction they lie in the subalgebra generated by $\mathfrak{p}$, and hence so does a.

Note that the surjectivity argument did not use the assumption that $\mathbf{k}$ has characteristic zero, but we will now use it in the injectivity argument for φ, to establish the following

(3.1.1) **Claim:** Every primitive element of $\operatorname{Sym}(\mathfrak{p})$ lies in $\mathfrak{p} = \operatorname{Sym}^1(\mathfrak{p})$.

Note that this claim fails in positive characteristic, e.g. if $\mathbf{k}$ has characteristic 2 then x^2 lies in $\operatorname{Sym}^2(\mathfrak{p})$, however

$$\Delta(x^2) = 1 \otimes x^2 + 2x \otimes x + x^2 \otimes 1 = 1 \otimes x^2 + x^2 \otimes 1.$$

To prove the claim (3.1.1), assume not, so that by gradedness, there must exist some primitive element $y \neq 0$ lying in some $\operatorname{Sym}^n(\mathfrak{p})$ with $n \geq 2$. This would mean that $f(y) = 0$, where the map f is defined as the composition

$$\operatorname{Sym}^n(\mathfrak{p}) \xrightarrow{\Delta} \bigoplus_{i+j=n} \operatorname{Sym}^i(\mathfrak{p}) \otimes \operatorname{Sym}^j(\mathfrak{p}) \xrightarrow{\text{projection}} \operatorname{Sym}^1(\mathfrak{p}) \otimes \operatorname{Sym}^{n-1}(\mathfrak{p})$$

of the coproduct Δ with the component projection of $\bigoplus_{i+j=n} \operatorname{Sym}^i(\mathfrak{p}) \otimes \operatorname{Sym}^j(\mathfrak{p})$ onto $\operatorname{Sym}^1(\mathfrak{p}) \otimes \operatorname{Sym}^{n-1}(\mathfrak{p})$. However, one can check on a basis that the multiplication backward $\operatorname{Sym}^1(\mathfrak{p}) \otimes \operatorname{Sym}^{n-1}(\mathfrak{p}) \xrightarrow{m} \operatorname{Sym}^n(\mathfrak{p})$ has the property that $m \circ f = n \cdot \mathrm{id}_{\operatorname{Sym}^n(\mathfrak{p})}$: Indeed,

$$(m \circ f)(x_1 \cdots x_n) = m\left(\sum_{j=1}^{n} x_j \otimes x_1 \cdots \widehat{x_j} \cdots x_n\right) = n \cdot x_1 \cdots x_n$$

for $x_1, \ldots, x_n$ in $\mathfrak{p}$. Then $n \cdot y = m(f(y)) = m(0) = 0$ leads to the contradiction that $y = 0$, since $\mathbf{k}$ has characteristic zero. Thus, (3.1.1) is proven.

Now one can argue the injectivity of the (graded) map[200] φ by assuming that one has a nonzero homogeneous element u in $\ker(\varphi)$ of minimum degree. In particular, $\deg(u) \geq 1$. Also since $\mathfrak{p} \hookrightarrow A$, one has that u is not in $\operatorname{Sym}^1(\mathfrak{p}) = \mathfrak{p}$, and hence u is not primitive by (3.1.1). Consequently $\Delta_+(u) \neq 0$, and one can find a nonzero component $u^{(i,j)}$ of $\Delta_+(u)$ lying in $\operatorname{Sym}(\mathfrak{p})_i \otimes \operatorname{Sym}(\mathfrak{p})_j$ for some $i, j > 0$. Since this forces $i, j < \deg(u)$, one has that φ maps both $\operatorname{Sym}(\mathfrak{p})_i, \operatorname{Sym}(\mathfrak{p})_j$ injectively into A_i, A_j. Hence the tensor product map

$$\operatorname{Sym}(\mathfrak{p})_i \otimes \operatorname{Sym}(\mathfrak{p})_j \xrightarrow{\varphi \otimes \varphi} A_i \otimes A_j$$

[200] The grading on $\operatorname{Sym}(\mathfrak{p})$ is induced from the grading on $\mathfrak{p}$, a homogeneous subspace of $I \subset A$ as it is the kernel of the graded map $I \xrightarrow{\Delta_+} A \otimes A$.

is also injective[201]. This implies $(\varphi \otimes \varphi)(u^{(i,j)}) \neq 0$, giving the contradiction that

$$0 = \Delta_+^A(0) = \Delta_+^A(\varphi(u)) = (\varphi \otimes \varphi)(\Delta_+^{\operatorname{Sym}(\mathfrak{p})}(u))$$

contains the nonzero $A_i \otimes A_j$-component $(\varphi \otimes \varphi)(u^{(i,j)})$.

(An alternative proof of the injectivity of φ proceeds as follows: By (3.1.1), the subspace of primitive elements of $\operatorname{Sym}(\mathfrak{p})$ is $\mathfrak{p}$, and clearly $\varphi\mid_{\mathfrak{p}}$ is injective. Hence, Exercise 1.4.35(b) (applied to the homomorphism φ) shows that φ is injective.) □

Before closing this section, we mention one nonobvious corollary of the Claim (3.1.1), when applied to the ring of symmetric functions $\Lambda_{\mathbb{Q}}$ with $\mathbb{Q}$-coefficients, since Proposition 2.4.1 says that $\Lambda_{\mathbb{Q}} = \mathbb{Q}[p_1, p_2, \ldots] = \operatorname{Sym}(V)$ where $V = \mathbb{Q}\{p_1, p_2, \ldots\}$.

Corollary 3.1.8. *The subspace $\mathfrak{p}$ of primitives in $\Lambda_{\mathbb{Q}}$ is one-dimensional in each degree $n = 1, 2, \ldots$, and spanned by $\{p_1, p_2, \ldots\}$.*

We note in passing that this corollary can also be obtained in a simpler fashion and a greater generality:

Exercise 3.1.9. Let $\mathbf{k}$ be any commutative ring. Show that the primitive elements of Λ are precisely the elements of the $\mathbf{k}$-linear span of $p_1, p_2, p_3, \ldots$.

3.2. **The decomposition theorem.**

Our goal here is Zelevinsky's theorem [227, Theorem 2.2] giving a canonical decomposition of any PSH as a tensor product into PSH's that each have only one primitive element in their PSH-basis. For the sake of stating it, we introduce some notation.

Definition 3.2.1. Given a PSH A with PSH-basis Σ, let $\mathcal{C} := \Sigma \cap \mathfrak{p}$ be the primitive elements in Σ. For each ρ in $\mathcal{C}$, let $A(\rho) \subset A$ be the $\mathbb{Z}$-span of

$$\Sigma(\rho) := \{\sigma \in \Sigma : \text{ there exists } n \geq 0 \text{ with } (\sigma, \rho^n) \neq 0\}.$$

Definition 3.2.2. The tensor product of two PSHs A_1 and A_2 with PSH-bases Σ_1 and Σ_2 is defined as the graded Hopf algebra $A_1 \otimes A_2$ with PSH-basis $\{\sigma_1 \otimes \sigma_2\}_{(\sigma_1, \sigma_2) \in \Sigma_1 \times \Sigma_2}$. It is easy to see that this is again a PSH. The tensor product of any finite family of PSHs is defined similarly[202].

Theorem 3.2.3. *Any PSH A has a canonical tensor product decomposition*

$$A = \bigotimes_{\rho \in \mathcal{C}} A(\rho)$$

with $A(\rho)$ a PSH, and ρ the only primitive element in its PSH-basis $\Sigma(\rho)$.

Although in all the applications, $\mathcal{C}$ will be finite, when $\mathcal{C}$ is infinite one should interpret the tensor product in the theorem as the inductive limit of tensor products over finite subsets of $\mathcal{C}$, that is, linear combinations of basic tensors $\bigotimes_\rho a_\rho$ in which there are only finitely many factors $a_\rho \neq 1$.

The first step toward the theorem uses a certain unique factorization property.

Lemma 3.2.4. *Let $\mathcal{P}$ be a set of pairwise orthogonal primitives in a PSH A. Then,*

$$(\rho_1 \cdots \rho_r, \pi_1 \cdots \pi_s) = 0$$

for ρ_i, π_j in $\mathcal{P}$ unless $r = s$ and one can reindex so that $\rho_i = \pi_i$.

[201] One needs to know that for two injective maps $V_i \overset{\varphi_i}{\hookrightarrow} W_i$ of $\mathbf{k}$-vector spaces V_i, W_i with $i = 1, 2$, the tensor product $\varphi_1 \otimes \varphi_2$ is also injective. Factoring it as $\varphi_1 \otimes \varphi_2 = (\mathrm{id} \otimes \varphi_2) \circ (\varphi_1 \otimes \mathrm{id})$, one sees that it suffices to show that for an injective map $V \overset{\varphi}{\hookrightarrow} W$ of free $\mathbf{k}$-modules, and any free $\mathbf{k}$-module U, the map $V \otimes U \xrightarrow{\varphi \otimes \mathrm{id}} W \otimes U$ is also injective. Since tensor products commute with direct sums, and U is (isomorphic to) a direct sum of copies of $\mathbf{k}$, this reduces to the easy-to-check case where $U = \mathbf{k}$.

Note that some kind of freeness or flatness hypothesis on U is needed here since, e.g. the injective $\mathbb{Z}$-module maps $\mathbb{Z} \xrightarrow{\varphi_1 = (\cdot \times 2)} \mathbb{Z}$ and $\mathbb{Z}/2\mathbb{Z} \xrightarrow{\varphi_2 = \mathrm{id}} \mathbb{Z}/2\mathbb{Z}$ have $\varphi_1 \otimes \varphi_2 = 0$ on $\mathbb{Z} \otimes_{\mathbb{Z}} \mathbb{Z}/2\mathbb{Z} \cong \mathbb{Z}/2\mathbb{Z} \neq 0$.

[202] For the empty family, it is the connected graded Hopf algebra $\mathbb{Z}$ with PSH-basis $\{1\}$.

Proof. Induct on r. For $r > 0$, one has

$$\begin{aligned}(\rho_1 \cdots \rho_r, \pi_1 \cdots \pi_s) &= (\rho_2 \cdots \rho_r, \rho_1^{\perp}(\pi_1 \cdots \pi_s)) \\ &= (\rho_2 \cdots \rho_r, \sum_{j=1}^{s} (\pi_1 \cdots \pi_{j-1} \cdot \rho_1^{\perp}(\pi_j) \cdot \pi_{j+1} \cdots \pi_s))\end{aligned}$$

from Proposition 2.8.2(iv) because ρ_1 is primitive[203]. On the other hand, since each π_j is primitive, one has $\rho_1^{\perp}(\pi_j) = (\rho_1, 1) \cdot \pi_j + (\rho_1, \pi_j) \cdot 1 = (\rho_1, \pi_j)$ which vanishes unless $\rho_1 = \pi_j$. Hence $(\rho_1 \cdots \rho_r, \pi_1 \cdots \pi_s) = 0$ unless $\rho_1 \in \{\pi_1, \ldots, \pi_s\}$, in which case after reindexing so that $\pi_1 = \rho_1$, it equals

$$n \cdot (\rho_1, \rho_1) \cdot (\rho_2 \cdots \rho_r, \pi_2 \cdots \pi_s)$$

if there are exactly n occurrences of ρ_1 among $\pi_1, \ldots, \pi_s$. Now apply induction. □

So far the positivity hypothesis for a PSH has played little role. Now we use it to introduce a certain partial order on the PSH A, and then a semigroup grading.

Definition 3.2.5. For a subset S of an abelian group, let $\mathbb{Z}S$ (resp. $\mathbb{N}S$) denote the subgroup of $\mathbb{Z}$-linear combinations (resp. submonoid of $\mathbb{N}$-linear combinations[204]) of the elements of S.

In a PSH A with PSH-basis Σ, the subset $\mathbb{N}\Sigma$ forms a submonoid, and lets one define a partial order on A via $a \leq b$ if $b - a$ lies in $\mathbb{N}\Sigma$.

We note a few trivial properties of this partial order:

- The positivity hypothesis implies that $\mathbb{N}\Sigma \cdot \mathbb{N}\Sigma \subset \mathbb{N}\Sigma$.
- Hence multiplication by an element $c \geq 0$ (meaning c lies in $\mathbb{N}\Sigma$) preserves the order: $a \leq b$ implies $ac \leq bc$ since $(b - a)c$ lies in $\mathbb{N}\Sigma$.
- Thus $0 \leq c \leq d$ and $0 \leq a \leq b$ together imply $ac \leq bc \leq bd$.

This allows one to introduce a semigroup grading on A.

Definition 3.2.6. Let $\mathbb{N}^{\mathcal{C}}_{\text{fin}}$ denote the additive submonoid of $\mathbb{N}^{\mathcal{C}}$ consisting of those $\alpha = (\alpha_\rho)_{\rho \in \mathcal{C}}$ with finite support.

Note that for any α in $\mathbb{N}^{\mathcal{C}}_{\text{fin}}$, one has that the product $\prod_{\rho \in \mathcal{C}} \rho^{\alpha_\rho} \geq 0$. Define

$$\Sigma(\alpha) := \{\sigma \in \Sigma : \sigma \leq \prod_{\rho \in \mathcal{C}} \rho^{\alpha_\rho}\},$$

that is, the subset of Σ on which $\prod_{\rho \in \mathcal{C}} \rho^{\alpha_\rho}$ has support. Also define

$$A_{(\alpha)} := \mathbb{Z}\Sigma(\alpha) \subset A.$$

Proposition 3.2.7. *The PSH A has an $\mathbb{N}^{\mathcal{C}}_{\text{fin}}$-semigroup-grading: one has an orthogonal direct sum decomposition*

$$A = \bigoplus_{\alpha \in \mathbb{N}^{\mathcal{C}}_{\text{fin}}} A_{(\alpha)}$$

for which

$$A_{(\alpha)} A_{(\beta)} \subset A_{(\alpha+\beta)}, \tag{3.2.1}$$

$$\Delta A_{(\alpha)} \subset \bigoplus_{\alpha = \beta + \gamma} A_{(\beta)} \otimes A_{(\gamma)}. \tag{3.2.2}$$

Proof. We will make free use of the fact that a PSH A is commutative, since it embeds in $A \otimes_{\mathbb{Z}} \mathbb{Q}$, which is commutative by Theorem 3.1.7.

Note that the orthogonality $\left(A_{(\alpha)}, A_{(\beta)}\right) = 0$ for $\alpha \neq \beta$ is equivalent to the assertion that

$$\left(\prod_{\rho \in \mathcal{C}} \rho^{\alpha_\rho}, \prod_{\rho \in \mathcal{C}} \rho^{\beta_\rho} \right) = 0,$$

[203] Strictly speaking, this argument needs further justification since A might not be of finite type (and if it is not, Proposition 2.8.2(iv) cannot be applied). It is more adequate to refer to the proof of Proposition 2.8.2(iv), which indeed goes through with ρ_1 taking the role of f.

[204] Recall that $\mathbb{N} := \{0, 1, 2, \ldots\}$.

which follows from Lemma 3.2.4.

Next let us deal with the assertion (3.2.1). It suffices to check that when τ, ω in Σ lie in $A_{(\alpha)}, A_{(\beta)}$, respectively, then $\tau\omega$ lies in $A_{(\alpha+\beta)}$. But note that any σ in Σ having $\sigma \leq \tau\omega$ will then have

$$\sigma \leq \tau\omega \leq \prod_{\rho\in\mathcal{C}} \rho^{\alpha_\rho} \cdot \prod_{\rho\in\mathcal{C}} \rho^{\beta_\rho} = \prod_{\rho\in\mathcal{C}} \rho^{\alpha_\rho+\beta_\rho}$$

so that σ lies in $A_{(\alpha+\beta)}$. This means that $\tau\omega$ lies in $A_{(\alpha+\beta)}$.

This lets us check that $\bigoplus_{\alpha\in\mathbb{N}^{\mathcal{C}}_{\text{fin}}} A_{(\alpha)}$ exhaust A. It suffices to check that any σ in Σ lies in some $A_{(\alpha)}$. Proceed by induction on $\deg(\sigma)$, with the case $\sigma = 1$ being trivial; the element 1 always lies in Σ, and hence lies in $A_{(\alpha)}$ for $\alpha = 0$. For σ lying in I, one either has $(\sigma, a) \neq 0$ for some a in I^2, or else σ lies in $(I^2)^\perp = \mathfrak{p}$ (by Proposition 3.1.2(a)), so that σ is in $\mathcal{C}$ and we are done. If $(\sigma, a) \neq 0$ with a in I^2, then σ appears in the support of some $\mathbb{Z}$-linear combination of elements $\tau\omega$ where τ, ω lie in Σ and have strictly smaller degree than σ has. There exists at least one such pair τ, ω for which $(\sigma, \tau\omega) \neq 0$, and therefore $\sigma \leq \tau\omega$. Then by induction τ, ω lie in some $A_{(\alpha)}, A_{(\beta)}$, respectively, so $\tau\omega$ lies in $A_{(\alpha+\beta)}$, and hence σ lies in $A_{(\alpha+\beta)}$ also.

Self-duality shows that (3.2.1) implies (3.2.2): if a, b, c lie in $A_{(\alpha)}, A_{(\beta)}, A_{(\gamma)}$, respectively, then $(\Delta a, b \otimes c)_{A\otimes A} = (a, bc)_A = 0$ unless $\alpha = \beta + \gamma$. □

Proposition 3.2.8. *For α, β in $\mathbb{N}^{\mathcal{C}}_{\text{fin}}$ with disjoint support, one has a bijection*

$$\begin{array}{ccc} \Sigma(\alpha) \times \Sigma(\beta) & \longrightarrow & \Sigma(\alpha+\beta), \\ (\sigma, \tau) & \longmapsto & \sigma\tau. \end{array}$$

Thus, the multiplication map $A_{(\alpha)} \otimes A_{(\beta)} \to A_{(\alpha+\beta)}$ is an isomorphism.

Proof. We first check that for σ_1, σ_2 in $\Sigma(\alpha)$ and τ_1, τ_2 in $\Sigma(\beta)$, one has

$$(\sigma_1\tau_1, \sigma_2\tau_2) = \delta_{(\sigma_1,\tau_1),(\sigma_2,\tau_2)}. \tag{3.2.3}$$

Note that this is equivalent to showing both

- that $\sigma\tau$ lie in $\Sigma(\alpha+\beta)$ so that the map is well-defined, since it shows $(\sigma\tau, \sigma\tau) = 1$, and
- that the map is injective.

One calculates

$$\begin{aligned} (\sigma_1\tau_1, \sigma_2\tau_2)_A &= (\sigma_1\tau_1, m(\sigma_2 \otimes \tau_2))_A \\ &= (\Delta(\sigma_1\tau_1), \sigma_2 \otimes \tau_2)_{A\otimes A} \\ &= (\Delta(\sigma_1)\Delta(\tau_1), \sigma_2 \otimes \tau_2)_{A\otimes A}. \end{aligned}$$

Note that due to (3.2.2), $\Delta(\sigma_1)\Delta(\tau_1)$ lies in $\sum A_{(\alpha'+\beta')} \otimes A_{(\alpha''+\beta'')}$ where

$$\begin{aligned} \alpha' + \alpha'' &= \alpha, \\ \beta' + \beta'' &= \beta. \end{aligned}$$

Since $\sigma_2 \otimes \tau_2$ lies in $A_{(\alpha)} \otimes A_{(\beta)}$, the only nonvanishing terms in the inner product come from those with

$$\begin{aligned} \alpha' + \beta' &= \alpha, \\ \alpha'' + \beta'' &= \beta. \end{aligned}$$

As α, β have disjoint support, this can only happen if

$$\alpha' = \alpha, \ \alpha'' = 0, \ \beta' = 0, \ \beta'' = \beta;$$

that is, the only nonvanishing term comes from $(\sigma_1 \otimes 1)(1 \otimes \tau_1) = \sigma_1 \otimes \tau_1$. Hence

$$(\sigma_1\tau_1, \sigma_2\tau_2)_A = (\sigma_1 \otimes \tau_1, \sigma_2 \otimes \tau_2)_{A\otimes A} = \delta_{(\sigma_1,\tau_1),(\sigma_2,\tau_2)}.$$

To see that the map is surjective, express

$$\begin{aligned} \prod_{\rho\in\mathcal{C}} \rho^{\alpha_\rho} &= \sum_i \sigma_i, \\ \prod_{\rho\in\mathcal{C}} \rho^{\beta_\rho} &= \sum_j \tau_j \end{aligned}$$

with $\sigma_i \in \Sigma(\alpha)$ and $\tau_j \in \Sigma(\beta)$. Then each product $\sigma_i \tau_j$ is in $\Sigma(\alpha+\beta)$ by (3.2.3), and

$$\prod_{\rho \in \mathcal{C}} \rho^{\alpha_\rho + \beta_\rho} = \sum_{i,j} \sigma_i \tau_j$$

shows that $\{\sigma_i \tau_j\}$ exhausts $\Sigma(\alpha+\beta)$. This gives surjectivity. □

Proof of Theorem 3.2.3. Recall from Definition 3.2.1 that for each ρ in $\mathcal{C}$, one defines $A(\rho) \subset A$ to be the $\mathbb{Z}$-span of

$$\Sigma(\rho) := \{\sigma \in \Sigma : \text{ there exists } n \geq 0 \text{ with } (\sigma, \rho^n) \neq 0\}.$$

In other words, $A(\rho) := \bigoplus_{n \geq 0} A_{(n \cdot e_\rho)}$ where e_ρ in $\mathbb{N}^{\mathcal{C}}_{\text{fin}}$ is the standard basis element indexed by ρ. Proposition 3.2.7 then shows that $A(\rho)$ is a Hopf subalgebra of A. Since every α in $\mathbb{N}^{\mathcal{C}}_{\text{fin}}$ can be expressed as the (finite) sum $\sum_\rho \alpha_\rho e_\rho$, and the e_ρ have disjoint support, iterating Proposition 3.2.8 shows that $A = \bigotimes_{\rho \in \mathcal{C}} A(\rho)$. Lastly, $\Sigma(\rho)$ is clearly a PSH-basis for $A(\rho)$, and if σ is any primitive element in $\Sigma(\rho)$ then $(\sigma, \rho^n) \neq 0$ lets one conclude via Lemma 3.2.4 that $\sigma = \rho$ (and $n = 1$). □

3.3. **Λ is the unique indecomposable PSH.** The goal here is to prove the rest of Zelevinsky's structure theory for PSH's. Namely, if A has only one primitive element ρ in its PSH-basis Σ, then A must be isomorphic as a PSH to the ring of symmetric functions Λ, after one rescales the grading of A. Note that every σ in Σ has $\sigma \leq \rho^n$ for some n, and hence has degree divisible by the degree of ρ. Thus one can divide all degrees by that of ρ and assume ρ has degree 1.

The idea is to find within A and Σ a set of elements that play the role of

$$\{h_n = s_{(n)}\}_{n=0,1,2,\ldots}, \qquad \{e_n = s_{(1^n)}\}_{n=0,1,2,\ldots}$$

within $A = \Lambda$ and its PSH-basis of Schur functions $\Sigma = \{s_\lambda\}$. Zelevinsky's argument does this by isolating some properties that turn out to characterize these elements:

(a) $h_0 = e_0 = 1$, and $h_1 = e_1 =: \rho$ has ρ^2 a sum of two elements of Σ, namely

$$\rho^2 = h_2 + e_2.$$

(b) For all $n = 0, 1, 2, \ldots$, there exist unique elements h_n, e_n in $A_n \cap \Sigma$ that satisfy

$$h_2^\perp e_n = 0,$$
$$e_2^\perp h_n = 0$$

with h_2, e_2 being the two elements of Σ introduced in (a).

(c) For $k = 0, 1, 2, \ldots, n$ one has

$$h_k^\perp h_n = h_{n-k} \text{ and } \sigma^\perp h_n = 0 \text{ for } \sigma \in \Sigma \setminus \{h_0, h_1, \ldots, h_n\},$$
$$e_k^\perp e_n = e_{n-k} \text{ and } \sigma^\perp e_n = 0 \text{ for } \sigma \in \Sigma \setminus \{e_0, e_1, \ldots, e_n\}.$$

In particular, $e_k^\perp h_n = 0 = h_k^\perp e_n$ for $k \geq 2$.

(d) Their coproducts are

$$\Delta(h_n) = \sum_{i+j=n} h_i \otimes h_j,$$
$$\Delta(e_n) = \sum_{i+j=n} e_i \otimes e_j.$$

We will prove Zelevinsky's result [227, Theorem 3.1] as a combination of the following two theorems.

Theorem 3.3.1. *Let A be a PSH with PSH-basis Σ containing only one primitive ρ, and assume that the grading has been rescaled so that ρ has degree 1. Then, after renaming $\rho = e_1 = h_1$, one can find unique sequences $\{h_n\}_{n=0,1,2,\ldots}$, $\{e_n\}_{n=0,1,2,\ldots}$ of elements of Σ having properties (a),(b),(c),(d) listed above.*

The second theorem uses the following notion.

Definition 3.3.2. A *PSH-morphism* $A \overset{\varphi}{\to} A'$ between two PSH's A, A' having PSH-bases Σ, Σ' is a graded Hopf algebra morphism for which $\varphi(\mathbb{N}\Sigma) \subset \mathbb{N}\Sigma'$. If $A = A'$ and $\Sigma = \Sigma'$ it will be called a *PSH-endomorphism*. If φ is an isomorphism and restricts to a bijection $\Sigma \to \Sigma'$, it will be called a *PSH-isomorphism*[205]; if it is both a PSH-isomorphism and an endomorphism, it is a *PSH-automorphism*.[206]

Theorem 3.3.3. *The elements $\{h_n\}_{n=0,1,2,\ldots}, \{e_n\}_{n=0,1,2,\ldots}$ in Theorem 3.3.1 also satisfy the following.*

(e) *The elements h_n, e_n in A satisfy the same relation (2.4.4)*

$$\sum_{i+j=n} (-1)^i e_i h_j = \delta_{0,n}$$

as their counterparts in Λ, along with the property that

$$A = \mathbb{Z}[h_1, h_2, \ldots] = \mathbb{Z}[e_1, e_2, \ldots].$$

(f) *There is exactly one nontrivial automorphism $A \overset{\omega}{\to} A$ as a PSH, swapping $h_n \leftrightarrow e_n$.*

(g) *There are exactly two PSH-isomorphisms $A \to \Lambda$:*
- *one sending h_n to the complete homogeneous symmetric functions $h_n(\mathbf{x})$, while sending e_n to the elementary symmetric functions $e_n(\mathbf{x})$,*
- *the second one (obtained by composing the first with ω) sending $h_n \mapsto e_n(\mathbf{x})$ and $e_n \mapsto h_n(\mathbf{x})$.*

Before embarking on the proof, we mention one more bit of convenient terminology: say that an element σ in Σ is a *constituent* of a in $\mathbb{N}\Sigma$ when $\sigma \leq a$, that is, σ appears with nonzero coefficient c_σ in the unique expansion $a = \sum_{\tau\in\Sigma} c_\tau \tau$.

Proof of Theorem 3.3.1. One fact that occurs frequently is this:

$$\text{Every } \sigma \text{ in } \Sigma \cap A_n \text{ is a constituent of } \rho^n. \tag{3.3.1}$$

This follows from Theorem 3.2.3, since ρ is the only primitive element of Σ: one has $A = A(\rho)$ and $\Sigma = \Sigma(\rho)$, so that σ is a constituent of some ρ^m, and homogeneity considerations force $m = n$.

Notice that A is of finite type (due to (3.3.1)). Thus, A^o is a graded Hopf algebra isomorphic to A.

Assertion (a). Note that

$$(\rho^2, \rho^2) = (\rho^\perp(\rho^2), \rho) = (2\rho, \rho) = 2$$

using the fact that $\rho^\perp$ is a derivation since ρ is primitive (Proposition 2.8.2(iv)). On the other hand, expressing $\rho^2 = \sum_{\sigma\in\Sigma} c_\sigma \sigma$ with c_σ in $\mathbb{N}$, one has $(\rho^2, \rho^2) = \sum_\sigma c_\sigma^2$. Hence exactly two of the $c_\sigma = 1$, so ρ^2 has exactly two distinct constituents. Denote them by h_2 and e_2. One concludes that $\Sigma \cap A_2 = \{h_2, e_2\}$ from (3.3.1).

Note also that the same argument shows $\Sigma \cap A_1 = \{\rho\}$, so that $A_1 = \mathbb{Z}\rho$. Since $\rho^\perp h_2$ lies in $A_1 = \mathbb{Z}\rho$ and $(\rho^\perp h_2, \rho) = (h_2, \rho^2) = 1$, we have $\rho^\perp h_2 = \rho$. Similarly $\rho^\perp e_2 = \rho$.

Assertion (b). We will show via induction on n the following three assertions for $n \geq 1$:

(3.3.2)
- There exists an element h_n in $\Sigma \cap A_n$ with $e_2^\perp h_n = 0$.
- This element h_n is unique.
- Furthermore $\rho^\perp h_n = h_{n-1}$.

In the base cases $n = 1, 2$, it is not hard to check that our previously labelled elements, h_1, h_2 (namely $h_1 := \rho$, and h_2 as named in part (a)) really *are* the unique elements satisfying these hypotheses.

[205] This definition is easily seen to be equivalent to saying that a PSH-isomorphism is an invertible PSH-morphism whose inverse is again a PSH-morphism.

[206] The reader should be warned that not every invertible PSH-endomorphism is necessarily a PSH-automorphism. For instance, it is an easy exercise to check that $\Lambda \otimes \Lambda \to \Lambda \otimes \Lambda$, $f \otimes g \mapsto \sum_{(f)} f_1 \otimes f_2 g$ is a well-defined invertible PSH-endomorphism of the PSH $\Lambda \otimes \Lambda$ with PSH-basis $(s_\lambda \otimes s_\mu)_{(\lambda,\mu)\in\mathrm{Par}\times\mathrm{Par}}$, but not a PSH-automorphism.

In the inductive step, it turns out that we will find h_n as a constituent of ρh_{n-1}. Thus we again use the derivation property of $\rho^\perp$ to compute that ρh_{n-1} has exactly two constituents:

$$\begin{aligned}(\rho h_{n-1}, \rho h_{n-1}) &= (\rho^\perp(\rho h_{n-1}), h_{n-1}) \\ &= (h_{n-1} + \rho \cdot \rho^\perp h_{n-1}, h_{n-1}) \\ &= (h_{n-1} + \rho h_{n-2}, h_{n-1}) \\ &= 1 + (h_{n-2}, \rho^\perp h_{n-1}) \\ &= 1 + (h_{n-2}, h_{n-2}) = 1 + 1 = 2\end{aligned}$$

where the inductive hypothesis $\rho^\perp h_{n-1} = h_{n-2}$ was used twice. We next show that exactly one of the two constituents of ρh_{n-1} is annihilated by $e_2^\perp$. Note that since e_2 lies in A_2, and A_1 has $\mathbb{Z}$-basis element ρ, there is a constant c in $\mathbb{Z}$ such that

$$\Delta(e_2) = e_2 \otimes 1 + c\rho \otimes \rho + 1 \otimes e_2. \tag{3.3.3}$$

On the other hand, (a) showed

$$1 = (e_2, \rho^2)_A = (\Delta(e_2), \rho \otimes \rho)_{A \otimes A}$$

so one must have $c = 1$. Therefore by Proposition 2.8.2(iv) again,

$$\begin{array}{rcccccc} e_2^\perp(\rho h_{n-1}) & = & e_2^\perp(\rho) h_{n-1} & + & \rho^\perp(\rho)\rho^\perp(h_{n-1}) & + & \rho e_2^\perp(h_{n-1}) \\ & = & 0 & + & h_{n-2} & + & 0 \\ & = & h_{n-2}, & & & & \end{array} \tag{3.3.4}$$

where the first term vanished due to degree considerations and the last term vanished by the inductive hypothesis. Bearing in mind that ρh_{n-1} lies in $\mathbb{N}\Sigma$, and in a PSH with PSH-basis Σ, any skewing operator $\sigma^\perp$ for σ in Σ will preserve $\mathbb{N}\Sigma$, one concludes from (3.3.4) that

- one of the two distinct constituents of the element ρh_{n-1} must be sent by $e_2^\perp$ to h_{n-2}, and
- the other constituent of ρh_{n-1} must be annihilated by $e_2^\perp$; call this second constituent h_n.

Lastly, to see that this h_n is unique, it suffices to show that any element σ of $\Sigma \cap A_n$ which is killed by $e_2^\perp$ must be a constituent of ρh_{n-1}. This holds for the following reason. We know $\sigma \leq \rho^n$ by (3.3.1), and hence $0 \neq (\rho^n, \sigma) = (\rho^{n-1}, \rho^\perp \sigma)$, implying that $\rho^\perp \sigma \neq 0$. On the other hand, since $0 = \rho^\perp e_2^\perp \sigma = e_2^\perp \rho^\perp \sigma$, one has that $\rho^\perp \sigma$ is annihilated by $e_2^\perp$, and hence $\rho^\perp \sigma$ must be a (positive) multiple of h_{n-1} by part of our inductive hypothesis. Therefore $(\sigma, \rho h_{n-1}) = (\rho^\perp \sigma, h_{n-1})$ is positive, that is, σ is a constituent of ρh_{n-1}.

The preceding argument, applied to $\sigma = h_n$, shows that $\rho^\perp h_n = c h_{n-1}$ for some c in $\{1, 2, \ldots\}$. Since $(\rho^\perp h_n, h_{n-1}) = (h_n, \rho h_{n-1}) = 1$, this c must be 1, so that $\rho^\perp h_n = h_{n-1}$. This completes the induction step in the proof of (3.3.2).

One can then argue, swapping the roles of e_n, h_n in the above argument, the existence and uniqueness of a sequence $\{e_n\}_{n=0}^\infty$ in Σ satisfying the properties analogous to (3.3.2), with $e_0 := 1, e_1 := \rho$.

Assertion (c). Iterating the property from (b) that $\rho^\perp h_n = h_{n-1}$ shows that $(\rho^k)^\perp h_n = h_{n-k}$ for $0 \leq k \leq n$. However one also has an expansion

$$\rho^k = c h_k + \sum_{\substack{\sigma \in \Sigma \cap A_k: \\ \sigma \neq h_k}} c_\sigma \sigma$$

for some integers $c, c_\sigma > 0$, since every σ in $\Sigma \cap A_k$ is a constituent of ρ^k. Hence

$$1 = (h_{n-k}, h_{n-k}) = ((\rho^k)^\perp h_n, (\rho^k)^\perp h_n) \geq c^2 (h_k^\perp h_n, h_k^\perp h_n)$$

using Proposition 2.8.2(ii). Hence if we knew that $h_k^\perp h_n \neq 0$ this would force

$$h_k^\perp h_n = (\rho^k)^\perp h_n = h_{n-k}$$

as well as $\sigma^\perp h_n = 0$ for all $\sigma \notin \{h_0, h_1, \ldots, h_n\}$. But

$$(\rho^{n-k})^\perp h_k^\perp h_n = h_k^\perp (\rho^{n-k})^\perp h_n = h_k^\perp h_k = 1 \neq 0$$

so $h_k^\perp h_n \neq 0$, as desired. The argument for $e_k^\perp e_n = e_{n-k}$ is symmetric.

The last assertion in (c) follows if one checks that $e_n \neq h_n$ for each $n \geq 2$, but this holds since $e_2^\perp(h_n) = 0$ but $e_2^\perp(e_n) = e_{n-2}$.

Assertion (d). Part (c) implies that

$$(\Delta h_n, \sigma \otimes \tau)_{A \otimes A} = (h_n, \sigma\tau)_A = (\sigma^\perp h_n, \tau)_A = 0$$

unless $\sigma = h_k$ for some $k = 0, 1, 2, \ldots, n$ and $\tau = h_{n-k}$. Also one can compute

$$(\Delta h_n, h_k \otimes h_{n-k}) = (h_n, h_k h_{n-k}) = (h_k^\perp h_n, h_{n-k}) \overset{(c)}{=} (h_{n-k}, h_{n-k}) = 1.$$

This is equivalent to the assertion for Δh_n in (d). The argument for Δe_n is symmetric. □

Before proving Theorem 3.3.3, we note some consequences of Theorem 3.3.1. Define for each partition $\lambda = (\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_\ell)$ the following two elements of A:

$$h_\lambda = h_{\lambda_1} h_{\lambda_2} \cdots h_{\lambda_\ell} = h_{\lambda_1} h_{\lambda_2} \cdots ,$$
$$e_\lambda = e_{\lambda_1} e_{\lambda_2} \cdots e_{\lambda_\ell} = e_{\lambda_1} e_{\lambda_2} \cdots .$$

Also, define the *lexicographic order* on Par_n by saying $\lambda <_{\mathrm{lex}} \mu$ if $\lambda \neq \mu$ and the smallest index i for which $\lambda_i \neq \mu_i$ has $\lambda_i < \mu_i$. Recall also that λ^t denotes the *conjugate* or *transpose* partition to λ, obtained by swapping rows and columns in the Ferrers diagram.

The following unitriangularity lemma will play a role in the proof of Theorem 3.3.3(e).

Lemma 3.3.4. *Under the hypotheses of Theorem 3.3.1, for λ, μ in Par_n, one has*

$$e_\mu^\perp h_\lambda = \begin{cases} 1, & \text{if } \mu = \lambda^t; \\ 0, & \text{if } \mu >_{\mathrm{lex}} \lambda^t. \end{cases} \tag{3.3.5}$$

Consequently

$$\det \left[(e_{\mu^t}, h_\lambda) \right]_{\lambda, \mu \in \mathrm{Par}_n} = 1. \tag{3.3.6}$$

Proof. Notice that A is of finite type (as shown in the proof of Theorem 3.3.1). Thus, A^o is a graded Hopf algebra isomorphic to A.

Also, notice that any $m \in \mathbb{N}$ and any $a_1, a_2, \ldots, a_\ell \in A$ satisfy

$$e_m^\perp (a_1 a_2 \cdots a_\ell) = \sum_{i_1 + \cdots + i_\ell = m} e_{i_1}^\perp (a_1) \cdots e_{i_\ell}^\perp (a_\ell) . \tag{3.3.7}$$

Indeed, this follows by induction over ℓ using Proposition 2.8.2(iv) (and the coproduct formula for $\Delta(e_n)$ in Theorem 3.3.1(d)).

In order to prove (3.3.5), induct on the length of μ. If λ has length ℓ, so that $\lambda_1^t = \ell$, then

$$\begin{aligned} e_\mu^\perp h_\lambda &= e_{(\mu_2, \mu_3, \ldots)}^\perp \left(e_{\mu_1}^\perp (h_{\lambda_1} \cdots h_{\lambda_\ell}) \right) \qquad \left(\text{since } e_\mu = e_{\mu_1} e_{(\mu_2, \mu_3, \ldots)} \text{ and thus } e_\mu^\perp = e_{(\mu_2, \mu_3, \ldots)}^\perp \circ e_{\mu_1}^\perp \right) \\ &= e_{(\mu_2, \mu_3, \ldots)}^\perp \sum_{i_1 + \cdots + i_\ell = \mu_1} e_{i_1}^\perp (h_{\lambda_1}) \cdots e_{i_\ell}^\perp (h_{\lambda_\ell}) \qquad (\text{by } (3.3.7)) \\ &= e_{(\mu_2, \mu_3, \ldots)}^\perp \sum_{\substack{i_1 + \cdots + i_\ell = \mu_1; \\ \text{each of } i_1, \ldots, i_\ell \text{ is } \leq 1}} e_{i_1}^\perp (h_{\lambda_1}) \cdots e_{i_\ell}^\perp (h_{\lambda_\ell}) \qquad \left(\text{since } e_k^\perp h_n = 0 \text{ for } k \geq 2 \right) \\ &= \begin{cases} 0, & \text{if } \mu_1 > \ell = \lambda_1^t; \\ e_{(\mu_2, \mu_3, \ldots)}^\perp h_{(\lambda_1 - 1, \ldots, \lambda_\ell - 1)}, & \text{if } \mu_1 = \ell = \lambda_1^t \end{cases} \end{aligned}$$

where the last equality used

$$e_k^\perp (h_n) = \begin{cases} h_{n-1}, & \text{if } k = 1; \\ 0, & \text{if } k \geq 2. \end{cases}$$

Now apply the induction hypothesis, since $(\lambda_1 - 1, \ldots, \lambda_\ell - 1)^t = (\lambda_2^t, \lambda_3^t, \ldots)$.

To prove (3.3.6), note that any λ, μ in Par_n satisfy $(e_{\mu^t}, h_\lambda) = (e_{\mu^t}^\perp (h_\lambda), 1) = e_{\mu^t}^\perp (h_\lambda)$ (since degree considerations enforce $e_{\mu^t}^\perp (h_\lambda) \in A_0 = \mathbf{k} \cdot 1$), and thus

$$(e_{\mu^t}, h_\lambda) = e_{\mu^t}^\perp (h_\lambda) = \begin{cases} 1, & \text{if } \mu^t = \lambda^t; \\ 0, & \text{if } \mu^t >_{\mathrm{lex}} \lambda^t \end{cases}$$

(by (3.3.5)). This means that the matrix $[(e_{\mu^t}, h_\lambda)]_{\lambda,\mu\in\mathrm{Par}_n}$ is unitriangular with respect to some total order on Par_n (namely, the lexicographic order on the conjugate partitions), and hence has determinant 1. □

The following proposition will be the crux of the proof of Theorem 3.3.3(f) and (g), and turns out to be closely related to Kerov's *asymptotic theory of characters of the symmetric groups* [108].

Proposition 3.3.5. *Given a PSH A with PSH-basis Σ containing only one primitive ρ, the two maps $A \to \mathbb{Z}$ defined on $A = \bigoplus_{n\geq 0} A_n$ via*

$$\delta_h = \bigoplus_n h_n^\perp,$$
$$\delta_e = \bigoplus_n e_n^\perp$$

are characterized as the only two $\mathbb{Z}$-linear maps $A \overset{\delta}{\to} \mathbb{Z}$ with the three properties of being

- ***positive**: $\delta(\mathbb{N}\Sigma) \subset \mathbb{N}$,*
- ***multiplicative**: $\delta(a_1a_2) = \delta(a_1)\delta(a_2)$ for all $a_1, a_2 \in A$, and*
- ***normalized**: $\delta(\rho) = 1$.*

Proof. Notice that A is of finite type (as shown in the proof of Theorem 3.3.1). Thus, A^o is a graded Hopf algebra isomorphic to A.

It should be clear from their definitions that δ_h, δ_e are $\mathbb{Z}$-linear, positive and normalized. To see that δ_h is multiplicative, by $\mathbb{Z}$-linearity, it suffices to check that for a_1, a_2 in A_{n_1}, A_{n_2} with $n_1 + n_2 = n$, one has

$$\delta_h(a_1a_2) = h_n^\perp(a_1a_2) = \sum_{i_1+i_2=n} h_{i_1}^\perp(a_1)h_{i_2}^\perp(a_2) = h_{n_1}^\perp(a_1)h_{n_2}^\perp(a_2) = \delta_h(a_1)\delta_h(a_2)$$

in which the second equality used Proposition 2.8.2(iv) and Theorem 3.3.1(d). The argument for δ_e is symmetric.

Conversely, given $A \overset{\delta}{\to} \mathbb{Z}$ which is $\mathbb{Z}$-linear, positive, multiplicative, and normalized, note that

$$\delta(h_2) + \delta(e_2) = \delta(h_2 + e_2) = \delta(\rho^2) = \delta(\rho)^2 = 1^2 = 1$$

and hence positivity implies that either $\delta(h_2) = 0$ or $\delta(e_2) = 0$. Assume the latter holds, and we will show that $\delta = \delta_h$.

Given any σ in $\Sigma \cap A_n \setminus \{h_n\}$, note that $e_2^\perp\sigma \neq 0$ by Theorem 3.3.1(b), and hence $0 \neq (e_2^\perp\sigma, \rho^{n-2}) = (\sigma, e_2\rho^{n-2})$. Thus σ is a constituent of $e_2\rho^{n-2}$, so positivity implies

$$0 \leq \delta(\sigma) \leq \delta(e_2\rho^{n-2}) = \delta(e_2)\delta(\rho^{n-2}) = 0.$$

Thus $\delta(\sigma) = 0$ for σ in $\Sigma \cap A_n \setminus \{h_n\}$. Since $\delta(\rho^n) = \delta(\rho)^n = 1^n = 1$, this forces $\delta(h_n) = 1$, for each $n \geq 0$ (including $n = 0$, as $1 = \delta(\rho) = \delta(\rho \cdot 1) = \delta(\rho)\delta(1) = 1 \cdot \delta(1) = \delta(1)$). Thus $\delta = \delta_h$. The argument when $\delta(h_2) = 0$ showing $\delta = \delta_e$ is symmetric. □

Proof of Theorem 3.3.3. Many of the assertions of parts (e) and (f) will come from constructing the unique nontrivial PSH-automorphism ω of A from the antipode S: for homogeneous a in A_n, define $\omega(a) := (-1)^nS(a)$. We now study some of the properties of S and ω.

Notice that A is of finite type (as shown in the proof of Theorem 3.3.1). Thus, A^o is a graded Hopf algebra isomorphic to A.

Since A is a PSH, it is commutative by Theorem 3.1.7 (applied to $A \otimes_\mathbb{Z} \mathbb{Q}$). This implies both that S is an algebra endomorphism by Proposition 1.4.10 (since Exercise 1.5.8(a) shows that the algebra anti-endomorphisms of a commutative algebra are the same as its algebra endomorphisms), and that $S^2 = \mathrm{id}_A$ by Corollary 1.4.12. Thus, ω is an algebra endomorphism and satisfies $\omega^2 = \mathrm{id}_A$.

Since A is self-dual and the defining diagram (1.4.3) satisfied by the antipode S is sent to itself when one replaces A by A^o and all maps by their adjoints, one concludes that $S = S^*$ (where S^* means the restricted adjoint $S^* : A^o \to A^o$), i.e., S is self-adjoint. Since S is an algebra endomorphism, and $S = S^*$, in fact S is also a coalgebra endomorphism, a bialgebra endomorphism, and a Hopf endomorphism (by Corollary 1.4.27). The same properties are shared by ω.

Since $\mathrm{id}_A = S^2 = SS^*$, one concludes that S is an isometry, and hence so is ω.

Since ρ is primitive, one has $S(\rho) = -\rho$ and $\omega(\rho) = \rho$. Therefore $\omega(\rho^n) = \rho^n$ for $n = 1, 2, \ldots$. Use this as follows to check that ω is a PSH-automorphism, which amounts to checking that every σ in Σ has $\omega(\sigma)$ in Σ:

$$(\omega(\sigma), \omega(\sigma)) = (\sigma, \sigma) = 1$$

so that $\pm\omega(\sigma)$ lies in Σ, but also if σ lies in A_n, then

$$(\omega(\sigma), \rho^n) = (\sigma, \omega(\rho^n)) = (\sigma, \rho^n) > 0.$$

In summary, ω is a PSH-automorphism of A, an isometry, and an involution.

Let us try to determine the action of ω on the $\{h_n\}$. By similar reasoning as in (3.3.3), one has

$$\Delta(h_2) = h_2 \otimes 1 + \rho \otimes \rho + 1 \otimes h_2.$$

Thus $0 = S(h_2) + S(\rho)\rho + h_2$, and combining this with $S(\rho) = -\rho$, one has $S(h_2) = e_2$. Thus also $\omega(h_2) = (-1)^2 S(h_2) = e_2$.

We claim that this forces $\omega(h_n) = e_n$, because $h_2^\perp \omega(h_n) = 0$ via the following calculation: for any a in A one has

$$\begin{aligned}
(h_2^\perp \omega(h_n), a) &= (\omega(h_n), h_2 a) \\
&= (h_n, \omega(h_2 a)) \\
&= (h_n, e_2 \omega(a)) \\
&= (e_2^\perp h_n, \omega(a)) = (0, \omega(a)) = 0.
\end{aligned}$$

Consequently the involution ω swaps h_n and e_n, while the antipode S has $S(h_n) = (-1)^n e_n$ and $S(e_n) = (-1)^n h_n$. Thus the coproduct formulas in (d) and definition of the antipode S imply the relation (2.4.4) between $\{h_n\}$ and $\{e_n\}$.

This relation (2.4.4) also lets one recursively express the h_n as polynomials with integer coefficients in the $\{e_n\}$, and vice-versa, so that $\{h_n\}$ and $\{e_n\}$ each generate the same $\mathbb{Z}$-subalgebra A' of A. We wish to show that A' exhausts A.

We argue that Lemma 3.3.4 implies that the *Gram matrix* $[(h_\mu, h_\lambda)]_{\mu,\lambda \in \mathrm{Par}_n}$ has determinant ± 1 as follows. Since $\{h_n\}$ and $\{e_n\}$ both generate A', there exists a $\mathbb{Z}$-matrix $(a_{\mu,\lambda})$ expressing $e_{\mu^t} = \sum_\lambda a_{\mu,\lambda} h_\lambda$, and one has

$$[(e_{\mu^t}, h_\lambda)] = [a_{\mu,\lambda}] \cdot [(h_\mu, h_\lambda)] .$$

Taking determinants of these three $\mathbb{Z}$-matrices, and using the fact that the determinant on the left is 1 (by (3.3.6)), both determinants on the right must also be ± 1.

Now we will show that every $\sigma \in \Sigma \cap A_n$ lies in A'_n. Uniquely express $\sigma = \sigma' + \sigma''$ in which σ' lies in the $\mathbb{R}$-span $\mathbb{R}A'_n$ and σ'' lies in the real perpendicular space $(\mathbb{R}A'_n)^\perp$ inside $\mathbb{R} \otimes_{\mathbb{Z}} A_n$. One can compute $\mathbb{R}$-coefficients $(c_\mu)_{\mu \in \mathrm{Par}_n}$ that express $\sigma' = \sum_\mu c_\mu h_\mu$ by solving the system

$$\left(\sum_\mu c_\mu h_\mu, h_\lambda \right) = (\sigma, h_\lambda) \text{ for } \lambda \in \mathrm{Par}_n .$$

This linear system is governed by the Gram matrix $[(h_\mu, h_\lambda)]_{\mu,\lambda \in \mathrm{Par}_n}$ with determinant ± 1, and its right side has $\mathbb{Z}$-entries since σ, h_λ lie in A. Hence the solution $(c_\mu)_{\mu \in \mathrm{Par}_n}$ will have $\mathbb{Z}$-entries, so σ' lies in A'. Furthermore, $\sigma'' = \sigma - \sigma'$ will lie in A, and hence by the orthogonality of σ', σ'',

$$1 = (\sigma, \sigma) = (\sigma', \sigma') + (\sigma'', \sigma'').$$

One concludes that either $\sigma'' = 0$, or $\sigma' = 0$. The latter cannot occur since it would mean that $\sigma = \sigma''$ is perpendicular to all of A'. But $\rho^n = h_1^n$ lies in A', and $(\sigma, \rho^n) \neq 0$. Thus $\sigma'' = 0$, meaning $\sigma = \sigma'$ lies in A'. This completes the proof of assertion (e). Note that in the process, having shown $\det(h_\mu, h_\lambda)_{\lambda,\mu \in \mathrm{Par}_n} = \pm 1$, one also knows that $\{h_\lambda\}_{\lambda \in \mathrm{Par}_n}$ are $\mathbb{Z}$-linearly independent, so that $\{h_1, h_2, \ldots\}$ are algebraically independent[207], and $A = \mathbb{Z}[h_1, h_2, \ldots]$ is the polynomial algebra generated by $\{h_1, h_2, \ldots\}$.

For assertion (f), we have seen that ω gives such a PSH-automorphism $A \to A$, swapping $h_n \leftrightarrow e_n$. Conversely, given a PSH-automorphism $A \overset{\varphi}{\to} A$, consider the positive, multiplicative, normalized $\mathbb{Z}$-linear map $\delta := \delta_h \circ \varphi : A \to \mathbb{Z}$. Proposition 3.3.5 shows that either

[207] by Exercise 2.2.14(c)

- $\delta = \delta_h$, which then forces $\varphi(h_n) = h_n$ for all n, so $\varphi = \mathrm{id}_A$, or
- $\delta = \delta_e$, which then forces $\varphi(e_n) = h_n$ for all n, so $\varphi = \omega$.

For assertion (g), given a PSH A with PSH-basis Σ having exactly one primitive ρ, since we have seen $A = \mathbb{Z}[h_1, h_2, \ldots]$, where h_n in A is as defined in Theorem 3.3.1, one can uniquely define an algebra morphism $A \overset{\varphi}{\to} \Lambda$ that sends the element h_n to the complete homogeneous symmetric function $h_n(\mathbf{x})$. Assertions (d) and (e) show that φ is a bialgebra isomorphism, and hence it is a Hopf isomorphism. To show that it is a PSH-isomorphism, we first note that it is an isometry because one can iterate Proposition 2.8.2(iv) together with assertions (c) and (d) to compute all inner products

$$(h_\mu, h_\lambda)_A = (1, h_\mu^\perp h_\lambda)_A = (1, h_{\mu_1}^\perp h_{\mu_2}^\perp \cdots (h_{\lambda_1} h_{\lambda_2} \cdots))_A$$

for μ, λ in Par_n. Hence

$$(h_\mu, h_\lambda)_A = (h_\mu(\mathbf{x}), h_\lambda(\mathbf{x}))_\Lambda = (\varphi(h_\mu), \varphi(h_\lambda))_\Lambda.$$

Once one knows φ is an isometry, then elements ω in $\Sigma \cap A_n$ are characterized in terms of the form $(\cdot, \cdot)$ by $(\omega, \omega) = 1$ and $(\omega, \rho^n) > 0$. Hence φ sends each σ in Σ to a Schur function s_λ, and is a PSH-isomorphism. □

4. Complex representations for $\mathfrak{S}_n$, wreath products, $GL_n(\mathbb{F}_q)$

After reviewing the basics that we will need from representation and character theory of finite groups, we give Zelevinsky's three main examples of PSH's arising as spaces of virtual characters for three towers of finite groups:

- *symmetric* groups,
- their *wreath products* with any finite group, and
- the finite *general linear* groups.

Much in this chapter traces its roots to Zelevinsky's book [227]. The results concerning the symmetric groups, however, are significantly older and spread across the literature: see, e.g., [206, §7.18], [73, §7.3], [142, §I.7], [186, §4.7], [113], for proofs using different tools.

4.1. **Review of complex character theory.** We shall now briefly discuss some basics of representation (and character) theory that will be used below. A good source for this material, including the crucial Mackey formula, is Serre [197, Chaps. 1-7].[208]

4.1.1. *Basic definitions, Maschke, Schur.* For a group G, a *representation of* G is a homomorphism $G \xrightarrow{\varphi} GL(V)$ for some vector space V over a field. We will take the field to be $\mathbb{C}$ from now on, and we will also assume that V is finite-dimensional over $\mathbb{C}$. Thus a representation of G is the same as a finite-dimensional (left) $\mathbb{C}G$-module V. (We use the notations $\mathbb{C}G$ and $\mathbb{C}[G]$ synonymously for the group algebra of G over $\mathbb{C}$. More generally, if S is a set, then $\mathbb{C}S = \mathbb{C}[S]$ denotes the free $\mathbb{C}$-module with basis S.)

We also assume that G is finite, so that Maschke's Theorem[209] says that $\mathbb{C}G$ is semisimple, meaning that every $\mathbb{C}G$-module $U \subset V$ has a $\mathbb{C}G$-module complement U' with $V = U \oplus U'$. Equivalently, *indecomposable* $\mathbb{C}G$-modules are the same thing as *simple* (=*irreducible*) $\mathbb{C}G$-modules.

Schur's Lemma implies that for two simple $\mathbb{C}G$-modules V_1, V_2, one has

$$\operatorname{Hom}_{\mathbb{C}G}(V_1, V_2) \cong \begin{cases} \mathbb{C}, & \text{if } V_1 \cong V_2; \\ 0, & \text{if } V_1 \not\cong V_2. \end{cases}$$

4.1.2. *Characters and Hom spaces.* A $\mathbb{C}G$-module V is completely determined up to isomorphism by its *character*

$$\begin{array}{ccc} G & \xrightarrow{\chi_V} & \mathbb{C}, \\ g & \longmapsto & \chi_V(g) := \operatorname{trace}(g : V \to V). \end{array}$$

This character χ_V is a *class function*, meaning it is constant on G-conjugacy classes. The space $R_{\mathbb{C}}(G)$ of class functions $G \to \mathbb{C}$ has a Hermitian, positive definite form

$$(f_1, f_2)_G := \frac{1}{|G|} \sum_{g \in G} f_1(g) \overline{f_2(g)}.$$

For any two $\mathbb{C}G$-modules V_1, V_2,

$$(\chi_{V_1}, \chi_{V_2})_G = \dim_{\mathbb{C}} \operatorname{Hom}_{\mathbb{C}G}(V_1, V_2). \tag{4.1.1}$$

The set of all *irreducible characters*

$$\operatorname{Irr}(G) = \{\chi_V : V \text{ is a simple } \mathbb{C}G\text{-module}\}$$

forms an orthonormal basis of $R_{\mathbb{C}}(G)$ with respect to this form, and spans a $\mathbb{Z}$-sublattice

$$R(G) := \mathbb{Z} \operatorname{Irr}(G) \subset R_{\mathbb{C}}(G)$$

sometimes called the *virtual characters* of G. For every $\mathbb{C}G$-module V, the character χ_V belongs to $R(G)$.

Instead of working with the Hermitian form $(\cdot, \cdot)_G$ on G, we could also (and some authors do) define a $\mathbb{C}$-bilinear form $\langle \cdot, \cdot \rangle_G$ on $R_{\mathbb{C}}(G)$ by

$$\langle f_1, f_2 \rangle_G := \frac{1}{|G|} \sum_{g \in G} f_1(g) f_2(g^{-1}).$$

[208] More advanced treatments of representation theory can be found in [222] and [69].

[209] ... which has a beautiful generalization to finite-dimensional Hopf algebras due to Larson and Sweedler; see Montgomery [157, §2.2].

This form is not identical with $(\cdot,\cdot)_G$ (indeed, $\langle\cdot,\cdot\rangle_G$ is bilinear while $(\cdot,\cdot)_G$ is Hermitian), but it still satisfies (4.1.1), and thus is identical with $(\cdot,\cdot)_G$ on $R(G)\times R(G)$. Hence, for all we are going to do until Section 4.9, we could just as well use the form $\langle\cdot,\cdot\rangle_G$ instead of $(\cdot,\cdot)_G$.

4.1.3. *Tensor products.* Given two groups G_1, G_2 and $\mathbb{C}G_i$-modules V_i for $i = 1, 2$, their tensor product $V_1 \otimes_{\mathbb{C}} V_2$ becomes a $\mathbb{C}[G_1 \times G_2]$-module via $(g_1, g_2)(v_1 \otimes v_2) = g_1(v_1) \otimes g_2(v_2)$. This module is called the *(outer) tensor product* of V_1 and V_2. When V_1, V_2 are both simple, then so is $V_1 \otimes V_2$, and every simple $\mathbb{C}[G_1 \times G_2]$-module arises this way (with V_1 and V_2 determined uniquely up to isomorphism).[210] Thus one has identifications and isomorphisms

$$\mathrm{Irr}(G_1 \times G_2) = \mathrm{Irr}(G_1) \times \mathrm{Irr}(G_2),$$
$$R(G_1 \times G_2) \cong R(G_1) \otimes_{\mathbb{Z}} R(G_2);$$

here, $\chi_{V_1} \otimes \chi_{V_2} \in R(G_1) \otimes_{\mathbb{Z}} R(G_2)$ is being identified with $\chi_{V_1 \otimes V_2} \in R(G_1 \times G_2)$ for all $\mathbb{C}G_1$-modules V_1 and all $\mathbb{C}G_2$-modules V_2. The latter isomorphism is actually a restriction of the isomorphism $R_{\mathbb{C}}(G_1 \times G_2) \cong R_{\mathbb{C}}(G_1) \otimes_{\mathbb{C}} R_{\mathbb{C}}(G_2)$ under which every pure tensor $\phi_1 \otimes \phi_2 \in R_{\mathbb{C}}(G_1) \otimes_{\mathbb{C}} R_{\mathbb{C}}(G_2)$ corresponds to the class function $G_1 \times G_2 \to \mathbb{C}$, $(g_1, g_2) \mapsto \phi_1(g_1) \otimes \phi_2(g_2)$.

Given two $\mathbb{C}G_1$-modules V_1 and W_1 and two $\mathbb{C}G_2$-modules V_2 and W_2, we have

$$(\chi_{V_1 \otimes V_2}, \chi_{W_1 \otimes W_2})_{G_1 \times G_2} = (\chi_{V_1}, \chi_{W_1})_{G_1} (\chi_{V_2}, \chi_{W_2})_{G_2}. \tag{4.1.2}$$

4.1.4. *Induction and restriction.* Given a subgroup $H < G$ and $\mathbb{C}H$-module U, one can use the fact that $\mathbb{C}G$ is a $(\mathbb{C}G, \mathbb{C}H)$-bimodule to form the *induced* $\mathbb{C}G$*-module*

$$\mathrm{Ind}_H^G U := \mathbb{C}G \otimes_{\mathbb{C}H} U.$$

The fact that $\mathbb{C}G$ is free as a (right-)$\mathbb{C}H$-module[211] on basis elements $\{t_g\}_{gH \in G/H}$ makes this tensor product easy to analyze. For example one can compute its character

$$\chi_{\mathrm{Ind}_H^G U}(g) = \frac{1}{|H|} \sum_{\substack{k \in G:\\ kgk^{-1} \in H}} \chi_U(kgk^{-1}). \tag{4.1.3}$$

[212] One can also recognize when a $\mathbb{C}G$-module V is isomorphic to $\mathrm{Ind}_H^G U$ for some $\mathbb{C}H$-module U: this happens if and only if there is an H-stable subspace $U \subset V$ having the property that $V = \bigoplus_{gH \in G/H} gU$.

The above construction of a $\mathbb{C}G$-module $\mathrm{Ind}_H^G U$ corresponding to any $\mathbb{C}H$-module U is part of a functor Ind_H^G from the category of $\mathbb{C}H$-modules to the category of $\mathbb{C}G$-modules[213]; this functor is called *induction*.

Besides induction on $\mathbb{C}H$-modules, one can define induction on class functions of H:

Exercise 4.1.1. Let G be a finite group, and H a subgroup of G. Let $f \in R_{\mathbb{C}}(H)$ be a class function. We define the *induction* $\mathrm{Ind}_H^G f$ of f to be the function $G \to \mathbb{C}$ given by

$$\left(\mathrm{Ind}_H^G f\right)(g) = \frac{1}{|H|} \sum_{\substack{k \in G:\\ kgk^{-1} \in H}} f\left(kgk^{-1}\right) \qquad \text{for all } g \in G. \tag{4.1.4}$$

(a) Prove that this induction $\mathrm{Ind}_H^G f$ is a class function on G, hence belongs to $R_{\mathbb{C}}(G)$.

(b) Let J be a system of right coset[214] representatives for $H \backslash G$, so that $G = \bigsqcup_{j \in J} Hj$. Prove that

$$\left(\mathrm{Ind}_H^G f\right)(g) = \sum_{\substack{j \in J:\\ jgj^{-1} \in H}} f\left(jgj^{-1}\right) \qquad \text{for all } g \in G.$$

[210] This is proven in [197, §3.2, Thm. 10]. The fact that $\mathbb{C}$ is algebraically closed is essential for this!

[211] ... which also has a beautiful generalization to finite-dimensional Hopf algebras due to Nichols and Zoeller; see [157, §3.1].

[212] See [197, §7.2, Prop. 20(ii)] for the proof of this equality. (Another proof is given in [69, Remark 5.9.2 (the Remark after Theorem 4.32 in the arXiv version)], but [69] uses a different definition of $\mathrm{Ind}_H^G U$; see Remark 4.1.5 for why it is equivalent to ours. Yet another proof of (4.1.3) is given in Exercise 4.1.14(k).)

[213] On morphisms, it sends any $f : U \to U'$ to $\mathrm{id}_{\mathbb{C}G} \otimes_{\mathbb{C}H} f : \mathbb{C}G \otimes_{\mathbb{C}H} U \to \mathbb{C}G \otimes_{\mathbb{C}H} U'$.

[214] A *right coset* of a subgroup H in a group G is defined to be a subset of G having the form Hj for some $j \in G$. Similarly, a *left coset* has the form jH for some $j \in G$.

The induction Ind_H^G defined in Exercise 4.1.1 is a $\mathbb{C}$-linear map $R_{\mathbb{C}}(H) \to R_{\mathbb{C}}(G)$. Since every $\mathbb{C}H$-module U satisfies

$$\chi_{\operatorname{Ind}_H^G U} = \operatorname{Ind}_H^G(\chi_U) \tag{4.1.5}$$

[215], this $\mathbb{C}$-linear map Ind_H^G restricts to a $\mathbb{Z}$-linear map $R(H) \to R(G)$ (also denoted Ind_H^G) which sends the character χ_U of any $\mathbb{C}H$-module U to the character $\chi_{\operatorname{Ind}_H^G U}$ of the induced $\mathbb{C}G$-module $\operatorname{Ind}_H^G U$.

Exercise 4.1.2. Let G, H and I be three finite groups such that $I < H < G$. Let U be a $\mathbb{C}I$-module. Prove that $\operatorname{Ind}_H^G \operatorname{Ind}_I^H U \cong \operatorname{Ind}_I^G U$. (This fact is often referred to as the *transitivity of induction*.)

Exercise 4.1.3. Let G_1 and G_2 be two groups. Let $H_1 < G_1$ and $H_2 < G_2$ be two subgroups. Let U_1 be a $\mathbb{C}H_1$-module, and U_2 be a $\mathbb{C}H_2$-module. Show that

$$\operatorname{Ind}_{H_1 \times H_2}^{G_1 \times G_2} (U_1 \otimes U_2) \cong \left(\operatorname{Ind}_{H_1}^{G_1} U_1\right) \otimes \left(\operatorname{Ind}_{H_2}^{G_2} U_2\right) \tag{4.1.6}$$

as $\mathbb{C}[G_1 \times G_2]$-modules.

The *restriction* operation $V \mapsto \operatorname{Res}_H^G V$ restricts a $\mathbb{C}G$-module V to a $\mathbb{C}H$-module. *Frobenius reciprocity* asserts the adjointness between Ind_H^G and Res_H^G

$$\operatorname{Hom}_{\mathbb{C}G}(\operatorname{Ind}_H^G U, V) \cong \operatorname{Hom}_{\mathbb{C}H}(U, \operatorname{Res}_H^G V), \tag{4.1.7}$$

as a special case ($S = A = \mathbb{C}G, R = \mathbb{C}H, B = U, C = V$) of the general *adjoint associativity*

$$\operatorname{Hom}_S(A \otimes_R B, C) \cong \operatorname{Hom}_R(B, \operatorname{Hom}_S(A, C)) \tag{4.1.8}$$

for S, R two rings, A an (S, R)-bimodule, B a left R-module, C a left S-module.

We can define not just the restriction of a $\mathbb{C}G$-module, but also the *restriction of a class function* $f \in R_{\mathbb{C}}(G)$. When H is a subgroup of G, the restriction $\operatorname{Res}_H^G f$ of an $f \in R_{\mathbb{C}}(G)$ is defined as the result of restricting the map $f : G \to \mathbb{C}$ to H. This $\operatorname{Res}_H^G f$ is easily seen to belong to $R_{\mathbb{C}}(H)$, and so Res_H^G is a $\mathbb{C}$-linear map $R_{\mathbb{C}}(G) \to R_{\mathbb{C}}(H)$. This map restricts to a $\mathbb{Z}$-linear map $R(G) \to R(H)$, since we have $\operatorname{Res}_H^G \chi_V = \chi_{\operatorname{Res}_H^G V}$ for any $\mathbb{C}G$-module V. Taking characters in (4.1.7) (and recalling $\operatorname{Res}_H^G \chi_V = \chi_{\operatorname{Res}_H^G V}$ and (4.1.5)), we obtain

$$(\operatorname{Ind}_H^G \chi_U, \chi_V)_G = (\chi_U, \operatorname{Res}_H^G \chi_V)_H. \tag{4.1.9}$$

By bilinearity, this yields the equality

$$\left(\operatorname{Ind}_H^G \alpha, \beta\right)_G = \left(\alpha, \operatorname{Res}_H^G \beta\right)_H$$

for any class functions $\alpha \in R_{\mathbb{C}}(H)$ and $\beta \in R_{\mathbb{C}}(G)$ (since $R(G)$ spans $R_{\mathbb{C}}(G)$ as a $\mathbb{C}$-vector space).

Exercise 4.1.4. Let G be a finite group, and let $H < G$. Let U be a $\mathbb{C}H$-module. If A and B are two algebras, P is a (B, A)-bimodule and Q is a left B-module, then $\operatorname{Hom}_B(P, Q)$ is a left A-module (since $\mathbb{C}G$ is a $(\mathbb{C}H, \mathbb{C}G)$-bimodule). As a consequence, $\operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, U)$ is a $\mathbb{C}G$-module. Prove that this $\mathbb{C}G$-module is isomorphic to $\operatorname{Ind}_H^G U$.

Remark 4.1.5. Some texts *define* the induction $\operatorname{Ind}_H^G U$ of a $\mathbb{C}H$-module U to be $\operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, U)$ (rather than to be $\mathbb{C}G \otimes_{\mathbb{C}H} U$, as we did).[216] As Exercise 4.1.4 shows, this definition is equivalent to ours as long as G is finite (but not otherwise).

Exercise 4.1.4 yields the following "wrong-way" version of Frobenius reciprocity:

Exercise 4.1.6. Let G be a finite group; let $H < G$. Let U be a $\mathbb{C}G$-module, and let V be a $\mathbb{C}H$-module. Prove that $\operatorname{Hom}_{\mathbb{C}G}\left(U, \operatorname{Ind}_H^G V\right) \cong \operatorname{Hom}_{\mathbb{C}H}\left(\operatorname{Res}_H^G U, V\right)$.

[215] This follows by comparing the value of $\chi_{\operatorname{Ind}_H^G U}(g)$ obtained from (4.1.3) with the value of $\left(\operatorname{Ind}_H^G(\chi_U)\right)(g)$ found using (4.1.4).

[216] Or they define it as a set of morphisms of H-sets from G to U (this is how [69, Def. 5.8.1 (Def. 4.28 in the arXiv version)] defines it); this is easily seen to be equivalent to $\operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, U)$.

4.1.5. *Mackey's formula.* Mackey gave an alternate description of a module which has been induced and then restricted. To state it, for a subgroup $H < G$ and g in G, let $H^g := g^{-1}Hg$ and ${}^gH := gHg^{-1}$. Given a $\mathbb{C}H$-module U, say defined by a homomorphism $H \overset{\varphi}{\to} GL(U)$, let U^g denote the $\mathbb{C}[gHg^{-1}]$-module on the same $\mathbb{C}$-vector space U defined by the composite homomorphism

$$\begin{array}{ccccc} {}^gH & \longrightarrow & H & \overset{\varphi}{\longrightarrow} & GL(U), \\ h & \longmapsto & g^{-1}hg. & & \end{array}$$

Theorem 4.1.7. *(Mackey's formula) Consider subgroups $H, K < G$, and any $\mathbb{C}H$-module U. If $\{g_1, \ldots, g_t\}$ are double coset representatives for $K\backslash G/H$, then*

$$\operatorname{Res}^G_K \operatorname{Ind}^G_H U \cong \bigoplus_{i=1}^t \operatorname{Ind}^K_{{}^{g_i}H\cap K} \left(\left(\operatorname{Res}^H_{H\cap K^{g_i}} U \right)^{g_i} \right).$$

Proof. In this proof, all tensor product symbols $\otimes$ should be interpreted as $\otimes_{\mathbb{C}H}$. Recall $\mathbb{C}G$ has $\mathbb{C}$-basis $\{t_g\}_{g\in G}$. For subsets $S \subset G$, let $\mathbb{C}[S]$ denote the $\mathbb{C}$-span of $\{t_g\}_{g\in S}$ in $\mathbb{C}G$.

Note that each double coset KgH gives rise to a sub-(K, H)-bimodule $\mathbb{C}[KgH]$ within $\mathbb{C}G$, and one has a $\mathbb{C}K$-module direct sum decomposition

$$\operatorname{Ind}^G_H U = \mathbb{C}G \otimes U = \bigoplus_{i=1}^t \mathbb{C}[Kg_iH] \otimes U.$$

Hence it suffices to check for any element g in G that

$$\mathbb{C}[KgH] \otimes U \cong \operatorname{Ind}^K_{{}^gH\cap K} \left(\left(\operatorname{Res}^H_{H\cap K^g} U \right)^g \right).$$

Note that ${}^gH \cap K$ is the subgroup of K consisting of the elements k in K for which $kgH = gH$. Hence by picking $\{k_1, \ldots, k_s\}$ to be coset representatives for $K/({}^gH \cap K)$, one disjointly decomposes the double coset

$$KgH = \bigsqcup_{j=1}^s k_j({}^gH \cap K)gH,$$

giving a $\mathbb{C}$-vector space direct sum decomposition

$$\begin{aligned} \mathbb{C}[KgH] \otimes U &= \bigoplus_{j=1}^s \mathbb{C}\left[k_j\left({}^gH \cap K\right) gH\right] \otimes U \\ &\cong \operatorname{Ind}^K_{{}^gH\cap K} \left(\mathbb{C}[\left({}^gH \cap K\right) gH] \otimes U \right). \end{aligned}$$

So it remains to check that one has a $\mathbb{C}[{}^gH \cap K]$-module isomorphism

$$\mathbb{C}[\left({}^gH \cap K\right) gH] \otimes U \cong \left(\operatorname{Res}^H_{H\cap K^g} U \right)^g.$$

Bearing in mind that, for each k in ${}^gH \cap K$ and h in H, one has $g^{-1}kg$ in H and hence

$$t_{kgh} \otimes u = t_g \cdot t_{g^{-1}kg\cdot h} \otimes u = t_g \otimes g^{-1}kgh \cdot u,$$

one sees that this isomorphism can be defined by mapping

$$t_{kgh} \otimes u \longmapsto g^{-1}kgh \cdot u.$$

□

4.1.6. *Inflation and fixed points.* There are two (adjoint) constructions on representations that apply when one has a normal subgroup $K \triangleleft G$. Given a $\mathbb{C}[G/K]$-module U, say defined by the homomorphism $G/K \overset{\varphi}{\to} GL(U)$, the *inflation* of U to a $\mathbb{C}G$-module $\operatorname{Infl}^G_{G/K} U$ has the same underlying space U, and is defined by the composite homomorphism $G \to G/K \overset{\varphi}{\to} GL(U)$. We will later use the easily-checked fact that when $H < G$ is any other subgroup, one has

$$\operatorname{Res}^G_H \operatorname{Infl}^G_{G/K} U = \operatorname{Infl}^H_{H/H\cap K} \operatorname{Res}^{G/K}_{H/H\cap K} U. \tag{4.1.10}$$

(We regard $H/H \cap K$ as a subgroup of G/K, since the canonical homomorphism $H/H \cap K \to G/K$ is injective.)

Inflation turns out to be adjoint to the *K-fixed space construction* sending a $\mathbb{C}G$-module V to the $\mathbb{C}[G/K]$-module

$$V^K := \{v \in V : k(v) = v \text{ for } k \in K\}.$$

Note that V^K is indeed a G-stable subspace: for any v in V^K and g in G, one has that $g(v)$ lies in V^K since an element k in K satisfies $kg(v) = g \cdot g^{-1}kg(v) = g(v)$ as $g^{-1}kg$ lies in K. One has this adjointness

$$\operatorname{Hom}_{\mathbb{C}G}(\operatorname{Infl}_{G/K}^G U, V) = \operatorname{Hom}_{\mathbb{C}[G/K]}(U, V^K), \tag{4.1.11}$$

because any $\mathbb{C}G$-module homomorphism φ on the left must have the property that $k\varphi(u) = \varphi(k(u)) = \varphi(u)$ for all k in K, so that φ actually lies on the right.

We will also need the following formula for the character χ_{V^K} in terms of the character χ_V:

$$\chi_{V^K}(gK) = \frac{1}{|K|} \sum_{k \in K} \chi_V(gk). \tag{4.1.12}$$

To see this, note that when one has a $\mathbb{C}$-linear endomorphism φ on a space V that preserves some $\mathbb{C}$-subspace $W \subset V$, if $V \overset{\pi}{\to} W$ is any idempotent projection onto W, then the trace of the restriction $\varphi|_W$ equals the trace of $\varphi \circ \pi$ on V. Applying this to $W = V^K$ and $\varphi = g$, with $\pi = \frac{1}{|K|}\sum_{k\in K} k$, gives (4.1.12).[217]

Another way to restate (4.1.12) is:

$$\chi_{V^K}(gK) = \frac{1}{|K|} \sum_{h \in gK} \chi_V(h). \tag{4.1.13}$$

Inflation and K-fixed space construction can also be defined on class functions. For inflation, this is particularly easy: Inflation $\operatorname{Infl}_{G/K}^G f$ of an $f \in R_{\mathbb{C}}(G/K)$ is defined as the composition $G \twoheadrightarrow G/K \xrightarrow{f} \mathbb{C}$. This is a class function of G and thus lies in $R_{\mathbb{C}}(G)$. Thus, inflation $\operatorname{Infl}_{G/K}^G$ is a $\mathbb{C}$-linear map $R_{\mathbb{C}}(G/K) \to R_{\mathbb{C}}(G)$. It restricts to a $\mathbb{Z}$-linear map $R(G/K) \to R(G)$, since it is clear that every $\mathbb{C}(G/K)$-module U satisfies $\operatorname{Infl}_{G/K}^G \chi_U = \chi_{\operatorname{Infl}_{G/K}^G U}$.

We can also use (4.1.12) (or (4.1.13)) as inspiration for defining a "K-fixed space construction" on class functions. Explicitly, for every class function $f \in R_{\mathbb{C}}(G)$, we define a class function $f^K \in R_{\mathbb{C}}(G/K)$ by

$$f^K(gK) = \frac{1}{|K|} \sum_{k \in K} f(gk) = \frac{1}{|K|} \sum_{h \in gK} f(h).$$

The map $(\cdot)^K : R_{\mathbb{C}}(G) \to R_{\mathbb{C}}(G/K),\ f \mapsto f^K$ is $\mathbb{C}$-linear, and restricts to a $\mathbb{Z}$-linear map $R(G) \to R(G/K)$. Again, we have a compatibility with the K-fixed point construction on modules: We have $\chi_{V^K} = (\chi_V)^K$ for every $\mathbb{C}G$-module V.

Taking characters in (4.1.11), we obtain

$$(\operatorname{Infl}_{G/K}^G \chi_U, \chi_V)_G = (\chi_U, \chi_V^K)_{G/K} \tag{4.1.14}$$

for any $\mathbb{C}[G/K]$-module U and any $\mathbb{C}G$-module V (since $\chi_{\operatorname{Infl}_{G/K}^G U} = \operatorname{Infl}_{G/K}^G \chi_U$ and $\chi_{V^K} = (\chi_V)^K$). By $\mathbb{Z}$-linearity, this implies that

$$\left(\operatorname{Infl}_{G/K}^G \alpha, \beta\right)_G = \left(\alpha, \beta^K\right)_{G/K}$$

for any class functions $\alpha \in R_{\mathbb{C}}(G/K)$ and $\beta \in R_{\mathbb{C}}(G)$.

There is also an analogue of (4.1.6):

Lemma 4.1.8. *Let G_1 and G_2 be two groups, and $K_1 < G_1$ and $K_2 < G_2$ be two respective subgroups. Let U_i be a $\mathbb{C}G_i$-module for each $i \in \{1, 2\}$. Then,*

$$(U_1 \otimes U_2)^{K_1 \times K_2} = U_1^{K_1} \otimes U_2^{K_2} \tag{4.1.15}$$

(as subspaces of $U_1 \otimes U_2$).

[217] For another proof of (4.1.12), see Exercise 4.1.14(l).

Proof. The subgroup $K_1 = K_1 \times 1$ of $G_1 \times G_2$ acts on $U_1 \otimes U_2$, and its fixed points are $(U_1 \otimes U_2)^{K_1} = U_1^{K_1} \otimes U_2$ (because for a $\mathbb{C}K_1$-module, tensoring with U_2 is the same as taking a direct power, which clearly commutes with taking fixed points). Similarly, $(U_1 \otimes U_2)^{K_2} = U_1 \otimes U_2^{K_2}$. Now,

$$(U_1 \otimes U_2)^{K_1 \times K_2} = (U_1 \otimes U_2)^{K_1} \cap (U_1 \otimes U_2)^{K_2} = \left(U_1^{K_1} \otimes U_2\right) \cap \left(U_1 \otimes U_2^{K_2}\right) = U_1^{K_1} \otimes U_2^{K_2}$$

according to the known linear-algebraic fact stating that if P and Q are subspaces of two vector spaces U and V, respectively, then $(P \otimes V) \cap (U \otimes Q) = P \otimes Q$. □

Exercise 4.1.9. (a) Let G_1 and G_2 be two groups. Let V_i and W_i be finite-dimensional $\mathbb{C}G_i$-modules for every $i \in \{1, 2\}$. Prove that the $\mathbb{C}$-linear map

$$\operatorname{Hom}_{\mathbb{C}G_1}(V_1, W_1) \otimes \operatorname{Hom}_{\mathbb{C}G_2}(V_2, W_2) \to \operatorname{Hom}_{\mathbb{C}[G_1 \times G_2]}(V_1 \otimes V_2, W_1 \otimes W_2)$$

sending each tensor $f \otimes g$ to the tensor product $f \otimes g$ of homomorphisms is a vector space isomorphism.

(b) Use part (a) to give a new proof of (4.1.2).

As an aside, (4.1.10) has a "dual" analogue:

Exercise 4.1.10. Let G be a finite group, and let $K \triangleleft G$ and $H < G$. Let U be a $\mathbb{C}H$-module. As usual, regard $H/(H \cap K)$ as a subgroup of G/K. Show that $\left(\operatorname{Ind}_H^G U\right)^K \cong \operatorname{Ind}_{H/(H\cap K)}^{G/K}\left(U^{H \cap K}\right)$ as $\mathbb{C}[G/K]$-modules.

Inflation also "commutes" with induction:

Exercise 4.1.11. Let G be a finite group, and let $K < H < G$ be such that $K \triangleleft G$. Thus, automatically, $K \triangleleft H$, and we regard the quotient H/K as a subgroup of G/K. Let V be a $\mathbb{C}[H/K]$-module. Show that $\operatorname{Infl}_{G/K}^G \operatorname{Ind}_{H/K}^{G/K} V \cong \operatorname{Ind}_H^G \operatorname{Infl}_{H/K}^H V$ as $\mathbb{C}G$-modules.

Exercise 4.1.12. Let G be a finite group, and let $K \triangleleft G$. Let V be a $\mathbb{C}G$-module. Let $I_{V,K}$ denote the $\mathbb{C}$-vector subspace of V spanned by all elements of the form $v - kv$ for $k \in K$ and $v \in V$.

(a) Show that $I_{V,K}$ is a $\mathbb{C}G$-submodule of V.

(b) Let V_K denote the quotient $\mathbb{C}G$-module $V/I_{V,K}$. (This module is occasionally called the *K-coinvariant module of V*, a name it sadly shares with at least two other non-equivalent constructions in algebra.) Show that $V_K \cong \operatorname{Infl}_{G/K}^G\left(V^K\right)$ as $\mathbb{C}G$-modules. (Use $\operatorname{char}\mathbb{C} = 0$.)

In the remainder of this subsection, we shall briefly survey generalized notions of induction and restriction, defined in terms of a group homomorphism ρ rather than in terms of a group G and a subgroup H. These generalized notions (defined by van Leeuwen in [128, §2.2]) will not be used in the rest of these notes, but they shed some new light on the facts about induction, restriction, inflation and fixed point construction discussed above. (In particular, they reveal that some of said facts have common generalizations.)

The reader might have noticed that the definitions of inflation and of restriction (both for characters and for modules) are similar. In fact, they both are particular cases of the following construction:

Remark 4.1.13. Let G and H be two finite groups, and let $\rho : H \to G$ be a group homomorphism.

- If $f \in R_{\mathbb{C}}(G)$, then the *ρ-restriction* $\operatorname{Res}_\rho f$ of f is defined as the map $f \circ \rho : H \to \mathbb{C}$. This map is easily seen to belong to $R_{\mathbb{C}}(H)$.
- If V is a $\mathbb{C}G$-module, then the *ρ-restriction* $\operatorname{Res}_\rho V$ of V is the $\mathbb{C}H$-module with ground space V and action given by

$$h \cdot v = \rho(h) \cdot v \qquad \text{for every } h \in H \text{ and } v \in V.$$

This construction generalizes both inflation and restriction: If H is a subgroup of G, and if $\rho : H \to G$ is the inclusion map, then $\operatorname{Res}_\rho f = \operatorname{Res}_H^G f$ (for any $f \in R_{\mathbb{C}}(G)$) and $\operatorname{Res}_\rho V = \operatorname{Res}_H^G V$ (for any $\mathbb{C}G$-module V). If, instead, we have $G = H/K$ for a normal subgroup K of H, and if $\rho : H \to G$ is the projection map, then $\operatorname{Res}_\rho f = \operatorname{Infl}_{H/K}^H f$ (for any $f \in R_{\mathbb{C}}(H/K)$) and $\operatorname{Res}_\rho V = \operatorname{Infl}_{H/K}^H V$ (for any $\mathbb{C}[H/K]$-module V).

A subtler observation is that induction and fixed point construction can be generalized by a common notion. This is the subject of Exercise 4.1.14 below.

Exercise 4.1.14. Let G and H be two finite groups, and let $\rho : H \to G$ be a group homomorphism. We introduce the following notations:

- If $f \in R_{\mathbb{C}}(H)$, then the *ρ-induction* $\operatorname{Ind}_\rho f$ of f is a map $G \to \mathbb{C}$ which is defined as follows:

$$(\operatorname{Ind}_\rho f)(g) = \frac{1}{|H|} \sum_{\substack{(h,k) \in H \times G; \\ k\rho(h)k^{-1}=g}} f(h) \qquad \text{for every } g \in G.$$

- If U is a $\mathbb{C}H$-module, then the *ρ-induction* $\operatorname{Ind}_\rho U$ of U is defined as the $\mathbb{C}G$-module $\mathbb{C}G \otimes_{\mathbb{C}H} U$, where $\mathbb{C}G$ is regarded as a $(\mathbb{C}G, \mathbb{C}H)$-bimodule according to the following rule: The left $\mathbb{C}G$-module structure on $\mathbb{C}G$ is plain multiplication inside $\mathbb{C}G$; the right $\mathbb{C}H$-module structure on $\mathbb{C}G$ is induced by the $\mathbb{C}$-algebra homomorphism $\mathbb{C}[\rho] : \mathbb{C}H \to \mathbb{C}G$ (thus, it is explicitly given by $\gamma\eta = \gamma \cdot (\mathbb{C}[\rho])\eta$ for all $\gamma \in \mathbb{C}G$ and $\eta \in \mathbb{C}H$).

Prove the following properties of this construction:

(a) For every $f \in R_{\mathbb{C}}(H)$, we have $\operatorname{Ind}_\rho f \in R_{\mathbb{C}}(G)$.

(b) For any finite-dimensional $\mathbb{C}H$-module U, we have $\chi_{\operatorname{Ind}_\rho U} = \operatorname{Ind}_\rho \chi_U$.

(c) If H is a subgroup of G, and if $\rho : H \to G$ is the inclusion map, then $\operatorname{Ind}_\rho f = \operatorname{Ind}_H^G f$ for every $f \in R_{\mathbb{C}}(H)$.

(d) If H is a subgroup of G, and if $\rho : H \to G$ is the inclusion map, then $\operatorname{Ind}_\rho U = \operatorname{Ind}_H^G U$ for every $\mathbb{C}H$-module U.

(e) If $G = H/K$ for some normal subgroup K of H, and if $\rho : H \to G$ is the projection map, then $\operatorname{Ind}_\rho f = f^K$ for every $f \in R_{\mathbb{C}}(H)$.

(f) If $G = H/K$ for some normal subgroup K of H, and if $\rho : H \to G$ is the projection map, then $\operatorname{Ind}_\rho U \cong U^K$ for every $\mathbb{C}H$-module U.

(g) Any class functions $\alpha \in R_{\mathbb{C}}(H)$ and $\beta \in R_{\mathbb{C}}(G)$ satisfy

$$(\operatorname{Ind}_\rho \alpha, \beta)_G = (\alpha, \operatorname{Res}_\rho \beta)_H \tag{4.1.16}$$

and

$$\langle \operatorname{Ind}_\rho \alpha, \beta \rangle_G = \langle \alpha, \operatorname{Res}_\rho \beta \rangle_H . \tag{4.1.17}$$

(See Remark 4.1.13 for the definition of $\operatorname{Res}_\rho \beta$.)

(h) We have $\operatorname{Hom}_{\mathbb{C}G}(\operatorname{Ind}_\rho U, V) \cong \operatorname{Hom}_{\mathbb{C}H}(U, \operatorname{Res}_\rho V)$ for every $\mathbb{C}H$-module U and every $\mathbb{C}G$-module V. (See Remark 4.1.13 for the definition of $\operatorname{Res}_\rho V$.)

(i) Similarly to how we made $\mathbb{C}G$ into a $(\mathbb{C}G, \mathbb{C}H)$-bimodule, let us make $\mathbb{C}G$ into a $(\mathbb{C}H, \mathbb{C}G)$-bimodule (so the right $\mathbb{C}G$-module structure is plain multiplication inside $\mathbb{C}G$, whereas the left $\mathbb{C}H$-module structure is induced by the $\mathbb{C}$-algebra homomorphism $\mathbb{C}[\rho] : \mathbb{C}H \to \mathbb{C}G$). If U is any $\mathbb{C}H$-module, then the $\mathbb{C}G$-module $\operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, U)$ (defined as in Exercise 4.1.4 using the $(\mathbb{C}H, \mathbb{C}G)$-bimodule structure on $\mathbb{C}G$) is isomorphic to $\operatorname{Ind}_\rho U$.

(j) We have $\operatorname{Hom}_{\mathbb{C}G}(U, \operatorname{Ind}_\rho V) \cong \operatorname{Hom}_{\mathbb{C}H}(\operatorname{Res}_\rho U, V)$ for every $\mathbb{C}G$-module U and every $\mathbb{C}H$-module V. (See Remark 4.1.13 for the definition of $\operatorname{Res}_\rho V$.)

Furthermore:

(k) Use the above to prove the formula (4.1.3).

(l) Use the above to prove the formula (4.1.12).

[**Hint:** Part (b) of this exercise is hard. To solve it, it is useful to have a way of computing the trace of a linear operator without knowing a basis of the vector space it is acting on. There is a way to do this using a "finite dual generating system", which is a somewhat less restricted notion than that of a basis[218]. Try to create a finite dual generating system for $\operatorname{Ind}_\rho U$ from one for U (and from the group G), and then use it to compute $\chi_{\operatorname{Ind}_\rho U}$.

[218] More precisely: Let $\mathbb{K}$ be a field, and V be a $\mathbb{K}$-vector space. A *finite dual generating system* for V means a triple $(I, (a_i)_{i \in I}, (f_i)_{i \in I})$, where

- I is a finite set;
- $(a_i)_{i \in I}$ is a family of elements of V;
- $(f_i)_{i \in I}$ is a family of elements of V^* (where V^* means $\operatorname{Hom}_{\mathbb{K}}(V, \mathbb{K})$)

such that every $v \in V$ satisfies $v = \sum_{i \in I} f_i(v) a_i$. For example, if $(e_j)_{j \in J}$ is a finite basis of the vector space V, and if $\left(e_j^*\right)_{j \in J}$ is the basis of V^* dual to this basis $(e_j)_{j \in J}$, then $\left(J, (e_j)_{j \in J}, \left(e_j^*\right)_{j \in J}\right)$ is a finite dual generating system for V; however, most finite dual generating systems are not obtained this way.

The solution of part (i) is a modification of the solution of Exercise 4.1.4, but complicated by the fact that H is no longer (necessarily) a subgroup of G. Part (f) can be solved by similar arguments, or using part (i), or using Exercise 4.1.12(b).]

The result of Exercise 4.1.14(h) generalizes (4.1.7) (because of Exercise 4.1.14(d)), but also generalizes (4.1.11) (due to Exercise 4.1.14(f)). Similarly, Exercise 4.1.14(g) generalizes both (4.1.9) and (4.1.14). Similarly, Exercise 4.1.14(i) generalizes Exercise 4.1.4, and Exercise 4.1.14(j) generalizes Exercise 4.1.6.

Similarly, Exercise 4.1.3 is generalized by the following exercise:

Exercise 4.1.15. Let G_1, G_2, H_1 and H_2 be four finite groups. Let $\rho_1 : H_1 \to G_1$ and $\rho_2 : H_2 \to G_2$ be two group homomorphisms. These two homomorphisms clearly induce a group homomorphism $\rho_1 \times \rho_2 : H_1 \times H_2 \to G_1 \times G_2$. Let U_1 be a $\mathbb{C}H_1$-module, and U_2 be a $\mathbb{C}H_2$-module. Show that

$$\operatorname{Ind}_{\rho_1 \times \rho_2} (U_1 \otimes U_2) \cong (\operatorname{Ind}_{\rho_1} U_1) \otimes (\operatorname{Ind}_{\rho_2} U_2)$$

as $\mathbb{C}[G_1 \times G_2]$-modules.

The Ind_ρ and Res_ρ operators behave "functorially" with respect to composition. Here is what this means:

Exercise 4.1.16. Let G, H and I be three finite groups. Let $\rho : H \to G$ and $\tau : I \to H$ be two group homomorphisms.

(a) We have $\operatorname{Ind}_\rho \operatorname{Ind}_\tau U \cong \operatorname{Ind}_{\rho\circ\tau} U$ for every $\mathbb{C}I$-module U.
(b) We have $\operatorname{Ind}_\rho \operatorname{Ind}_\tau f = \operatorname{Ind}_{\rho\circ\tau} f$ for every $f \in R_{\mathbb{C}}(I)$.
(c) We have $\operatorname{Res}_\tau \operatorname{Res}_\rho V = \operatorname{Res}_{\rho\circ\tau} V$ for every $\mathbb{C}G$-module V.
(d) We have $\operatorname{Res}_\tau \operatorname{Res}_\rho f = \operatorname{Res}_{\rho\circ\tau} f$ for every $f \in R_{\mathbb{C}}(G)$.

Exercise 4.1.16(a), of course, generalizes Exercise 4.1.2.

4.1.7. *Semidirect products.* Recall that a *semidirect product* is a group $G \ltimes K$ having two subgroups G, K with

- $K \lhd (G \ltimes K)$ is a normal subgroup,
- $G \ltimes K = GK = KG$, and
- $G \cap K = \{e\}$.

In this setting one has two interesting adjoint constructions, applied in Section 4.5.

Proposition 4.1.17. *Fix a $\mathbb{C}[G \ltimes K]$-module V.*

(i) *For any $\mathbb{C}G$-module U, one has $\mathbb{C}[G \ltimes K]$-module structure*

$$\Phi(U) := U \otimes V,$$

determined via

$$k(u \otimes v) = u \otimes k(v),$$
$$g(u \otimes v) = g(u) \otimes g(v).$$

(ii) *For any $\mathbb{C}[G \ltimes K]$-module W, one has $\mathbb{C}G$-module structure*

$$\Psi(W) := \operatorname{Hom}_{\mathbb{C}K}(\operatorname{Res}^{G\ltimes K}_K V, \operatorname{Res}^{G\ltimes K}_K W),$$

determined via $g(\varphi) = g \circ \varphi \circ g^{-1}$.

(iii) *The maps*

$$\mathbb{C}G - \text{mods} \underset{\Psi}{\overset{\Phi}{\rightleftharpoons}} \mathbb{C}[G \ltimes K] - \text{mods}$$

The crucial observation is now that if $\left(I, (a_i)_{i\in I}, (f_i)_{i\in I}\right)$ is a finite dual generating system for a vector space V, and if T is an endomorphism of V, then

$$\operatorname{trace} T = \sum_{i\in I} f_i (Ta_i).$$

Prove this!

are adjoint in the sense that one has an isomorphism

$$\begin{tikzcd}
\operatorname{Hom}_{\mathbb{C}G}(U, \Psi(W)) \arrow[r] \arrow[d, equal] & \operatorname{Hom}_{\mathbb{C}[G \ltimes K]}(\Phi(U), W) \arrow[d, equal] \\
\operatorname{Hom}_{\mathbb{C}G}(U, \operatorname{Hom}_{\mathbb{C}K}(\operatorname{Res}^{G \ltimes K}_{K} V, \operatorname{Res}^{G \ltimes K}_{K} W)) & \operatorname{Hom}_{\mathbb{C}[G \ltimes K]}(U \otimes V, W), \\
\varphi \arrow[r, mapsto] & \overline{\varphi}(u \otimes v) := \varphi(u)(v).
\end{tikzcd}$$

(iv) *One has a $\mathbb{C}G$-module isomorphism*

$$(\Psi \circ \Phi)(U) \cong U \otimes \operatorname{End}_{\mathbb{C}K}(\operatorname{Res}^{G \ltimes K}_{K} V).$$

In particular, if $\operatorname{Res}^{G \ltimes K}_{K} V$ is a simple $\mathbb{C}K$-module, then $(\Psi \circ \Phi)(U) \cong U$.

Proof. These are mostly straightforward exercises in the definitions. To check assertion (iv), for example, note that K acts only in the right tensor factor in $\operatorname{Res}^{G \ltimes K}_{K}(U \otimes V)$, and hence as $\mathbb{C}G$-modules one has

$$\begin{aligned}
(\Psi \circ \Phi)(U) &= \operatorname{Hom}_{\mathbb{C}K}(\operatorname{Res}^{G \ltimes K}_{K} V,\ \operatorname{Res}^{G \ltimes K}_{K}(U \otimes V)) \\
&= \operatorname{Hom}_{\mathbb{C}K}(\operatorname{Res}^{G \ltimes K}_{K} V,\ U \otimes \operatorname{Res}^{G \ltimes K}_{K} V) \\
&= U \otimes \operatorname{Hom}_{\mathbb{C}K}(\operatorname{Res}^{G \ltimes K}_{K} V,\ \operatorname{Res}^{G \ltimes K}_{K} V) \\
&= U \otimes \operatorname{End}_{\mathbb{C}K}(\operatorname{Res}^{G \ltimes K}_{K} V).
\end{aligned}$$

□

4.2. **Three towers of groups.** Here we consider three towers of groups

$$G_* = (G_0 < G_1 < G_2 < G_3 < \cdots)$$

where either

- $G_n = \mathfrak{S}_n$, the *symmetric group*[219], or
- $G_n = \mathfrak{S}_n[\Gamma]$, the *wreath product* of the symmetric group with some arbitrary finite group Γ, or
- $G_n = GL_n(\mathbb{F}_q)$, the finite general linear group[220].

Here the wreath product $\mathfrak{S}_n[\Gamma]$ can be thought of informally as the group of *monomial* $n \times n$ matrices whose nonzero entries lie in Γ, that is, $n \times n$ matrices having exactly one nonzero entry in each row and column, and that entry is an element of Γ. E.g.

$$\begin{bmatrix} 0 & g_2 & 0 \\ g_1 & 0 & 0 \\ 0 & 0 & g_3 \end{bmatrix} \begin{bmatrix} 0 & 0 & g_6 \\ 0 & g_5 & 0 \\ g_4 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & g_2 g_5 & 0 \\ 0 & 0 & g_1 g_6 \\ g_3 g_4 & 0 & 0 \end{bmatrix}.$$

More formally, $\mathfrak{S}_n[\Gamma]$ is the semidirect product $\mathfrak{S}_n \ltimes \Gamma^n$ in which $\mathfrak{S}_n$ acts on Γ^n via $\sigma(\gamma_1, \ldots, \gamma_n) = (\gamma_{\sigma^{-1}(1)}, \ldots, \gamma_{\sigma^{-1}(n)})$.

For each of the three towers G_*, there are embeddings $G_i \times G_j \hookrightarrow G_{i+j}$ and we introduce maps $\operatorname{ind}^{i+j}_{i,j}$ taking $\mathbb{C}[G_i \times G_j]$-modules to $\mathbb{C}G_{i+j}$-modules, as well as maps $\operatorname{res}^{i+j}_{i,j}$ carrying modules in the reverse direction which are adjoint:

$$\operatorname{Hom}_{\mathbb{C}G_{i+j}}(\operatorname{ind}^{i+j}_{i,j} U, V) = \operatorname{Hom}_{\mathbb{C}[G_i \times G_j]}(U, \operatorname{res}^{i+j}_{i,j} V). \tag{4.2.1}$$

Definition 4.2.1. For $G_n = \mathfrak{S}_n$, one embeds $\mathfrak{S}_i \times \mathfrak{S}_j$ into $\mathfrak{S}_{i+j}$ as the permutations that permute $\{1, 2, \ldots, i\}$ and $\{i+1, i+2, \ldots, i+j\}$ separately. Here one defines

$$\begin{aligned}
\operatorname{ind}^{i+j}_{i,j} &:= \operatorname{Ind}^{\mathfrak{S}_{i+j}}_{\mathfrak{S}_i \times \mathfrak{S}_j}, \\
\operatorname{res}^{i+j}_{i,j} &:= \operatorname{Res}^{\mathfrak{S}_{i+j}}_{\mathfrak{S}_i \times \mathfrak{S}_j}.
\end{aligned}$$

[219] The symmetric group $\mathfrak{S}_0$ is the group of all permutations of the empty set $\{1, 2, \ldots, 0\} = \varnothing$. It is a trivial group. (Note that $\mathfrak{S}_1$ is also a trivial group.)

[220] The group $GL_0(\mathbb{F}_q)$ is a trivial group, consisting of the empty 0×0 matrix.

For $G_n = \mathfrak{S}_n[\Gamma]$, similarly embed $\mathfrak{S}_i[\Gamma] \times \mathfrak{S}_j[\Gamma]$ into $\mathfrak{S}_{i+j}[\Gamma]$ as block monomial matrices whose two diagonal blocks have sizes i, j respectively, and define

$$\begin{aligned} \mathrm{ind}_{i,j}^{i+j} &:= \mathrm{Ind}_{\mathfrak{S}_i[\Gamma]\times\mathfrak{S}_j[\Gamma]}^{\mathfrak{S}_{i+j}[\Gamma]}, \\ \mathrm{res}_{i,j}^{i+j} &:= \mathrm{Res}_{\mathfrak{S}_i[\Gamma]\times\mathfrak{S}_j[\Gamma]}^{\mathfrak{S}_{i+j}[\Gamma]}. \end{aligned}$$

For $G_n = GL_n(\mathbb{F}_q)$, which we will denote just GL_n, similarly embed $GL_i \times GL_j$ into GL_{i+j} as block diagonal matrices whose two diagonal blocks have sizes i, j respectively. However, one also introduces as an intermediate the *parabolic subgroup* $P_{i,j}$ consisting of the block upper-triangular matrices of the form

$$\begin{bmatrix} g_i & \ell \\ 0 & g_j \end{bmatrix}$$

where g_i, g_j lie in GL_i, GL_j, respectively, and ℓ in $\mathbb{F}_q^{i\times j}$ is arbitrary. One has a quotient map $P_{i,j} \to GL_i \times GL_j$ whose kernel $K_{i,j}$ is the set of matrices of the form

$$\begin{bmatrix} I_i & \ell \\ 0 & I_j \end{bmatrix}$$

with ℓ again arbitrary. Here one defines

$$\begin{aligned} \mathrm{ind}_{i,j}^{i+j} &:= \mathrm{Ind}_{P_{i,j}}^{GL_{i+j}} \mathrm{Infl}_{GL_i\times GL_j}^{P_{i,j}}, \\ \mathrm{res}_{i,j}^{i+j} &:= \left(\mathrm{Res}_{P_{i,j}}^{GL_{i+j}}(-)\right)^{K_{i,j}}. \end{aligned}$$

In the case $G_n = GL_n$, the operation $\mathrm{ind}_{i,j}^{i+j}$ is sometimes called *parabolic induction* or *Harish-Chandra induction*. The operation $\mathrm{res}_{i,j}^{i+j}$ is essentially just the $K_{i,j}$-fixed point construction $V \mapsto V^{K_{i,j}}$. However writing it as the above two-step composite makes it more obvious, (via (4.1.7) and (4.1.11)) that $\mathrm{res}_{i,j}^{i+j}$ is again adjoint to $\mathrm{ind}_{i,j}^{i+j}$.

Definition 4.2.2. For each of the three towers G_*, define a graded $\mathbb{Z}$-module

$$A := A(G_*) = \bigoplus_{n\geq 0} R(G_n)$$

with a bilinear form $(\cdot,\cdot)_A$ whose restriction to $A_n := R(G_n)$ is the usual form $(\cdot,\cdot)_{G_n}$, and such that $\Sigma := \bigsqcup_{n\geq 0} \mathrm{Irr}(G_n)$ gives an orthonormal $\mathbb{Z}$-basis. Notice that $A_0 = \mathbb{Z}$ has its basis element 1 equal to the unique irreducible character of the trivial group G_0.

Bearing in mind that $A_n = R(G_n)$ and

$$A_i \otimes A_j = R(G_i) \otimes R(G_j) \cong R(G_i \times G_j),$$

one then has candidates for product and coproduct defined by

$$\begin{array}{rccc} m := \mathrm{ind}_{i,j}^{i+j}: & A_i \otimes A_j & \longrightarrow & A_{i+j}, \\ \Delta := \bigoplus_{i+j=n} \mathrm{res}_{i,j}^{i+j}: & A_n & \longrightarrow & \bigoplus_{i+j=n} A_i \otimes A_j. \end{array}$$

The coassociativity of Δ is an easy consequence of transitivity of the constructions of restriction and fixed points[221]. We could derive the associativity of m from the transitivity of induction and inflation, but this would be more complicated[222]; we will instead prove it differently.

[221] More precisely, using this transitivity, it is easily reduced to proving that $K_{i+j,k} \cdot (K_{i,j} \times \{I_k\}) = K_{i,j+k} \cdot (\{I_i\} \times K_{j,k})$ (an equality between subgroups of GL_{i+j+k}) for any three nonnegative integers i, j, k. But this equality can be proven by realizing that both of its sides equal the set of all block matrices of the form $\begin{pmatrix} I_i & \ell & \ell' \\ 0 & I_j & \ell'' \\ 0 & 0 & I_k \end{pmatrix}$ with ℓ, ℓ' and ℓ'' being matrices of sizes $i \times j$, $i \times k$ and $j \times k$, respectively.

[222] See Exercise 4.3.11(c) for such a derivation.

We first show that the maps m and Δ are adjoint with respect to the forms $(\cdot,\cdot)_A$ and $(\cdot,\cdot)_{A\otimes A}$. In fact, if U, V, W are modules over $\mathbb{C}G_i$, $\mathbb{C}G_j$, $\mathbb{C}G_{i+j}$, respectively, then we can write the $\mathbb{C}[G_i \times G_j]$-module $\operatorname{res}^{i+j}_{i,j} W$ as a direct sum $\bigoplus_k X_k \otimes Y_k$ with X_k being $\mathbb{C}G_i$-modules and Y_k being $\mathbb{C}G_j$-modules; we then have

$$\operatorname{res}^{i+j}_{i,j} \chi_W = \sum_k \chi_{X_k} \otimes \chi_{Y_k} \tag{4.2.2}$$

and

$$\begin{aligned}
(m(\chi_U \otimes \chi_V), \chi_W)_A &= \left(\operatorname{ind}^{i+j}_{i,j}(\chi_{U\otimes V}), \chi_W\right)_A = \left(\operatorname{ind}^{i+j}_{i,j}(\chi_{U\otimes V}), \chi_W\right)_{G_{i+j}} \\
&= \left(\chi_{U\otimes V}, \operatorname{res}^{i+j}_{i,j}\chi_W\right)_{G_i\times G_j} = \left(\chi_{U\otimes V}, \sum_k \chi_{X_k}\otimes\chi_{Y_k}\right)_{G_i\times G_j} \\
&= \sum_k (\chi_{U\otimes V}, \chi_{X_k\otimes Y_k})_{G_i\times G_j} = \sum_k (\chi_U, \chi_{X_k})_{G_i}(\chi_V, \chi_{Y_k})_{G_j}
\end{aligned}$$

(the third equality sign follows by taking dimensions in (4.2.1) and recalling (4.1.1); the fourth equality sign follows from (4.2.2); the sixth one follows from (4.1.2)) and

$$\begin{aligned}
(\chi_U \otimes \chi_V, \Delta(\chi_W))_{A\otimes A} &= \left(\chi_U \otimes \chi_V, \operatorname{res}^{i+j}_{i,j}\chi_W\right)_{A\otimes A} = \left(\chi_U\otimes\chi_V, \sum_k \chi_{X_k}\otimes\chi_{Y_k}\right)_{A\otimes A} \\
&= \sum_k (\chi_U, \chi_{X_k})_A (\chi_V, \chi_{Y_k})_A = \sum_k (\chi_U, \chi_{X_k})_{G_i}(\chi_V, \chi_{Y_k})_{G_j}
\end{aligned}$$

(the first equality sign follows by removing all terms in $\Delta(\chi_W)$ whose scalar product with $\chi_U \otimes \chi_V$ vanishes for reasons of gradedness; the second equality sign follows from (4.2.2)), which in comparison yield $(m(\chi_U\otimes\chi_V), \chi_W)_A = (\chi_U\otimes\chi_V, \Delta(\chi_W))_{A\otimes A}$, thus showing that m and Δ are adjoint maps. Therefore, m is associative (since Δ is coassociative).

Endowing $A = \bigoplus_{n\geq 0} R(G_n)$ with the obvious unit and counit maps, it thus becomes a graded, finite-type $\mathbb{Z}$-algebra and $\mathbb{Z}$-coalgebra.

The next section addresses the issue of why they form a bialgebra. However, assuming this for the moment, it should be clear that each of these algebras A is a PSH having $\Sigma = \bigsqcup_{n\geq 0} \operatorname{Irr}(G_n)$ as its PSH-basis. Σ is self-dual because m, Δ are defined by adjoint maps, and it is positive because m, Δ take irreducible representations to genuine representations not just virtual ones, and hence have characters which are nonnegative sums of irreducible characters.

Exercise 4.2.3. Let i, j and k be three nonnegative integers. Let U be a $\mathbb{C}\mathfrak{S}_i$-module, let V be a $\mathbb{C}\mathfrak{S}_j$-module, and let W be a $\mathbb{C}\mathfrak{S}_k$-module. Show that there are canonical $\mathbb{C}[\mathfrak{S}_i \times \mathfrak{S}_j \times \mathfrak{S}_k]$-module isomorphisms

$$\begin{aligned}
\operatorname{Ind}^{\mathfrak{S}_{i+j+k}}_{\mathfrak{S}_{i+j}\times\mathfrak{S}_k}\left(\operatorname{Ind}^{\mathfrak{S}_{i+j}}_{\mathfrak{S}_i\times\mathfrak{S}_j}(U\otimes V)\otimes W\right) &\cong \operatorname{Ind}^{\mathfrak{S}_{i+j+k}}_{\mathfrak{S}_i\times\mathfrak{S}_j\times\mathfrak{S}_k}(U\otimes V\otimes W) \\
&\cong \operatorname{Ind}^{\mathfrak{S}_{i+j+k}}_{\mathfrak{S}_i\times\mathfrak{S}_{j+k}}\left(U\otimes \operatorname{Ind}^{\mathfrak{S}_{j+k}}_{\mathfrak{S}_j\times\mathfrak{S}_k}(V\otimes W)\right).
\end{aligned}$$

(Similar statements hold for the other two towers of groups and their respective ind functors, although the one for the GL_* tower is harder to prove. See Exercise 4.3.11(a) for a more general result.)

4.3. **Bialgebra and double cosets.**

To show that the algebra and coalgebras $A = A(G_*)$ are bialgebras, the central issue is checking the pentagonal diagram in (1.3.4), that is, as maps $A\otimes A \to A\otimes A$, one has

$$\Delta\circ m = (m\otimes m)\circ(\operatorname{id}\otimes T\otimes\operatorname{id})\circ(\Delta\otimes\Delta). \tag{4.3.1}$$

In checking this, it is convenient to have a lighter notation for various subgroups of the groups G_n corresponding to compositions α.

Definition 4.3.1. (a) An *almost-composition* is a (finite) tuple $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_\ell)$ of nonnegative integers. Its *length* is defined to be ℓ and denoted by $\ell(\alpha)$; its *size* is defined to be $\alpha_1 + \alpha_2 + \cdots + \alpha_\ell$ and denoted by $|\alpha|$; its *parts* are its entries $\alpha_1, \alpha_2, \ldots, \alpha_\ell$. The almost-compositions of size n are called the *almost-compositions of n*.

(b) A *composition* is a finite tuple of positive integers. Of course, any composition is an almost-composition, and so all notions defined for almost-compositions (like size and length) make sense for compositions.

Note that any partition of n (written without trailing zeroes) is a composition of n. We write $\varnothing$ (and sometimes, sloppily, (0), when there is no danger of mistaking it for the almost-composition (0)) for the empty composition $()$.

Definition 4.3.2. Given an almost-composition $\alpha = (\alpha_1, \ldots, \alpha_\ell)$ of n, define a subgroup

$$G_\alpha \cong G_{\alpha_1} \times \cdots \times G_{\alpha_\ell} < G_n$$

via the block-diagonal embedding with diagonal blocks of sizes $(\alpha_1, \ldots, \alpha_\ell)$. This G_α is called a *Young subgroup* $\mathfrak{S}_\alpha$ when $G_n = \mathfrak{S}_n$, and a *Levi subgroup* when $G_n = GL_n$. In the case when $G_n = \mathfrak{S}_n[\Gamma]$, we also denote G_α by $\mathfrak{S}_\alpha[\Gamma]$. In the case where $G_n = GL_n$, also define the *parabolic subgroup* P_α to be the subgroup of G_n consisting of block-upper triangular matrices whose diagonal blocks have sizes $(\alpha_1, \ldots, \alpha_\ell)$, and let K_α be the kernel of the obvious surjection $P_\alpha \to G_\alpha$ which sends a block upper-triangular matrix to the tuple of its diagonal blocks whose sizes are $\alpha_1, \alpha_2, \ldots, \alpha_\ell$. Notice that $P_{(i,j)} = P_{i,j}$ for any i and j with $i + j = n$; similarly, $K_{(i,j)} = K_{i,j}$ for any i and j with $i + j = n$. We will also abbreviate $G_{(i,j)} = G_i \times G_j$ by $G_{i,j}$.

When $(\alpha_1, \alpha_2, \ldots, \alpha_\ell)$ is an almost-composition, we abbreviate $G_{(\alpha_1,\alpha_2,\ldots,\alpha_\ell)}$ by $G_{\alpha_1,\alpha_2,\ldots,\alpha_\ell}$ (and similarly for the P's).

Definition 4.3.3. Let K and H be two groups, $\tau : K \to H$ a group homomorphism, and U a $\mathbb{C}H$-module. Then, U^τ is defined as the $\mathbb{C}K$-module with ground space U and action given by $k \cdot u = \tau(k) \cdot u$ for all $k \in K$ and $u \in U$. [223] This very simple construction generalizes the definition of U^g for an element $g \in G$, where G is a group containing H as a subgroup; in fact, in this situation we have $U^g = U^\tau$, where $K = {}^gH$ and $\tau : K \to H$ is the map $k \mapsto g^{-1}kg$.

Using homogeneity, checking the bialgebra condition (4.3.1) in the homogeneous component $(A \otimes A)_n$ amounts to the following: for each pair of representations U_1, U_2 of G_{r_1}, G_{r_2} with $r_1 + r_2 = n$, and for each (c_1, c_2) with $c_1 + c_2 = n$, one must verify that

$$\begin{aligned} &\operatorname{res}^n_{c_1,c_2}\left(\operatorname{ind}^n_{r_1,r_2}(U_1 \otimes U_2)\right) \\ &\cong \bigoplus_A \left(\operatorname{ind}^{c_1}_{a_{11},a_{21}} \otimes \operatorname{ind}^{c_2}_{a_{12},a_{22}}\right)\left(\left(\operatorname{res}^{r_1}_{a_{11},a_{12}} U_1 \otimes \operatorname{res}^{r_2}_{a_{21},a_{22}} U_2\right)^{\tau_A^{-1}}\right) \end{aligned} \tag{4.3.2}$$

where the direct sum is over all matrices $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ in $\mathbb{N}^{2\times 2}$ with row sums (r_1, r_2) and column sums (c_1, c_2), and where τ_A is the obvious isomorphism between the subgroups

$$\begin{aligned} &G_{a_{11},a_{12},a_{21},a_{22}} \quad (< G_{r_1,r_2}) \qquad \text{and} \\ &G_{a_{11},a_{21},a_{12},a_{22}} \quad (< G_{c_1,c_2}) \end{aligned} \tag{4.3.3}$$

(we are using the inverse τ_A^{-1} of this isomorphism τ_A to identify modules for the first subgroup with modules for the second subgroup, according to Definition 4.3.3).

As one might guess, (4.3.2) comes from the Mackey formula (Theorem 4.1.7), once one identifies the appropriate double coset representatives. This is just as easy to do in a slightly more general setting.

Definition 4.3.4. Given almost-compositions α, β of n having lengths ℓ, m and a matrix A in $\mathbb{N}^{\ell\times m}$ with row sums α and column sums β, define a permutation w_A in $\mathfrak{S}_n$ as follows. Disjointly decompose $[n] = \{1, 2, \ldots, n\}$ into consecutive intervals of numbers

$$[n] = I_1 \sqcup \cdots \sqcup I_\ell \qquad \text{such that } |I_i| = \alpha_i$$

(so the smallest α_1 elements of $[n]$ go into I_1, the next-smallest α_2 elements of $[n]$ go into I_2, and so on). Likewise, disjointly decompose $[n]$ into consecutive intervals of numbers

$$[n] = J_1 \sqcup \cdots \sqcup J_m \qquad \text{such that } |J_j| = \beta_j.$$

[223] We have already met this $\mathbb{C}K$-module U^τ in Remark 4.1.13, where it was called $\operatorname{Res}_\tau U$.

For every $j \in [m]$, disjointly decompose J_j into consecutive intervals of numbers $J_j = J_{j,1} \sqcup J_{j,2} \sqcup \cdots \sqcup J_{j,\ell}$ such that every $i \in [\ell]$ satisfies $|J_{j,i}| = a_{ij}$. For every $i \in [\ell]$, disjointly decompose I_i into consecutive intervals of numbers $I_i = I_{i,1} \sqcup I_{i,2} \sqcup \cdots \sqcup I_{i,m}$ such that every $j \in [m]$ satisfies $|I_{i,j}| = a_{ij}$. Now, for every $i \in [\ell]$ and $j \in [m]$, let $\pi_{i,j}$ be the increasing bijection from $J_{j,i}$ to $I_{i,j}$ (this is well-defined since these two sets both have cardinality a_{ij}). The disjoint union of these bijections $\pi_{i,j}$ over all i and j is a bijection $[n] \to [n]$ (since the disjoint union of the sets $J_{j,i}$ over all i and j is $[n]$, and so is the disjoint union of the sets $I_{i,j}$), that is, a permutation of $[n]$; this permutation is what we call w_A.

Example 4.3.5. Taking $n = 9$ and $\alpha = (4,5), \beta = (3,4,2)$, one has

$$\begin{aligned} I_1 &= \{1,2,3,4\}, \quad I_2 = \{5,6,7,8,9\}, \\ J_1 &= \{1,2,3\}, \quad J_2 = \{4,5,6,7\}, \quad J_3 = \{8,9\}. \end{aligned}$$

Then one possible matrix A having row and column sums α, β is $A = \begin{bmatrix} 2 & 2 & 0 \\ 1 & 2 & 2 \end{bmatrix}$, and its associated permutation w_A written in two-line notation is

$$\begin{pmatrix} 1 & 2 & 3 & | & 4 & 5 & 6 & 7 & | & 8 & 9 \\ \underline{1} & \underline{2} & \underline{\underline{5}} & | & \underline{3} & \underline{4} & \underline{\underline{6}} & \underline{\underline{7}} & | & \underline{\underline{8}} & \underline{\underline{9}} \end{pmatrix}$$

with vertical lines dividing the sets J_j on top, and with elements of I_i underlined i times on the bottom.

Remark 4.3.6. Given almost-compositions α and β of n having lengths ℓ and m, and a permutation $w \in \mathfrak{S}_n$. It is easy to see that there exists a matrix $A \in \mathbb{N}^{\ell \times m}$ satisfying $w_A = w$ if and only if the restriction of w to each J_j and the restriction of w^{-1} to each I_i are increasing. In this case, the matrix A is determined by $a_{ij} = |w(J_j) \cap I_i|$.

Among our three towers G_* of groups, the symmetric group tower ($G_n = \mathfrak{S}_n$) is the simplest one. We will now see that it also embeds into the two others, in the sense that $\mathfrak{S}_n$ embeds into $\mathfrak{S}_n[\Gamma]$ for every Γ and into $GL_n(\mathbb{F}_q)$ for every q.

First, for every $n \in \mathbb{N}$ and any group Γ, we embed the group $\mathfrak{S}_n$ into $\mathfrak{S}_n[\Gamma]$ by means of the canonical embedding $\mathfrak{S}_n \to \mathfrak{S}_n \ltimes \Gamma^n = \mathfrak{S}_n[\Gamma]$. If we regard elements of $\mathfrak{S}_n[\Gamma]$ as $n \times n$ monomial matrices with nonzero entries in Γ, then this boils down to identifying every $\pi \in \mathfrak{S}_n$ with the permutation matrix of π (in which the 1's are read as the neutral element of Γ). If α is an almost-composition of n, then this embedding $\mathfrak{S}_n \to \mathfrak{S}_n[\Gamma]$ makes the subgroup $\mathfrak{S}_\alpha$ of $\mathfrak{S}_n$ become a subgroup of $\mathfrak{S}_n[\Gamma]$, more precisely a subgroup of $\mathfrak{S}_\alpha[\Gamma] < \mathfrak{S}_n[\Gamma]$.

For every $n \in \mathbb{N}$ and every q, we embed the group $\mathfrak{S}_n$ into $GL_n(\mathbb{F}_q)$ by identifying every permutation $\pi \in \mathfrak{S}_n$ with its permutation matrix in $GL_n(\mathbb{F}_q)$. If α is an almost-composition of n, then this embedding makes the subgroup $\mathfrak{S}_\alpha$ of $\mathfrak{S}_n$ become a subgroup of $GL_n(\mathbb{F}_q)$. If we let $G_n = GL_n(\mathbb{F}_q)$, then $\mathfrak{S}_\alpha < G_\alpha < P_\alpha$.

The embeddings we have just defined commute with the group embeddings $G_n < G_{n+1}$ on both sides.

Proposition 4.3.7. *The permutations $\{w_A\}$, as A runs over all matrices in $\mathbb{N}^{\ell \times m}$ having row sums α and column sums β, give*

(a) *a system of double coset representatives for $\mathfrak{S}_\alpha \backslash \mathfrak{S}_n / \mathfrak{S}_\beta$;*
(b) *a system of double coset representatives for $\mathfrak{S}_\alpha[\Gamma] \backslash \mathfrak{S}_n[\Gamma] / \mathfrak{S}_\beta[\Gamma]$;*
(c) *a system of double coset representatives for $P_\alpha \backslash GL_n / P_\beta$.*

Proof. (a) We give an algorithm to show that every double coset $\mathfrak{S}_\alpha w \mathfrak{S}_\beta$ contains some w_A. Start by altering w within its coset $w\mathfrak{S}_\beta$, that is, by permuting the *positions* within each set J_j, to obtain a representative w' for $w\mathfrak{S}_\beta$ in which each set $w'(J_j)$ appears in increasing order in the second line of the two-line notation for w'. Then alter w' within its coset $\mathfrak{S}_\alpha w'$, that is, by permuting the *values* within each set I_i, to obtain a representative w_A having the elements of each set I_i appearing in increasing order in the second line; because the values within each set I_i are consecutive, this alteration will not ruin the property that one had each set

$w'(J_j)$ appearing in increasing order. For example, one might have

$$\begin{aligned}
w &= \begin{pmatrix} 1 & 2 & 3 & | & 4 & 5 & 6 & 7 & | & 8 & 9 \\ \underline{4} & \underline{\underline{8}} & \underline{2} & | & \underline{\underline{5}} & \underline{3} & \underline{\underline{9}} & \underline{1} & | & \underline{\underline{7}} & \underline{\underline{6}} \end{pmatrix}, \\
w' &= \begin{pmatrix} 1 & 2 & 3 & | & 4 & 5 & 6 & 7 & | & 8 & 9 \\ \underline{2} & \underline{4} & \underline{\underline{8}} & | & \underline{1} & \underline{3} & \underline{\underline{5}} & \underline{\underline{9}} & | & \underline{\underline{6}} & \underline{\underline{7}} \end{pmatrix} \in w\mathfrak{S}_\beta, \\
w_A &= \begin{pmatrix} 1 & 2 & 3 & | & 4 & 5 & 6 & 7 & | & 8 & 9 \\ \underline{1} & \underline{2} & \underline{\underline{5}} & | & \underline{3} & \underline{4} & \underline{\underline{6}} & \underline{\underline{7}} & | & \underline{\underline{8}} & \underline{\underline{9}} \end{pmatrix} \in \mathfrak{S}_\alpha w' \subset \mathfrak{S}_\alpha w' \mathfrak{S}_\beta = \mathfrak{S}_\alpha w \mathfrak{S}_\beta.
\end{aligned}$$

Next note that $\mathfrak{S}_\alpha w_A \mathfrak{S}_\beta = \mathfrak{S}_\alpha w_B \mathfrak{S}_\beta$ implies $A = B$, since the quantities

$$a_{i,j}(w) := |w(J_j) \cap I_i|$$

are easily seen to be constant on double cosets $\mathfrak{S}_\alpha w \mathfrak{S}_\beta$.

(b) Double coset representatives for $\mathfrak{S}_\alpha \backslash \mathfrak{S}_n / \mathfrak{S}_\beta$ should also provide double coset representatives for $\mathfrak{S}_\alpha[\Gamma] \backslash \mathfrak{S}_n[\Gamma] / \mathfrak{S}_\beta[\Gamma]$, since

$$\mathfrak{S}_\alpha[\Gamma] = \mathfrak{S}_\alpha \Gamma^n = \Gamma^n \mathfrak{S}_\alpha.$$

Thus, part (b) follows from part (a).

(c) In our proof of part (a) above, we showed that $\mathfrak{S}_\alpha w_A \mathfrak{S}_\beta = \mathfrak{S}_\alpha w_B \mathfrak{S}_\beta$ implies $A = B$. A similar argument shows that $P_\alpha w_A P_\beta = P_\alpha w_B P_\beta$ implies $A = B$: for g in GL_n, the rank $r_{ij}(g)$ of the matrix obtained by restricting g to rows $I_i \sqcup I_{i+1} \sqcup \cdots \sqcup I_\ell$ and columns $J_1 \sqcup J_2 \sqcup \cdots \sqcup J_j$ is constant on double cosets $P_\alpha g P_\beta$, and for a permutation matrix w one can recover $a_{i,j}(w)$ from the formula

$$a_{i,j}(w) = r_{i,j}(w) - r_{i,j-1}(w) - r_{i+1,j}(w) + r_{i+1,j-1}(w).$$

Thus it only remains to show that every double coset $P_\alpha g P_\beta$ contains some w_A. Since $\mathfrak{S}_\alpha < P_\alpha$, and we have seen already that every double coset $\mathfrak{S}_\alpha w \mathfrak{S}_\beta$ contains some w_A, it suffices to show that every double coset $P_\alpha g P_\beta$ contains some permutation w. However, we claim that this is already true for the smaller double cosets BgB where $B = P_{1^n}$ is the *Borel subgroup* of upper triangular invertible matrices, that is, one has the usual *Bruhat decomposition*

$$GL_n = \bigsqcup_{w \in \mathfrak{S}_n} BwB.$$

To prove this decomposition, we show how to find a permutation w in each double coset BgB. The freedom to alter g within its coset gB allows one to scale columns and add scalar multiples of earlier columns to later columns. We claim that using such column operations, one can always find a representative g' for coset gB in which

- the bottommost nonzero entry of each column is 1 (call this entry a *pivot*),
- the entries to right of each pivot within its row are all 0, and
- there is one pivot in each row and each column, so that their positions are the positions of the 1's in some permutation matrix w.

In fact, we will see below that $BgB = BwB$ in this case. The algorithm which produces g' from g is simple: starting with the leftmost column, find its bottommost nonzero entry, and scale the column to make this entry a 1, creating the pivot in this column. Now use this pivot to clear out all entries in its row to its right, using column operations that subtract multiples of this column from later columns. Having done this, move on to the next column to the right, and repeat, scaling to create a pivot, and using it to eliminate entries to its right.[224]

[224] To see that this works, we need to check three facts:

(a) We will find a nonzero entry in every column during our algorithm.

(b) Our column operations preserve the zeroes lying to the right of already existing pivots.

(c) Every row contains exactly one pivot at the end of the algorithm.

But fact **(a)** simply says that our matrix can never have an all-zero column during the algorithm; this is clear (since the rank of the matrix remains constant during the algorithm and was n at its beginning). Fact **(b)** holds because all our operations either scale columns (which clearly preserves zero entries) or subtract a multiple of the column c containing the current pivot from a later column d (which will preserve every zero lying to the right of an already existing pivot, because any already existing pivot must lie in a column $b < c$ and therefore both columns c and d have zeroes in its row). Fact **(c)** follows from noticing that

For example, the typical matrix g lying in the double coset BwB where

$$w = \begin{pmatrix} 1 & 2 & 3 & | & 4 & 5 & 6 & 7 & | & 8 & 9 \\ \underline{4} & \underline{\underline{8}} & \underline{2} & | & \underline{\underline{5}} & \underline{3} & \underline{\underline{9}} & \underline{1} & | & \underline{\underline{7}} & \underline{\underline{6}} \end{pmatrix}$$

from before is one that can be altered within its coset gB to look like this:

$$g' = \begin{bmatrix} * & * & * & * & * & * & 1 & 0 & 0 \\ * & * & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ * & * & 0 & * & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & * & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & * & 0 & 0 & 0 & * & 0 & * & 1 \\ 0 & * & 0 & 0 & 0 & * & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \in gB.$$

Having found this g' in gB, a similar algorithm using left multiplication by B shows that w lies in $Bg' \subset Bg'B = BgB$. This time no scalings are required to create the pivot entries: starting with the bottom row, one uses its pivot to eliminate all the entries above it in the same column (shown by stars $*$ above) by adding multiples of the bottom row to higher rows. Then do the same using the pivot in the next-to-bottom row, etc. The result is the permutation matrix for w. □

Remark 4.3.8. The Bruhat decomposition $GL_n = \bigsqcup_{w \in \mathfrak{S}_n} BwB$ is related to the so-called *LPU factorization* – one of a myriad of matrix factorizations appearing in linear algebra.[225] It is actually a fairly general phenomenon, and requires neither the finiteness of $\mathbb{F}$, nor the invertibility, nor even the squareness of the matrices (see Exercise 4.3.9(b) for an analogue holding in a more general setup).

Exercise 4.3.9. Let $\mathbb{F}$ be any field.

(a) For any $n \in \mathbb{N}$ and any $A \in GL_n(\mathbb{F})$, prove that there exist a lower-triangular matrix $L \in GL_n(\mathbb{F})$, an upper-triangular matrix $U \in GL_n(\mathbb{F})$ and a permutation matrix $P \in \mathfrak{S}_n \subset GL_n(\mathbb{F})$ (here, we identify permutations with the corresponding permutation matrices) such that $A = LPU$.

(b) Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$. Let $F_{n,m}$ denote the set of all $n \times m$-matrices $B \in \{0,1\}^{n \times m}$ such that each row of B contains at most one 1 and each column of B contains at most one 1. We regard $F_{n,m}$ as a subset of $\mathbb{F}^{n \times m}$ by means of regarding $\{0,1\}$ as a subset of $\mathbb{F}$.

For every $k \in \mathbb{N}$, we let B_k denote the subgroup of $GL_k(\mathbb{F})$ consisting of all upper-triangular matrices.

Prove that

$$\mathbb{F}^{n \times m} = \bigsqcup_{f \in F_{n,m}} B_n f B_m.$$

Corollary 4.3.10. *For each of the three towers of groups G_*, the product and coproduct structures on $A = A(G_*)$ endow it with a bialgebra structure, and hence they form PSH's.*

Proof. The first two towers $G_n = \mathfrak{S}_n$ and $G_n = \mathfrak{S}_n[\Gamma]$ have product, coproduct defined by induction, restriction along embeddings $G_i \times G_j < G_{i+j}$. Hence the desired bialgebra equality (4.3.2) follows from Mackey's Theorem 4.1.7, taking $G = G_n, H = G_{(r_1,r_2)}, K = G_{(c_1,c_2)}, U = U_1 \otimes U_2$ with double coset

there are n pivots altogether at the end of the algorithm, but no row can contain two of them (since the entries to the right of a pivot in its row are 0).

[225] Specifically, an *LPU factorization* of a matrix $A \in GL_n(\mathbb{F})$ (for an arbitrary field $\mathbb{F}$) means a way to write A as a product $A = LPU$ with $L \in GL_n(\mathbb{F})$ being lower-triangular, $U \in GL_n(\mathbb{F})$ being upper-triangular, and $P \in \mathfrak{S}_n \subset GL_n(\mathbb{F})$ being a permutation matrix. Such a factorization always exists (although it is generally not unique). This can be derived from the Bruhat decomposition (see Exercise 4.3.9(a) for a proof). See also [212] for related discussion.

representatives[226]

$$\{g_1, \ldots, g_t\} = \left\{ w_{A^t} \ : \ A \in \mathbb{N}^{2\times 2}, \ A \text{ has row sums } (r_1, r_2) \text{ and column sums } (c_1, c_2) \right\}$$

and checking for a given double coset

$$KgH = (G_{c_1,c_2}) w_{A^t} (G_{r_1,r_2})$$

indexed by a matrix A in $\mathbb{N}^{2\times 2}$ with row sums (r_1, r_2) and column sums (c_1, c_2), that the two subgroups appearing on the left in (4.3.3) are exactly

$$\begin{aligned} H \cap K^{w_{A^t}} &= G_{r_1,r_2} \cap (G_{c_1,c_2})^{w_{A^t}}, \\ {}^{w_{A^t}} H \cap K &= {}^{w_{A^t}} (G_{r_1,r_2}) \cap G_{c_1,c_2}, \end{aligned}$$

respectively. One should also apply (4.1.6) and check that the isomorphism τ_A between the two subgroups in (4.3.3) is the conjugation isomorphism by w_{A^t} (that is, $\tau_A(g) = w_{A^t} g w_{A^t}^{-1}$ for every $g \in H \cap K^{w_{A^t}}$). We leave all of these bookkeeping details to the reader to check. [227]

For the tower with $G_n = GL_n$, there is slightly more work to be done to check the equality (4.3.2). Via Mackey's Theorem 4.1.7 and Proposition 4.3.7(c), the left side is

$$\begin{aligned} & \operatorname{res}^n_{c_1,c_2} \left(\operatorname{ind}^n_{r_1,r_2} (U_1 \otimes U_2) \right) \\ &= \left(\operatorname{Res}^{G_n}_{P_{c_1,c_2}} \operatorname{Ind}^{G_n}_{P_{r_1,r_2}} \operatorname{Infl}^{P_{r_1,r_2}}_{G_{r_1,r_2}} (U_1 \otimes U_2) \right)^{K_{c_1,c_2}} \\ &= \bigoplus_A \left(\operatorname{Ind}^{P_{c_1,c_2}}_{w_{A^t} P_{r_1,r_2} \cap P_{c_1,c_2}} \left(\left(\operatorname{Res}^{P_{r_1,r_2}}_{P_{r_1,r_2} \cap P^{w_{A^t}}_{c_1,c_2}} \operatorname{Infl}^{P_{r_1,r_2}}_{G_{r_1,r_2}} (U_1 \otimes U_2) \right)^{\tau_A^{-1}} \right) \right)^{K_{c_1,c_2}} \end{aligned} \tag{4.3.4}$$

where A runs over the usual 2×2 matrices. The right side is a direct sum over this same set of matrices A:

$$\begin{aligned} & \bigoplus_A \left(\operatorname{ind}^{c_1}_{a_{11},a_{21}} \otimes \operatorname{ind}^{c_2}_{a_{12},a_{22}} \right) \left(\left(\operatorname{res}^{r_1}_{a_{11},a_{12}} U_1 \otimes \operatorname{res}^{r_2}_{a_{21},a_{22}} U_2 \right)^{\tau_A^{-1}} \right) \\ &= \bigoplus_A \left(\operatorname{Ind}^{G_{c_1}}_{P_{a_{11},a_{21}}} \otimes \operatorname{Ind}^{G_{c_2}}_{P_{a_{12},a_{22}}} \right) \circ \left(\operatorname{Infl}^{P_{a_{11},a_{21}}}_{G_{a_{11},a_{21}}} \otimes \operatorname{Infl}^{P_{a_{12},a_{22}}}_{G_{a_{12},a_{22}}} \right) \\ & \qquad \left(\left(\left(\operatorname{Res}^{G_{r_1}}_{P_{a_{11},a_{12}}} U_1 \right)^{K_{a_{11},a_{12}}} \otimes \left(\operatorname{Res}^{G_{r_2}}_{P_{a_{21},a_{22}}} U_2 \right)^{K_{a_{21},a_{22}}} \right)^{\tau_A^{-1}} \right) \\ &= \bigoplus_A \operatorname{Ind}^{G_{c_1,c_2}}_{P_{a_{11},a_{21}} \times P_{a_{12},a_{22}}} \\ & \qquad \operatorname{Infl}^{P_{a_{11},a_{21}} \times P_{a_{12},a_{22}}}_{G_{a_{11},a_{21},a_{12},a_{22}}} \left(\left(\left(\operatorname{Res}^{G_{r_1,r_2}}_{P_{a_{11},a_{12}} \times P_{a_{21},a_{22}}} (U_1 \otimes U_2) \right)^{K_{a_{11},a_{12}} \times K_{a_{21},a_{22}}} \right)^{\tau_A^{-1}} \right) \end{aligned} \tag{4.3.5}$$

(by (4.1.6), (4.1.15) and their obvious analogues for restriction and inflation). Thus it suffices to check for each 2×2 matrix A that any $\mathbb{C}G_{c_1,c_2}$-module of the form $V_1 \otimes V_2$ has the same inner product with the A-summands of (4.3.4) and (4.3.5). Abbreviate $w := w_{A^t}$ and $\tau := \tau_A^{-1}$.

[226] Proposition 4.3.7 gives as a system of double coset representatives for $G_{(c_1,c_2)} \backslash G_n / G_{(r_1,r_2)}$ the elements

$$\begin{aligned} & \left\{ w_A \ : \ A \in \mathbb{N}^{2\times 2}, \ A \text{ has row sums } (c_1, c_2) \text{ and column sums } (r_1, r_2) \right\} \\ &= \left\{ w_{A^t} \ : \ A \in \mathbb{N}^{2\times 2}, \ A \text{ has row sums } (r_1, r_2) \text{ and column sums } (c_1, c_2) \right\} \end{aligned}$$

where A^t denotes the transpose matrix of A.

[227] It helps to recognize w_{A^t} as the permutation written in two-line notation as

$$\begin{pmatrix} 1 & 2 & \ldots & a_{11} & | & a_{11}+1 & a_{11}+2 & \ldots & r_1 & | & r_1+1 & r_1+2 & \ldots & a'_{22} & | & a'_{22}+1 & a'_{22}+2 & \ldots & n \\ 1 & 2 & \ldots & a_{11} & | & c_1+1 & c_1+2 & \ldots & a'_{22} & | & a_{11}+1 & a_{11}+2 & \ldots & c_1 & | & a'_{22}+1 & a'_{22}+2 & \ldots & n \end{pmatrix},$$

where $a'_{22} = r_1 + a_{21} = c_1 + a_{12} = n - a_{22}$. In matrix form, w_{A^t} is the block matrix $\begin{bmatrix} I_{a_{11}} & 0 & 0 & 0 \\ 0 & 0 & I_{a_{21}} & 0 \\ 0 & I_{a_{12}} & 0 & 0 \\ 0 & 0 & 0 & I_{a_{22}} \end{bmatrix}$.

Notice that ${}^w P_{r_1,r_2}$ is the group of all matrices having the block form

$$\begin{bmatrix} g_{11} & h & i & j \\ 0 & g_{21} & 0 & k \\ d & e & g_{12} & \ell \\ 0 & f & 0 & g_{22} \end{bmatrix} \tag{4.3.6}$$

in which the diagonal blocks g_{ij} for $i, j = 1, 2$ are invertible of size $a_{ij} \times a_{ij}$, while the blocks $h, i, j, k, \ell, d, e, f$ are all arbitrary matrices[228] of the appropriate (rectangular) block sizes. Hence, ${}^w P_{r_1,r_2} \cap P_{c_1,c_2}$ is the group of all matrices having the block form

$$\begin{bmatrix} g_{11} & h & i & j \\ 0 & g_{21} & 0 & k \\ 0 & 0 & g_{12} & \ell \\ 0 & 0 & 0 & g_{22} \end{bmatrix} \tag{4.3.7}$$

in which the diagonal blocks g_{ij} for $i, j = 1, 2$ are invertible of size $a_{ij} \times a_{ij}$, while the blocks h, i, j, k, ℓ are all arbitrary matrices of the appropriate (rectangular) block sizes; then ${}^w P_{r_1,r_2} \cap G_{c_1,c_2}$ is the subgroup where the blocks i, j, k all vanish. The canonical projection ${}^w P_{r_1,r_2} \cap P_{c_1,c_2} \to {}^w P_{r_1,r_2} \cap G_{c_1,c_2}$ (obtained by restricting the projection $P_{c_1,c_2} \to G_{c_1,c_2}$) has kernel ${}^w P_{r_1,r_2} \cap P_{c_1,c_2} \cap K_{c_1,c_2}$. Consequently,

$$({}^w P_{r_1,r_2} \cap P_{c_1,c_2}) \,/\, ({}^w P_{r_1,r_2} \cap P_{c_1,c_2} \cap K_{c_1,c_2}) = {}^w P_{r_1,r_2} \cap G_{c_1,c_2}. \tag{4.3.8}$$

Similarly,

$$\left(P_{r_1,r_2} \cap P^w_{c_1,c_2}\right) / \left(P_{r_1,r_2} \cap P^w_{c_1,c_2} \cap K_{r_1,r_2}\right) = G_{r_1,r_2} \cap P^w_{c_1,c_2}. \tag{4.3.9}$$

Computing first the inner product of $V_1 \otimes V_2$ with the A-summand of (4.3.4), and using adjointness properties, one gets

$$\begin{aligned} &\Big(\left(\operatorname{Res}^{P_{r_1,r_2}}_{P_{r_1,r_2} \cap P^w_{c_1,c_2}} \operatorname{Infl}^{P_{r_1,r_2}}_{G_{r_1,r_2}} (U_1 \otimes U_2) \right)^{\tau}, \\ &\qquad \operatorname{Res}^{P_{c_1,c_2}}_{{}^w P_{r_1,r_2} \cap P_{c_1,c_2}} \operatorname{Infl}^{P_{c_1,c_2}}_{G_{c_1,c_2}} (V_1 \otimes V_2) \Big)_{{}^w P_{r_1,r_2} \cap P_{c_1,c_2}} \\ &\overset{(4.1.10)}{=} \Big(\left(\operatorname{Infl}^{P_{r_1,r_2} \cap P^w_{c_1,c_2}}_{G_{r_1,r_2} \cap P^w_{c_1,c_2}} \operatorname{Res}^{G_{r_1,r_2}}_{G_{r_1,r_2} \cap P^w_{c_1,c_2}} (U_1 \otimes U_2) \right)^{\tau}, \\ &\qquad \operatorname{Infl}^{{}^w P_{r_1,r_2} \cap P_{c_1,c_2}}_{{}^w P_{r_1,r_2} \cap G_{c_1,c_2}} \operatorname{Res}^{G_{c_1,c_2}}_{{}^w P_{r_1,r_2} \cap G_{c_1,c_2}} (V_1 \otimes V_2) \Big)_{{}^w P_{r_1,r_2} \cap P_{c_1,c_2}} \end{aligned}$$

(by (4.3.9) and (4.3.8)). One can compute this inner product by first recalling that ${}^w P_{r_1,r_2} \cap P_{c_1,c_2}$ is the group of matrices having the block form (4.3.7) in which the diagonal blocks g_{ij} for $i, j = 1, 2$ are invertible of size $a_{ij} \times a_{ij}$, while the blocks h, i, j, k, ℓ are all arbitrary matrices of the appropriate (rectangular) block sizes; then ${}^w P_{r_1,r_2} \cap G_{c_1,c_2}$ is the subgroup where the blocks i, j, k all vanish. The inner product above then becomes

$$\begin{aligned} &\frac{1}{\left| {}^w P_{r_1,r_2} \cap P_{c_1,c_2} \right|} \sum_{\substack{(g_{ij}) \\ (h,i,j,k,\ell)}} \chi_{U_1} \begin{pmatrix} g_{11} & i \\ 0 & g_{12} \end{pmatrix} \chi_{U_2} \begin{pmatrix} g_{21} & k \\ 0 & g_{22} \end{pmatrix} \\ &\qquad\qquad\qquad \overline{\chi}_{V_1} \begin{pmatrix} g_{11} & h \\ 0 & g_{21} \end{pmatrix} \overline{\chi}_{V_2} \begin{pmatrix} g_{12} & \ell \\ 0 & g_{22} \end{pmatrix}. \end{aligned} \tag{4.3.10}$$

[228] The blocks i and j have nothing to do with the indices i, j in g_{ij}.

If one instead computes the inner product of $V_1 \otimes V_2$ with the A-summand of (4.3.5), using adjointness properties and (4.1.13) one gets

$$\begin{aligned} &\Bigg(\left(\left(\operatorname{Res}^{G_{r_1,r_2}}_{P_{a_{11},a_{12}} \times P_{a_{21},a_{22}}} (U_1 \otimes U_2) \right)^{K_{a_{11},a_{12}} \times K_{a_{21},a_{22}}} \right)^{\tau}, \\ &\qquad \left(\operatorname{Res}^{G_{c_1,c_2}}_{P_{a_{11},a_{21}} \times P_{a_{12},a_{22}}} (V_1 \otimes V_2) \right)^{K_{a_{11},a_{21}} \times K_{a_{12},a_{22}}} \Bigg)_{G_{a_{11},a_{21},a_{12},a_{22}}} \\ &= \frac{1}{|G_{a_{11},a_{21},a_{12},a_{22}}|} \sum_{(g_{ij})} \frac{1}{|K_{a_{11},a_{12}} \times K_{a_{21},a_{22}}|} \sum_{(i,k)} \chi_{U_1} \begin{pmatrix} g_{11} & i \\ 0 & g_{12} \end{pmatrix} \chi_{U_2} \begin{pmatrix} g_{21} & k \\ 0 & g_{22} \end{pmatrix} \\ &\qquad \frac{1}{|K_{a_{11},a_{21}} \times K_{a_{12},a_{22}}|} \sum_{(h,\ell)} \overline{\chi}_{V_1} \begin{pmatrix} g_{11} & h \\ 0 & g_{21} \end{pmatrix} \overline{\chi}_{V_2} \begin{pmatrix} g_{12} & \ell \\ 0 & g_{22} \end{pmatrix}. \end{aligned}$$

But this right hand side can be seen to equal (4.3.10), after one notes that

$$|{}^w P_{r_1,r_2} \cap P_{c_1,c_2}| = |G_{a_{11},a_{21},a_{12},a_{22}}| \cdot |K_{a_{11},a_{12}} \times K_{a_{21},a_{22}}| \cdot |K_{a_{11},a_{21}} \times K_{a_{12},a_{22}}| \cdot \#\{j \in \mathbb{F}_q^{a_{11} \times a_{22}}\}$$

and that the summands in (4.3.10) are independent of the matrix j in the summation. □

We can also define a $\mathbb{C}$-vector space $A_{\mathbb{C}}$ as the direct sum $\bigoplus_{n \geq 0} R_{\mathbb{C}}(G_n)$. In the same way as we have made $A = \bigoplus_{n \geq 0} R(G_n)$ into a $\mathbb{Z}$-bialgebra, we can turn $A_{\mathbb{C}} = \bigoplus_{n \geq 0} R_{\mathbb{C}}(G_n)$ into a $\mathbb{C}$-bialgebra[229]. There is a $\mathbb{C}$-bilinear form $(\cdot,\cdot)_{A_{\mathbb{C}}}$ on $A_{\mathbb{C}}$ which can be defined either as the $\mathbb{C}$-bilinear extension of the $\mathbb{Z}$-bilinear form $(\cdot,\cdot)_A : A \times A \to \mathbb{Z}$ to $A_{\mathbb{C}}$, or (equivalently) as the $\mathbb{C}$-bilinear form on $A_{\mathbb{C}}$ which restricts to $\langle \cdot,\cdot \rangle_{\mathfrak{S}_n}$ on every homogeneous component $R_{\mathbb{C}}(G_n)$ and makes different homogeneous components mutually orthogonal. The obvious embedding of A into the $\mathbb{C}$-bialgebra $A_{\mathbb{C}}$ (obtained from the embeddings $R(G_n) \to R_{\mathbb{C}}(G_n)$ for all n) respects the bialgebra operations[230], and the $\mathbb{C}$-bialgebra $A_{\mathbb{C}}$ can be identified with $A \otimes_{\mathbb{Z}} \mathbb{C}$ (the result of extending scalars to $\mathbb{C}$ in A), because every finite group G satisfies $R_{\mathbb{C}}(G) \cong R(G) \otimes_{\mathbb{Z}} \mathbb{C}$. The embedding of A into $A_{\mathbb{C}}$ also respects the bilinear forms.

Exercise 4.3.11. Let G_* be one of the three towers.

For every almost-composition $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_\ell)$ of $n \in \mathbb{N}$, let us define a map $\operatorname{ind}^n_\alpha$ which takes $\mathbb{C}G_\alpha$-modules to $\mathbb{C}G_n$-modules as follows: If $G_* = \mathfrak{S}_*$ or $G_* = \mathfrak{S}_*[\Gamma]$, we set

$$\operatorname{ind}^n_\alpha := \operatorname{Ind}^{G_n}_{G_\alpha}.$$

If $G_* = GL_*$, then we set

$$\operatorname{ind}^n_\alpha := \operatorname{Ind}^{G_n}_{P_\alpha} \operatorname{Infl}^{P_\alpha}_{G_\alpha}.$$

(Note that $\operatorname{ind}^n_\alpha = \operatorname{ind}^n_{i,j}$ if α has the form (i,j).)

Similarly, for every almost-composition $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_\ell)$ of $n \in \mathbb{N}$, let us define a map $\operatorname{res}^n_\alpha$ which takes $\mathbb{C}G_n$-modules to $\mathbb{C}G_\alpha$-modules as follows: If $G_* = \mathfrak{S}_*$ or $G_* = \mathfrak{S}_*[\Gamma]$, we set

$$\operatorname{res}^n_\alpha := \operatorname{Res}^{G_n}_{G_\alpha}.$$

If $G_* = GL_*$, then we set

$$\operatorname{res}^n_\alpha := \left(\operatorname{Res}^{G_n}_{P_\alpha} (-) \right)^{K_\alpha}.$$

(Note that $\operatorname{res}^n_\alpha = \operatorname{res}^n_{i,j}$ if α has the form (i,j).)

[229] The definitions of m and Δ for this $\mathbb{C}$-bialgebra look the same as for A: For instance, m is still defined to be $\operatorname{ind}^{i+j}_{i,j}$ on $(A_{\mathbb{C}})_i \otimes (A_{\mathbb{C}})_j$, where $\operatorname{ind}^{i+j}_{i,j}$ is defined by the same formulas as in Definition 4.2.1. However, the operators of induction, restriction, inflation and K-fixed space construction appearing in these formulas now act on class functions as opposed to modules.

The fact that these maps m and Δ satisfy the axioms of a $\mathbb{C}$-bialgebra is easy to check: they are merely the $\mathbb{C}$-linear extensions of the maps m and Δ of the $\mathbb{Z}$-bialgebra A (this is because, for instance, induction of class functions and induction of modules are related by the identity (4.1.5)), and thus satisfy the same axioms as the latter.

[230] This is because, for example, induction of class functions harmonizes with induction of modules (i.e., the equality (4.1.5) holds).

(a) If $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_\ell)$ is an almost-composition of an integer $n \in \mathbb{N}$ satisfying $\ell \geq 1$, and if V_i is a $\mathbb{C}G_{\alpha_i}$-module for every $i \in \{1, 2, \ldots, \ell\}$, then show that

$$\begin{aligned}
&\operatorname{ind}^n_{\alpha_1+\alpha_2+\cdots+\alpha_{\ell-1},\alpha_\ell} \left(\operatorname{ind}^{\alpha_1+\alpha_2+\cdots+\alpha_{\ell-1}}_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})} \left(V_1 \otimes V_2 \otimes \cdots \otimes V_{\ell-1}\right) \otimes V_\ell \right) \\
&\cong \operatorname{ind}^n_\alpha \left(V_1 \otimes V_2 \otimes \cdots \otimes V_\ell\right) \\
&\cong \operatorname{ind}^n_{\alpha_1,\alpha_2+\alpha_3+\cdots+\alpha_\ell} \left(V_1 \otimes \operatorname{ind}^{\alpha_2+\alpha_3+\cdots+\alpha_\ell}_{(\alpha_2,\alpha_3,\ldots,\alpha_\ell)} \left(V_2 \otimes V_3 \otimes \cdots \otimes V_\ell\right) \right).
\end{aligned}$$

(b) Solve Exercise 4.2.3 again using Exercise 4.3.11(a).

(c) We proved above that the map $m : A \otimes A \to A$ (where $A = A(G_*)$) is associative, by using the adjointness of m and Δ. Give a new proof of this fact, which makes no use of Δ.

(d) If $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_\ell)$ is an almost-composition of an $n \in \mathbb{N}$, and if $\chi_i \in R(G_{\alpha_i})$ for every $i \in \{1, 2, \ldots, \ell\}$, then show that

$$\chi_1 \chi_2 \cdots \chi_\ell = \operatorname{ind}^n_\alpha \left(\chi_1 \otimes \chi_2 \otimes \cdots \otimes \chi_\ell\right)$$

in $A = A(G_*)$.

(e) If $n \in \mathbb{N}$, $\ell \in \mathbb{N}$ and $\chi \in R(G_n)$, then show that

$$\Delta^{(\ell-1)} \chi = \sum \operatorname{res}^n_\alpha \chi$$

in $A^{\otimes \ell}$, where $A = A(G_*)$. Here, the sum on the right hand side runs over all almost-compositions α of n having length ℓ.

4.4. **Symmetric groups.** Finally, some payoff. Consider the tower of symmetric groups $G_n = \mathfrak{S}_n$, and $A = A(G_*) =: A(\mathfrak{S})$. Denote by $\underline{1}_{\mathfrak{S}_n}, \operatorname{sgn}_{\mathfrak{S}_n}$ the *trivial* and *sign* characters on $\mathfrak{S}_n$. For a partition λ of n, denote by $\underline{1}_{\mathfrak{S}_\lambda}, \operatorname{sgn}_{\mathfrak{S}_\lambda}$ the trivial and sign characters restricted to the Young subgroup $\mathfrak{S}_\lambda = \mathfrak{S}_{\lambda_1} \times \mathfrak{S}_{\lambda_2} \times \cdots$, and denote by $\underline{1}_\lambda$ the class function which is the characteristic function for the $\mathfrak{S}_n$-conjugacy class of permutations of cycle type λ.

Theorem 4.4.1. (a) *Irreducible complex characters $\{\chi^\lambda\}$ of $\mathfrak{S}_n$ are indexed by partitions λ in* Par_n*, and one has a PSH-isomorphism, the* Frobenius characteristic map[231]*,*

$$A = A(\mathfrak{S}) \xrightarrow{\operatorname{ch}} \Lambda$$

that for $n \geq 0$ and $\lambda \in \operatorname{Par}_n$ sends

$$\begin{array}{rcl}
\underline{1}_{\mathfrak{S}_n} & \longmapsto & h_n, \\
\operatorname{sgn}_{\mathfrak{S}_n} & \longmapsto & e_n, \\
\chi^\lambda & \longmapsto & s_\lambda, \\
\operatorname{Ind}^{\mathfrak{S}_n}_{\mathfrak{S}_\lambda} \underline{1}_{\mathfrak{S}_\lambda} & \longmapsto & h_\lambda, \\
\operatorname{Ind}^{\mathfrak{S}_n}_{\mathfrak{S}_\lambda} \operatorname{sgn}_{\mathfrak{S}_\lambda} & \longmapsto & e_\lambda, \\
\underline{1}_\lambda & \longmapsto & \frac{p_\lambda}{z_\lambda}
\end{array}$$

(where ch *is extended to a $\mathbb{C}$-linear map $A_\mathbb{C} \to \Lambda_\mathbb{C}$), and for $n \geq 1$ sends*

$$\underline{1}_{(n)} \longmapsto \frac{p_n}{n}.$$

Here, z_λ is defined as in Proposition 2.5.15.

(b) *For each $n \geq 0$, the involution on class functions $f : \mathfrak{S}_n \to \mathbb{C}$ sending $f \longmapsto \operatorname{sgn}_{\mathfrak{S}_n} * f$ where*

$$(\operatorname{sgn}_{\mathfrak{S}_n} * f)(g) := \operatorname{sgn}(g) f(g)$$

preserves the $\mathbb{Z}$-sublattice $R(\mathfrak{S}_n)$ of genuine characters. The direct sum of these involutions induces an involution on $A = A(\mathfrak{S}) = \bigoplus_{n \geq 0} R(\mathfrak{S}_n)$ that corresponds under ch *to the involution ω on Λ.*

[231] It is unrelated to the Frobenius endomorphisms from Exercise 2.9.9.

Proof. (a) Corollary 4.3.10 implies that the set $\Sigma = \bigsqcup_{n\geq 0} \operatorname{Irr}(\mathfrak{S}_n)$ gives a PSH-basis for A. Since a character χ of $\mathfrak{S}_n$ has

$$\Delta(\chi) = \bigoplus_{i+j=n} \operatorname{Res}^{\mathfrak{S}_n}_{\mathfrak{S}_i \times \mathfrak{S}_j} \chi, \tag{4.4.1}$$

such an element $\chi \in \Sigma \cap A_n$ is never primitive for $n \geq 2$. Hence the unique irreducible character $\rho = \underline{1}_{\mathfrak{S}_1}$ of $\mathfrak{S}_1$ is the only element of $\mathcal{C} = \Sigma \cap \mathfrak{p}$.

Thus Theorem 3.3.3(g) tells us that there are two PSH-isomorphisms $A \to \Lambda$, each of which sends Σ to the PSH-basis of Schur functions $\{s_\lambda\}$ for Λ. It also tells us that we can pin down one of the two isomorphisms to call ch, by insisting that it map the two characters $\underline{1}_{\mathfrak{S}_2}, \operatorname{sgn}_{\mathfrak{S}_2}$ in $\operatorname{Irr}(\mathfrak{S}_2)$ to h_2, e_2 (and not e_2, h_2).

Bearing in mind the coproduct formula (4.4.1), and the fact that $\underline{1}_{\mathfrak{S}_n}, \operatorname{sgn}_{\mathfrak{S}_n}$ restrict, respectively, to trivial and sign characters of $\mathfrak{S}_i \times \mathfrak{S}_j$ for $i + j = n$, one finds that for $n \geq 2$ one has $\operatorname{sgn}^{\perp}_{\mathfrak{S}_2}$ annihilating $\underline{1}_{\mathfrak{S}_n}$, and $\underline{1}^{\perp}_{\mathfrak{S}_2}$ annihilating $\operatorname{sgn}_{\mathfrak{S}_n}$. Therefore Theorem 3.3.1(b) (applied to Λ) implies $\underline{1}_{\mathfrak{S}_n}, \operatorname{sgn}_{\mathfrak{S}_n}$ are sent under ch to h_n, e_n. Then the fact that $\operatorname{Ind}^{\mathfrak{S}_n}_{\mathfrak{S}_\lambda} \underline{1}_{\mathfrak{S}_\lambda}, \operatorname{Ind}^{\mathfrak{S}_n}_{\mathfrak{S}_\lambda} \operatorname{sgn}_{\mathfrak{S}_\lambda}$ are sent to h_λ, e_λ follows via induction products.

Recall that the $\mathbb{C}$-vector space $A_{\mathbb{C}} = \bigoplus_{n\geq 0} R_{\mathbb{C}}(\mathfrak{S}_n)$ is a $\mathbb{C}$-bialgebra, and can be identified with $A \otimes_{\mathbb{Z}} \mathbb{C}$. The multiplication and the comultiplication of $A_{\mathbb{C}}$ are $\mathbb{C}$-linear extensions of those of A, and are still given by the same formulas $m = \operatorname{ind}^{i+j}_{i,j}$ and $\Delta = \bigoplus_{i+j=n} \operatorname{res}^{i+j}_{i,j}$ as those of A (but now, induction and restriction are defined for class functions, not just for representations). The $\mathbb{C}$-bilinear form $(\cdot,\cdot)_{A_{\mathbb{C}}}$ on $A_{\mathbb{C}}$ extends both the $\mathbb{Z}$-bilinear form $(\cdot,\cdot)_A$ on A and the $\mathbb{C}$-bilinear forms $\langle\cdot,\cdot\rangle_{\mathfrak{S}_n}$ on all $R_{\mathbb{C}}(\mathfrak{S}_n)$.

For the assertion about $\underline{1}_{(n)}$, note that it is primitive in $A_{\mathbb{C}}$ for $n \geq 1$, because as a class function, the indicator function of n-cycles vanishes upon restriction to $\mathfrak{S}_i \times \mathfrak{S}_j$ for $i + j = n$ if both $i, j \geq 1$; these subgroups contain no n-cycles. Hence Corollary 3.1.8 implies that $\operatorname{ch}(\underline{1}_{(n)})$ is a scalar multiple of p_n. To pin down the scalar, note $p_n = m_{(n)}$ so $(h_n, p_n)_\Lambda = (h_n, m_n)_\Lambda = 1$, while $\operatorname{ch}^{-1}(h_n) = \underline{1}_{\mathfrak{S}_n}$ has

$$(\underline{1}_{\mathfrak{S}_n}, \underline{1}_{(n)}) = \frac{1}{n!} \cdot (n-1)! = \frac{1}{n}.$$

[232] Thus $\operatorname{ch}(\underline{1}_{(n)}) = \frac{p_n}{n}$. The fact that $\operatorname{ch}(\underline{1}_\lambda) = \frac{p_\lambda}{z_\lambda}$ then follows via induction product calculations[233]. Part (b) follows from Exercise 4.4.4 below. □

Remark 4.4.2. The paper of Liulevicius [133] gives a very elegant alternate approach to the Frobenius map as a Hopf isomorphism $A(\mathfrak{S}) \xrightarrow{\operatorname{ch}} \Lambda$, inspired by equivariant K-theory and vector bundles over spaces which are finite sets of points!

Exercise 4.4.3. If P is a subset of a group G, we denote by $\underline{1}_P$ the map $G \to \mathbb{C}$ which sends every element of P to 1 and all remaining elements of G to 0. [234] For any finite group G and any $h \in G$, we introduce the following notations:

- Let $Z_G(h)$ denote the centralizer of h in G.
- Let $\operatorname{Conj}_G(h)$ denote the conjugacy class of h in G.
- Define a map $\alpha_{G,h} : G \to \mathbb{C}$ by $\alpha_{G,h} = |Z_G(h)| \underline{1}_{\operatorname{Conj}_G(h)}$. This map $\alpha_{G,h}$ is a class function[235].

(a) Prove that $\alpha_{G,h}(g) = \sum_{k\in G} \left[khk^{-1} = g\right]$ for every finite group G and any $h \in G$ and $g \in G$. Here, we are using the Iverson bracket notation (that is, for any statement $\mathcal{A}$, we define $[\mathcal{A}]$ to be the integer 1 if $\mathcal{A}$ is true, and 0 otherwise).

(b) Prove that if H is a subgroup of a finite group G, and if $h \in H$, then $\operatorname{Ind}^G_H \alpha_{H,h} = \alpha_{G,h}$.

(c) Prove that if G_1 and G_2 are finite groups, and if $h_1 \in G_1$ and $h_2 \in G_2$, then the canonical isomorphism $R_{\mathbb{C}}(G_1) \otimes R_{\mathbb{C}}(G_2) \to R_{\mathbb{C}}(G_1 \times G_2)$ sends $\alpha_{G_1,h_1} \otimes \alpha_{G_2,h_2}$ to $\alpha_{G_1 \times G_2,(h_1,h_2)}$.

[232] The first equality sign in this computation uses the fact that the number of all n-cycles in $\mathfrak{S}_n$ is $(n-1)!$. This is because any n-cycle in $\mathfrak{S}_n$ can be uniquely written in the form $(i_1, i_2, \ldots, i_{n-1}, n)$ (in cycle notation) with $(i_1, i_2, \ldots, i_{n-1})$ being a permutation in $\mathfrak{S}_{n-1}$ (written in one-line notation).

[233] For instance, one can use (4.1.3) to show that $z_\lambda \underline{1}_\lambda = \lambda_1 \lambda_2 \cdots \lambda_\ell \cdot \underline{1}_{(\lambda_1)} \underline{1}_{(\lambda_2)} \cdots \underline{1}_{(\lambda_\ell)}$ if $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_\ell)$ with $\ell = \ell(\lambda)$. See Exercise 4.4.3(d) for the details.

[234] This is not in conflict with the notation $\underline{1}_G$ for the trivial character of G, since $\underline{1}_P = \underline{1}_G$ for $P = G$. Note that $\underline{1}_P$ is a class function when P is a union of conjugacy classes of G.

[235] In fact, $\underline{1}_{\operatorname{Conj}_G(h)}$ is a class function (since $\operatorname{Conj}_G(h)$ is a conjugacy class), and so $\alpha_{G,h}$ (being the scalar multiple $|Z_G(h)| \underline{1}_{\operatorname{Conj}_G(h)}$ of $\underline{1}_{\operatorname{Conj}_G(h)}$) must also be a class function.

(d) Fill in the details of the proof of $\operatorname{ch}(\underline{1}_\lambda) = \frac{p_\lambda}{z_\lambda}$ in the proof of Theorem 4.4.1.

(e) Obtain an alternative proof of Remark 2.5.16.

(f) If G and H are two finite groups, and if $\rho : H \to G$ is a group homomorphism, then prove that $\operatorname{Ind}_\rho \alpha_{H,h} = \alpha_{G,\rho(h)}$ for every $h \in H$, where $\operatorname{Ind}_\rho \alpha_{H,h}$ is defined as in Exercise 4.1.14.

Exercise 4.4.4. If G is a group and U_1 and U_2 are two $\mathbb{C}G$-modules, then the tensor product $U_1 \otimes U_2$ is a $\mathbb{C}[G \times G]$-module, which can be made into a $\mathbb{C}G$-module by letting $g \in G$ act as $(g, g) \in G \times G$. This $\mathbb{C}G$-module $U_1 \otimes U_2$ is called the *inner tensor product*[236] of U_1 and U_2, and is a restriction of the outer tensor product $U_1 \otimes U_2$ using the inclusion map $G \to G \times G$, $g \mapsto (g, g)$.

Let $n \geq 0$, and let $\operatorname{sgn}_{\mathfrak{S}_n}$ be the 1-dimensional $\mathbb{C}\mathfrak{S}_n$-module $\mathbb{C}$ on which every $g \in \mathfrak{S}_n$ acts as multiplication by $\operatorname{sgn}(g)$. If V is a $\mathbb{C}\mathfrak{S}_n$-module, show that the involution on $A(\mathfrak{S}) = \bigoplus_{n \geq 0} R(\mathfrak{S}_n)$ defined in Theorem 4.4.1(b) sends $\chi_V \mapsto \chi_{\operatorname{sgn}_{\mathfrak{S}_n} \otimes V}$ where $\operatorname{sgn}_{\mathfrak{S}_n} \otimes V$ is the inner tensor product of $\operatorname{sgn}_{\mathfrak{S}_n}$ and V. Use this to show that this involution is a nontrivial PSH-automorphism of $A(\mathfrak{S})$, and deduce Theorem 4.4.1(b).

Exercise 4.4.5. Let $n \in \mathbb{N}$. For every permutation $\sigma \in \mathfrak{S}_n$, we let $\operatorname{type} \sigma$ denote the cycle type of σ. Extend $\operatorname{ch} : A = A(\mathfrak{S}) \to \Lambda$ to a $\mathbb{C}$-linear map $A_\mathbb{C} \to \Lambda_\mathbb{C}$. We shall call the latter map ch, too.

(a) Prove that every class function $f \in R_\mathbb{C}(\mathfrak{S}_n)$ satisfies

$$\operatorname{ch}(f) = \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_n} f(\sigma)\, p_{\operatorname{type} \sigma}.$$

(b) Let H be a subgroup of $\mathfrak{S}_n$. Prove that every class function $f \in R_\mathbb{C}(H)$ satisfies

$$\operatorname{ch}\left(\operatorname{Ind}_H^{\mathfrak{S}_n} f\right) = \frac{1}{|H|} \sum_{h \in H} f(h)\, p_{\operatorname{type} h}.$$

Exercise 4.4.6. (a) Show that for every $n \geq 0$, every $g \in \mathfrak{S}_n$ and every finite-dimensional $\mathbb{C}\mathfrak{S}_n$-module V, we have $\chi_V(g) \in \mathbb{Z}$.

(b) Show that for every $n \geq 0$ and every finite-dimensional $\mathbb{C}\mathfrak{S}_n$-module V, there exists a $\mathbb{Q}\mathfrak{S}_n$-module W such that $V \cong \mathbb{C} \otimes_\mathbb{Q} W$. (In the representation theorists' parlance, this says that all representations of $\mathfrak{S}_n$ are *defined over* $\mathbb{Q}$. This part of the exercise requires some familiarity with representation theory.)

Remark 4.4.7. Parts (a) and (b) of Exercise 4.4.6 both follow from an even stronger result: For every $n \geq 0$ and every finite-dimensional $\mathbb{C}\mathfrak{S}_n$-module V, there exists a $\mathbb{Z}\mathfrak{S}_n$-module W which is finitely generated and free as a $\mathbb{Z}$-module and satisfies $V \cong \mathbb{C} \otimes_\mathbb{Z} W$ as $\mathbb{C}\mathfrak{S}_n$-modules. This follows from the combinatorial approach to the representation theory of $\mathfrak{S}_n$, in which the irreducible representations of $\mathbb{C}\mathfrak{S}_n$ (the *Specht modules*) are constructed using Young tableaux and tabloids. See the literature on the symmetric group, e.g., [186], [73, §7], [223] or [115, Section 2.2] for this approach.

The connection between Λ and $A(\mathfrak{S})$ as established in Theorem 4.4.1 benefits both the study of Λ and that of $A(\mathfrak{S})$. The following two exercises show some applications to Λ:

Exercise 4.4.8. If G is a group and U_1 and U_2 are two $\mathbb{C}G$-modules, then let $U_1 \boxtimes U_2$ denote the inner tensor product of U_1 and U_2 (as defined in Exercise 4.4.4). Consider also the binary operation $*$ on $\Lambda_\mathbb{Q}$ defined in Exercise 2.9.4(h).

(a) Show that $\operatorname{ch}(\chi_{U_1 \boxtimes U_2}) = \operatorname{ch}(\chi_{U_1}) * \operatorname{ch}(\chi_{U_2})$ for any $n \in \mathbb{N}$ and any two $\mathbb{C}\mathfrak{S}_n$-modules U_1 and U_2.

(b) Use this to obtain a new solution for Exercise 2.9.4(h).

(c) Show that $s_\mu * s_\nu \in \sum_{\lambda \in \operatorname{Par}} \mathbb{N} s_\lambda$ for any two partitions μ and ν.

[**Hint:** For any group G, introduce a binary operation $*$ on $R_\mathbb{C}(G)$ which satisfies $\chi_{U_1 \boxtimes U_2} = \chi_{U_1} * \chi_{U_2}$ for any two $\mathbb{C}G$-modules U_1 and U_2.]

Exercise 4.4.9. Define a $\mathbb{Q}$-bilinear map $\boxdot : \Lambda_\mathbb{Q} \times \Lambda_\mathbb{Q} \to \Lambda_\mathbb{Q}$, which will be written in infix notation (that is, we will write $a \boxdot b$ instead of $\boxdot(a, b)$), by setting

$$p_\lambda \boxdot p_\mu = \prod_{i=1}^{\ell(\lambda)} \prod_{j=1}^{\ell(\mu)} p_{\operatorname{lcm}(\lambda_i, \mu_j)}^{\gcd(\lambda_i, \mu_j)} \qquad \text{for any partitions } \lambda \text{ and } \mu.$$

[236] Do not confuse this with the inner product of characters.

[237]

(a) Show that $\Lambda_{\mathbb{Q}}$, equipped with the binary operation $\boxdot$, becomes a commutative $\mathbb{Q}$-algebra with unity p_1.

(b) For every $r \in \mathbb{Z}$, define the $\mathbb{Q}$-algebra homomorphism $\epsilon_r : \Lambda_{\mathbb{Q}} \to \mathbb{Q}$ as in Exercise 2.9.4(c). Show that $1 \boxdot f = \epsilon_1(f)\, 1$ for every $f \in \Lambda_{\mathbb{Q}}$ (where 1 denotes the unity of Λ).

(c) Show that $s_\mu \boxdot s_\nu \in \sum_{\lambda \in \operatorname{Par}} \mathbb{N} s_\lambda$ for any two partitions μ and ν.

(d) Show that $f \boxdot g \in \Lambda$ for any $f \in \Lambda$ and $g \in \Lambda$.

[**Hint:** For every set X, let $\mathfrak{S}_X$ denote the group of all permutations of X. For two sets X and Y, there is a canonical group homomorphism $\mathfrak{S}_X \times \mathfrak{S}_Y \to \mathfrak{S}_{X \times Y}$, which is injective if X and Y are nonempty. For positive integers n and m, this yields an embedding $\mathfrak{S}_n \times \mathfrak{S}_m \to \mathfrak{S}_{\{1,2,\ldots,n\}\times\{1,2,\ldots,m\}}$, which, once $\mathfrak{S}_{\{1,2,\ldots,n\}\times\{1,2,\ldots,m\}}$ is identified with $\mathfrak{S}_{nm}$ (using an arbitrary but fixed bijection $\{1,2,\ldots,n\} \times \{1,2,\ldots,m\} \to \{1,2,\ldots,nm\}$), can be regarded as an embedding $\mathfrak{S}_n \times \mathfrak{S}_m \to \mathfrak{S}_{nm}$ and thus allows defining a $\mathbb{C}\mathfrak{S}_{nm}$-module $\operatorname{Ind}_{\mathfrak{S}_n \times \mathfrak{S}_m}^{\mathfrak{S}_{nm}} (U \otimes V)$ for any $\mathbb{C}\mathfrak{S}_n$-module U and any $\mathbb{C}\mathfrak{S}_m$-module V. This gives a binary operation on $A(\mathfrak{S})$. Show that this operation corresponds to $\boxdot$ under the PSH-isomorphism $\operatorname{ch} : A(\mathfrak{S}) \to \Lambda$.]

Remark 4.4.10. The statements (and the idea of the solution) of Exercise 4.4.9 are due to Manuel Maia and Miguel Méndez (see [144] and, more explicitly, [155]), who call the operation $\boxdot$ the *arithmetic product*. Li [131, Thm. 3.5] denotes it by $\boxtimes$ and relates it to the enumeration of unlabelled graphs.

4.5. **Wreath products.** Next consider the tower of groups $G_n = \mathfrak{S}_n[\Gamma]$ for a finite group Γ, and the Hopf algebra $A = A(G_*) =: A(\mathfrak{S}[\Gamma])$. Recall (from Theorem 4.4.1) that irreducible complex representations χ^λ of $\mathfrak{S}_n$ are indexed by partitions λ in Par_n. Index the irreducible complex representations of Γ as $\operatorname{Irr}(\Gamma) = \{\rho_1, \ldots, \rho_d\}$.

Definition 4.5.1. Define for a partition λ in Par_n and ρ in $\operatorname{Irr}(\Gamma)$ a representation $\chi^{\lambda,\rho}$ of $\mathfrak{S}_n[\Gamma]$ in which σ in $\mathfrak{S}_n$ and $\gamma = (\gamma_1, \ldots, \gamma_n)$ in Γ^n act on the space $\chi^\lambda \otimes (\rho^{\otimes n})$ as follows:

$$\begin{aligned}
\sigma(u \otimes (v_1 \otimes \cdots \otimes v_n)) &= \sigma(u) \otimes (v_{\sigma^{-1}(1)} \otimes \cdots \otimes v_{\sigma^{-1}(n)}); \\
\gamma(u \otimes (v_1 \otimes \cdots \otimes v_n)) &= u \otimes (\gamma_1 v_1 \otimes \cdots \otimes \gamma_n v_n).
\end{aligned} \tag{4.5.1}$$

Theorem 4.5.2. *The irreducible $\mathbb{C}\mathfrak{S}_n[\Gamma]$-modules are the induced characters*

$$\chi^{\underline{\lambda}} := \operatorname{Ind}_{\mathfrak{S}_{\operatorname{degs}(\underline{\lambda})}[\Gamma]}^{\mathfrak{S}_n[\Gamma]} \left(\chi^{\lambda^{(1)},\rho_1} \otimes \cdots \otimes \chi^{\lambda^{(d)},\rho_d} \right)$$

as $\underline{\lambda}$ runs through all functions

$$\begin{array}{ccc} \operatorname{Irr}(\Gamma) & \xrightarrow{\underline{\lambda}} & \operatorname{Par}, \\ \rho_i & \longmapsto & \lambda^{(i)} \end{array}$$

with the property that $\sum_{i=1}^d |\lambda^{(i)}| = n$. Here, $\operatorname{degs}(\underline{\lambda})$ denotes the d-tuple $\left(\left|\lambda^{(1)}\right|, \left|\lambda^{(2)}\right|, \ldots, \left|\lambda^{(d)}\right|\right) \in \mathbb{N}^d$, and $\mathfrak{S}_{\operatorname{degs}(\underline{\lambda})}$ is defined as the subgroup $\mathfrak{S}_{\left|\lambda^{(1)}\right|} \times \mathfrak{S}_{\left|\lambda^{(2)}\right|} \times \cdots \times \mathfrak{S}_{\left|\lambda^{(d)}\right|}$ of $\mathfrak{S}_n$.

Furthermore, one has a PSH-isomorphism

$$\begin{array}{ccc} A(\mathfrak{S}[\Gamma]) & \longrightarrow & \Lambda^{\otimes d}, \\ \chi^{\underline{\lambda}} & \longmapsto & s_{\lambda^{(1)}} \otimes \cdots \otimes s_{\lambda^{(d)}}. \end{array}$$

Proof. We know from Corollary 4.3.10 that $A(\mathfrak{S}[\Gamma])$ is a PSH, with PSH-basis Σ given by the union of all irreducible characters of all groups $\mathfrak{S}_n[\Gamma]$. Therefore Theorem 3.2.3 tells us that $A(\mathfrak{S}[\Gamma]) \cong \bigotimes_{\rho \in \mathcal{C}} A(\mathfrak{S}[\Gamma])(\rho)$ where $\mathcal{C}$ is the set of irreducible characters which are also primitive. Just as in the case of $\mathfrak{S}_n$, it is clear from the definition of the coproduct that an irreducible character ρ of $\mathfrak{S}_n[\Gamma]$ is primitive if and only if $n = 1$, that in this case $\mathfrak{S}_n[\Gamma] = \Gamma$, and ρ lies in $\operatorname{Irr}(\Gamma) = \{\rho_1, \ldots, \rho_d\}$.

The remaining assertions of the theorem will then follow from the definition of the induction product algebra structure on $A(\mathfrak{S}[\Gamma])$, once we have shown that, for every $\rho \in \operatorname{Irr}(\Gamma)$, there is a PSH-isomorphism sending

$$\begin{array}{ccc} A(\mathfrak{S}) & \longrightarrow & A(\mathfrak{S}[\Gamma])(\rho), \\ \chi^\lambda & \longmapsto & \chi^{\lambda,\rho}. \end{array} \tag{4.5.2}$$

[237] This is well-defined, since $(p_\lambda)_{\lambda \in \operatorname{Par}}$ is a $\mathbb{Q}$-module basis of $\Lambda_{\mathbb{Q}}$.

Such an isomorphism comes from applying Proposition 4.1.17 to the semidirect product $\mathfrak{S}_n[\Gamma] = \mathfrak{S}_n \ltimes \Gamma^n$, so that $K = \Gamma^n, G = \mathfrak{S}_n$, and fixing $V = \rho^{\otimes n}$ as $\mathbb{C}\mathfrak{S}_n[\Gamma]$-module with structure as defined in (4.5.1) (but with λ set to (n), so that χ^λ is the trivial 1-dimensional $\mathbb{C}\mathfrak{S}_n$-module). One obtains for each n, maps

$$R(\mathfrak{S}_n) \underset{\Psi}{\overset{\Phi}{\rightleftarrows}} R(\mathfrak{S}_n[\Gamma])$$

where

$$\begin{array}{rcl} \chi & \overset{\Phi}{\longmapsto} & \chi \otimes (\rho^{\otimes n}), \\ \alpha & \overset{\Psi}{\longmapsto} & \operatorname{Hom}_{\mathbb{C}\Gamma^n}(\rho^{\otimes n}, \alpha). \end{array}$$

Taking the direct sum of these maps for all n gives maps $A(\mathfrak{S}) \underset{\Psi}{\overset{\Phi}{\rightleftarrows}} A(\mathfrak{S}[\Gamma])$.

These maps are coalgebra morphisms because of their interaction with restriction to $\mathfrak{S}_i \times \mathfrak{S}_j$. Since Proposition 4.1.17(iii) gives the adjointness property that

$$(\chi, \Psi(\alpha))_{A(\mathfrak{S})} = (\Phi(\chi), \alpha)_{A(\mathfrak{S}[\Gamma])},$$

one concludes from the self-duality of $A(\mathfrak{S}), A(\mathfrak{S}[\Gamma])$ that Φ, Ψ are also algebra morphisms. Since they take genuine characters to genuine characters, they are PSH-morphisms. Since ρ being a simple $\mathbb{C}\Gamma$-module implies that $V = \rho^{\otimes n}$ is a simple $\mathbb{C}\Gamma^n$-module, Proposition 4.1.17(iv) shows that

$$(\Psi \circ \Phi)(\chi) = \chi \tag{4.5.3}$$

for all $\mathfrak{S}_n$-characters χ. Hence Φ is an injective PSH-morphism. Using adjointness, (4.5.3) also shows that Φ sends $\mathbb{C}\mathfrak{S}_n$-simples χ to $\mathbb{C}[\mathfrak{S}_n[\Gamma]]$-simples $\Phi(\chi)$:

$$(\Phi(\chi), \Phi(\chi))_{A(\mathfrak{S}[\Gamma])} = ((\Psi \circ \Phi)(\chi), \chi)_{A(\mathfrak{S})} = (\chi, \chi)_{A(\mathfrak{S})} = 1.$$

Since $\Phi(\chi) = \chi \otimes (\rho^{\otimes n})$ has $V = \rho^{\otimes n}$ as a constituent upon restriction to Γ^n, Frobenius Reciprocity shows that the irreducible character $\Phi(\chi)$ is a constituent of $\operatorname{Ind}_{\Gamma^n}^{\mathfrak{S}_n[\Gamma]} \rho^{\otimes n} = \rho^n$. Hence the entire image of Φ lies in $A(\mathfrak{S}[\Gamma])(\rho)$ (due to how we defined $A(\rho)$ in the proof of Theorem 3.2.3), and so Φ must restrict to an isomorphism as desired in (4.5.2). □

One of Zelevinsky's sample applications of the theorem is this branching rule.

Corollary 4.5.3. *Given* $\underline{\lambda} = (\lambda^{(1)}, \ldots, \lambda^{(d)})$ *with* $\sum_{i=1}^d |\lambda^{(i)}| = n$, *one has*

$$\operatorname{Res}^{\mathfrak{S}_n[\Gamma]}_{\mathfrak{S}_{n-1}[\Gamma] \times \Gamma} \left(\chi^{\underline{\lambda}}\right) = \sum_{i=1}^{d} \sum_{\substack{\lambda^{(i)}_- \subseteq \lambda^{(i)}: \\ |\lambda^{(i)}/\lambda^{(i)}_-|=1}} \chi^{(\lambda^{(1)}, \ldots, \lambda^{(i)}_-, \ldots, \lambda^{(d)})} \otimes \rho_i.$$

(We are identifying functions $\underline{\lambda} : \operatorname{Irr}(\Gamma) \to \operatorname{Par}$ *with the corresponding d-tuples* $\left(\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(d)}\right)$ *here.)*

Example 4.5.4. For Γ a two-element group, so $\operatorname{Irr}(\Gamma) = \{\rho_1, \rho_2\}$ and $d = 2$, then

$$\operatorname{Res}^{\mathfrak{S}_6[\Gamma]}_{\mathfrak{S}_5[\Gamma] \times \Gamma} \left(\chi^{((3,1),(1,1))}\right) = \chi^{((3),(1,1))} \otimes \rho_1 + \chi^{((2,1),(1,1))} \otimes \rho_1 + \chi^{((3,1),(1))} \otimes \rho_2.$$

Proof of Corollary 4.5.3. By Theorem 4.5.2, this is equivalent to computing in the Hopf algebra $A := \Lambda^{\otimes d}$ the component of the coproduct of $s_{\lambda^{(1)}} \otimes \cdots \otimes s_{\lambda^{(d)}}$ that lies in $A_{n-1} \otimes A_1$. Working within each tensor factor Λ, we conclude from Proposition 2.3.6(iv) that the $\Lambda_{|\lambda|-1} \otimes \Lambda_1$-component of $\Delta(s_\lambda)$ is

$$\sum_{\substack{\lambda_- \subseteq \lambda: \\ |\lambda/\lambda_-|=1}} s_{\lambda_-} \otimes \rho.$$

One must apply this in each of the d tensor factors of $A = \Lambda^{\otimes d}$, then sum on i. □

4.6. **General linear groups.** We now consider the tower of finite general linear groups $G_n = GL_n = GL_n(\mathbb{F}_q)$ and $A = A(G_*) =: A(GL)$. Corollary 4.3.10 tells us that $A(GL)$ is a PSH, with PSH-basis Σ given by the union of all irreducible characters of all groups GL_n. Therefore Theorem 3.2.3 tells us that

$$A(GL) \cong \bigotimes_{\rho \in \mathcal{C}} A(GL)(\rho) \tag{4.6.1}$$

where $\mathcal{C} = \Sigma \cap \mathfrak{p}$ is the set of primitive irreducible characters.

Definition 4.6.1. Call an irreducible representation ρ of GL_n *cuspidal* for $n \geq 1$ if it lies in $\mathcal{C}$, that is, its restriction to proper parabolic subgroups $P_{i,j}$ with $i+j = n$ and $i, j > 0$ contain no nonzero vectors which are $K_{i,j}$-invariant. Given an irreducible character σ of GL_n, say that $d(\sigma) = n$, and let $\mathcal{C}_n := \{\rho \in \mathcal{C} : d(\rho) = n\}$ for $n \geq 1$ denote the subset of cuspidal characters of GL_n.

Just as was the case for $\mathfrak{S}_1$ and $\mathfrak{S}_1[\Gamma] = \Gamma$, *every* irreducible character ρ of $GL_1(\mathbb{F}_q) = \mathbb{F}_q^\times$ is cuspidal. However, this does not exhaust the cuspidal characters. In fact, one can predict the number of cuspidal characters in $\mathcal{C}_n$, using knowledge of the number of conjugacy classes in GL_n. Let $\mathcal{F}$ denote the set of all nonconstant monic irreducible polynomials $f(x) \neq x$ in $\mathbb{F}_q[x]$. Let $\mathcal{F}_n := \{f \in \mathcal{F} : \deg(f) = n\}$ for $n \geq 1$.

Proposition 4.6.2. *The number $|\mathcal{C}_n|$ of cuspidal characters of $GL_n(\mathbb{F}_q)$ is the number of $|\mathcal{F}_n|$ of irreducible monic degree n polynomials $f(x) \neq x$ in $\mathbb{F}_q[x]$ with nonzero constant term.*

Proof. We show $|\mathcal{C}_n| = |\mathcal{F}_n|$ for $n \geq 1$ by strong induction on n. For the base case[238] $n = 1$, just as with the families $G_n = \mathfrak{S}_n$ and $G_n = \mathfrak{S}_n[\Gamma]$, when $n = 1$ any irreducible character χ of $G_1 = GL_1(\mathbb{F}_q)$ gives a primitive element of $A = A(GL)$, and hence is cuspidal. Since $GL_1(\mathbb{F}_q) = \mathbb{F}_q^\times$ is abelian, there are $|\mathbb{F}_q^\times| = q-1$ such cuspidal characters in $\mathcal{C}_1$, which agrees with the fact that there are $q-1$ monic (irreducible) linear polynomials $f(x) \neq x$ in $\mathbb{F}_q[x]$, namely $\mathcal{F}_1 := \{f(x) = x - c : c \in \mathbb{F}_q^\times\}$.

In the inductive step, use the fact that the number $|\Sigma_n|$ of irreducible complex characters χ of $GL_n(\mathbb{F}_q)$ equals its number of conjugacy classes. These conjugacy classes are uniquely represented by *rational canonical forms*, which are parametrized by functions $\underline{\lambda} : \mathcal{F} \to \mathrm{Par}$ with the property that $\sum_{f \in \mathcal{F}} \deg(f) |\underline{\lambda}(f)| = n$. On the other hand, (4.6.1) tells us that $|\Sigma_n|$ is similarly parametrized by the functions $\underline{\lambda} : \mathcal{C} \to \mathrm{Par}$ having the property that $\sum_{\rho \in \mathcal{C}} d(\rho) |\underline{\lambda}(\rho)| = n$. Thus we have parallel disjoint decompositions

$$\begin{array}{lll} \mathcal{C} & = \bigsqcup_{n \geq 1} \mathcal{C}_n & \text{where } \mathcal{C}_n = \{\rho \in \mathcal{C} : d(\rho) = n\}, \\ \mathcal{F} & = \bigsqcup_{n \geq 1} \mathcal{F}_n & \text{where } \mathcal{F}_n = \{f \in \mathcal{F} : \deg(f) = n\}, \end{array}$$

and hence an equality for all $n \geq 1$

$$\left| \left\{ \mathcal{C} \xrightarrow{\underline{\lambda}} \mathrm{Par} : \quad \sum_{\rho \in \mathcal{C}} d(\rho) |\underline{\lambda}(\rho)| = n \right\} \right| = |\Sigma_n| = \left| \left\{ \mathcal{F} \xrightarrow{\underline{\lambda}} \mathrm{Par} : \quad \sum_{f \in \mathcal{F}} \deg(f) |\underline{\lambda}(f)| = n \right\} \right|.$$

Since there is only one partition λ having $|\lambda| = 1$ (namely, $\lambda = (1)$), this leads to parallel recursions

$$|\mathcal{C}_n| = |\Sigma_n| - \left| \left\{ \bigsqcup_{i=1}^{n-1} \mathcal{C}_i \xrightarrow{\underline{\lambda}} \mathrm{Par} : \quad \sum_{\rho \in \mathcal{C}} d(\rho) |\underline{\lambda}(\rho)| = n \right\} \right|,$$

$$|\mathcal{F}_n| = |\Sigma_n| - \left| \left\{ \bigsqcup_{i=1}^{n-1} \mathcal{F}_i \xrightarrow{\underline{\lambda}} \mathrm{Par} : \quad \sum_{f \in \mathcal{F}} \deg(f) |\underline{\lambda}(f)| = n \right\} \right|,$$

and induction implies that $|\mathcal{C}_n| = |\mathcal{F}_n|$. □

We shall use the notation $\underline{1}_H$ for the trivial character of a group H whenever H is a finite group. This generalizes the notations $\underline{1}_{\mathfrak{S}_n}$ and $\underline{1}_{\mathfrak{S}_\lambda}$ introduced above.

[238] Actually, we don't need any base case for our strong induction. We nevertheless handle the case $n = 1$ as a warmup.

Example 4.6.3. Taking $q = 2$, let us list the sets $\mathcal{F}_n$ of monic irreducible polynomials $f(x) \neq x$ in $\mathbb{F}_2[x]$ of degree n for $n \leq 3$, so that we know how many cuspidal characters of $GL_n(\mathbb{F}_q)$ in $\mathcal{C}_n$ to expect:

$$\begin{aligned}
\mathcal{F}_1 &= \{x + 1\}; \\
\mathcal{F}_2 &= \{x^2 + x + 1\}; \\
\mathcal{F}_3 &= \{x^3 + x + 1, x^3 + x^2 + 1\}.
\end{aligned}$$

Thus we expect

- one cuspidal character of $GL_1(\mathbb{F}_2)$, namely $\rho_1 (= \underline{1}_{GL_1(\mathbb{F}_2)})$,
- one cuspidal character ρ_2 of $GL_2(\mathbb{F}_2)$, and
- two cuspidal characters ρ_3, ρ_3' of $GL_3(\mathbb{F}_2)$.

We will say more about ρ_2, ρ_3, ρ_3' in the next section.

Exercise 4.6.4. Let $\mu : \{1, 2, 3, \ldots\} \to \mathbb{Z}$ denote the *number-theoretic Möbius function*, defined by setting $\mu(m) = (-1)^d$ if $m = p_1 \cdots p_d$ for d distinct primes $p_1, p_2, \ldots, p_d$, and $\mu(m) = 0$ if m is not squarefree.

(a) Show that for $n \geq 2$, we have

$$|\mathcal{C}_n| (= |\mathcal{F}_n|) = \frac{1}{n} \sum_{d | n} \mu\left(\frac{n}{d}\right) q^d. \tag{4.6.2}$$

(Here, the summation sign $\sum_{d|n}$ means a sum over all positive divisors d of n.)

(b) Show that (4.6.2) also counts the *necklaces* with n beads of q colors (= the equivalence classes under the $\mathbb{Z}/n\mathbb{Z}$-action of cyclic rotation on sequences $(a_1, \ldots, a_n)$ in $\mathbb{F}_q^n$) which are *primitive* in the sense that no nontrivial rotation fixes any of the sequences within the equivalence class. For example, when $q = 2$, here are systems of distinct representatives of these primitive necklaces for $n = 2, 3, 4$:

$$\begin{aligned}
n = 2 : &\quad \{(0, 1)\}; \\
n = 3 : &\quad \{(0, 0, 1), (0, 1, 1)\}; \\
n = 4 : &\quad \{(0, 0, 0, 1), (0, 0, 1, 1), (0, 1, 1, 1)\}.
\end{aligned}$$

The result of Exercise 4.6.4(a) was stated by Gauss for prime q, and by Witt for general q; it is discussed in [37], [182, Section 7.6.2] and (for prime q) [84, (4.12.3)]. Exercise 4.6.4(b) is also well-known. See [182, Section 7.6.2] for a bijection explaining why the answers to both parts of Exercise 4.6.4 are the same.

4.7. **Steinberg's unipotent characters.** Not surprisingly, the (cuspidal) character $\iota := \underline{1}_{GL_1}$ of $GL_1(\mathbb{F}_q)$ plays a distinguished role. The parabolic subgroup $P_{(1^n)}$ of $GL_n(\mathbb{F}_q)$ is the Borel subgroup B of upper triangular matrices, and we have $\iota^n = \operatorname{Ind}_B^{GL_n} \underline{1}_B = \mathbb{C}[GL_n/B]$ (identifying representations with their characters as usual)[239]. The subalgebra $A(GL)(\iota)$ of $A(GL)$ is the $\mathbb{Z}$-span of the irreducible characters σ that appear as constituents of $\iota^n = \operatorname{Ind}_B^{GL_n} \underline{1}_B = \mathbb{C}[GL_n/B]$ for some n.

Definition 4.7.1. An irreducible character σ of GL_n appearing as a constituent of $\operatorname{Ind}_B^{GL_n} \underline{1}_B = \mathbb{C}[GL_n/B]$ is called a *unipotent character*. Equivalently, by Frobenius reciprocity, σ is unipotent if it contains a nonzero B-invariant vector.

In particular, $\underline{1}_{GL_n}$ is a unipotent character of GL_n for each n.

Proposition 4.7.2. *One can choose $\Lambda \cong A(GL)(\iota)$ in Theorem 3.3.3(g) so that $h_n \longmapsto \underline{1}_{GL_n}$.*

[239] *Proof.* Exercise 4.3.11(d) (applied to $G_* = GL_*$, $\ell = n$, $\alpha = (1^n) = \left(\underbrace{1, 1, \ldots, 1}_{n \text{ times}}\right)$ and $\chi_i = \iota$) gives

$$\iota^n = \operatorname{ind}_{(1^n)}^n \iota^{\otimes n} = \underbrace{\operatorname{Ind}_{P_{(1^n)}}^{G_n}}_{= \operatorname{Ind}_B^{GL_n}} \underbrace{\operatorname{Infl}_{G_{(1^n)}}^{P_{(1^n)}} \iota^{\otimes n}}_{= \underline{1}_{P_{(1^n)}} = \underline{1}_B} = \operatorname{Ind}_B^{GL_n} \underline{1}_B = \mathbb{C}\left[GL_n/B\right],$$

where the last equality follows from the general fact that if G is a finite group and H is a subgroup of G, then $\operatorname{Ind}_H^G \underline{1}_H \cong \mathbb{C}\left[G/H\right]$ as $\mathbb{C}G$-modules.

Proof. Theorem 3.3.1(a) tells us $\iota^2 = \mathrm{Ind}_B^{GL_2} \underline{1}_B$ must have exactly two irreducible constituents, one of which is $\underline{1}_{GL_2}$; call the other one St_2. Choose the isomorphism so as to send $h_2 \longmapsto \underline{1}_{GL_2}$. Then $h_n \mapsto \underline{1}_{GL_n}$ follows from the claim that $\mathrm{St}_2^{\perp}(\underline{1}_{GL_n}) = 0$ for $n \geq 2$: one has

$$\Delta(\underline{1}_{GL_n}) = \sum_{i+j=n} \left(\mathrm{Res}^{G_n}_{P_{i,j}} \underline{1}_{GL_n}\right)^{K_{i,j}} = \sum_{i+j=n} \underline{1}_{GL_i} \otimes \underline{1}_{GL_j}$$

so that $\mathrm{St}_2^{\perp}(\underline{1}_{GL_n}) = (\mathrm{St}_2, \underline{1}_{GL_2})\underline{1}_{GL_{n-2}} = 0$ since $\mathrm{St}_2 \neq \underline{1}_{GL_2}$.

□

This subalgebra $A(GL)(\iota)$, and the unipotent characters χ_q^{λ} corresponding under this isomorphism to the Schur functions s_λ, were introduced by Steinberg [208]. He wrote down χ_q^λ as a virtual sum of induced characters $\mathrm{Ind}_{P_\alpha}^{GL_n} \underline{1}_{P_\alpha} (= \underline{1}_{G_{\alpha_1}} \cdots \underline{1}_{G_{\alpha_\ell}})$, modelled on the Jacobi-Trudi determinantal expression for $s_\lambda = \det(h_{\lambda_i - i + j})$. Note that $\mathrm{Ind}_{P_\alpha}^{GL_n} \underline{1}_{P_\alpha}$ is the transitive permutation representation $\mathbb{C}[G/P_\alpha]$ for GL_n permuting the *finite partial flag variety* G/P_α, that is, the set of *α-flags of subspaces*

$$\{0\} \subset V_{\alpha_1} \subset V_{\alpha_1+\alpha_2} \subset \cdots \subset V_{\alpha_1+\alpha_2+\cdots+\alpha_{\ell-1}} \subset \mathbb{F}_q^n$$

where $\dim_{\mathbb{F}_q} V_d = d$ in each case. This character has dimension equal to $|G/P_\alpha|$, with formula given by the *q-multinomial coefficient* (see e.g. Stanley [206, §1.7]):

$$\begin{bmatrix} n \\ \alpha \end{bmatrix}_q = \frac{[n]!_q}{[\alpha_1]!_q \cdots [\alpha_\ell]!_q}$$

where $[n]!_q := [n]_q [n-1]_q \cdots [2]_q [1]_q$ and $[n]_q := 1 + q + \cdots + q^{n-1} = \frac{q^n - 1}{q-1}$.

Our terminology St_2 is motivated by the $n = 2$ special case of the *Steinberg character* St_n, which is the unipotent character corresponding under the isomorphism in Proposition 4.7.2 to $e_n = s_{(1^n)}$. It can be defined by the virtual sum

$$\mathrm{St}_n := \chi_q^{(1^n)} = \sum_{\alpha} (-1)^{n-\ell(\alpha)} \mathrm{Ind}_{P_\alpha}^{GL_n} \underline{1}_{P_\alpha}$$

in which the sum runs through all compositions α of n. This turns out to be the genuine character for $GL_n(\mathbb{F}_q)$ acting on the top homology group of its *Tits building*: the simplicial complex whose vertices are nonzero proper subspaces V of $\mathbb{F}_q^n$, and whose simplices correspond to flags of nested subspaces. One needs to know that this Tits building has only top homology, so that one can deduce the above character formula from the Hopf trace formula; see Björner [22].

4.8. **Examples: $GL_2(\mathbb{F}_2)$ and $GL_3(\mathbb{F}_2)$.** Let's get our hands dirty.

Example 4.8.1. For $n = 2$, there are two unipotent characters, $\chi_q^{(2)} = \underline{1}_{GL_2}$ and

$$\mathrm{St}_2 := \chi_q^{(1,1)} = \underline{1}_{GL_1}^2 - \underline{1}_{GL_2} = \mathrm{Ind}_B^{GL_2} \underline{1}_B - \underline{1}_{GL_2} \tag{4.8.1}$$

since the Jacobi-Trudi formula (2.4.16) gives $s_{(1,1)} = \det \begin{bmatrix} h_1 & h_2 \\ 1 & h_1 \end{bmatrix} = h_1^2 - h_2$. The description (4.8.1) for this Steinberg character St_2 shows that it has dimension

$$|GL_2/B| - 1 = (q+1) - 1 = q$$

and that one can think of it as follows: consider the permutation action of GL_2 on the $q+1$ lines $\{\ell_0, \ell_1, \ldots, \ell_q\}$ in the projective space $\mathbb{P}^1_{\mathbb{F}_q} = GL_2(\mathbb{F}_q)/B$, and take the invariant subspace perpendicular to the sum of basis elements $e_{\ell_0} + \cdots + e_{\ell_q}$.

Example 4.8.2. Continuing the previous example, but taking $q = 2$, we find that we have constructed two unipotent characters: $\underline{1}_{GL_2} = \chi_{q=2}^{(2)}$ of dimension 1, and $\mathrm{St}_2 = \chi_{q=2}^{(1,1)}$ of dimension $q = 2$. This lets us identify the unique cuspidal character ρ_2 of $GL_2(\mathbb{F}_2)$, using knowledge of the character table of $GL_2(\mathbb{F}_2) \cong \mathfrak{S}_3$:

		$\begin{bmatrix}1&0\\0&1\end{bmatrix}$	$\begin{bmatrix}1&1\\0&1\end{bmatrix}, \begin{bmatrix}1&0\\1&1\end{bmatrix}, \begin{bmatrix}0&1\\1&0\end{bmatrix}$	$\begin{bmatrix}1&1\\1&0\end{bmatrix}, \begin{bmatrix}0&1\\1&1\end{bmatrix}$
$\underline{1}_{GL_2} = \chi^{(2)}_{q=2}$	unipotent	1	1	1
$\mathrm{St}_2 = \chi^{(1,1)}_{q=2}$	unipotent	2	0	-1
ρ_2	cuspidal	1	-1	1

In other words, the cuspidal character ρ_2 of $GL_2(\mathbb{F}_2)$ corresponds under the isomorphism $GL_2(\mathbb{F}_2) \cong \mathfrak{S}_3$ to the sign character $\mathrm{sgn}_{\mathfrak{S}_3}$.

Example 4.8.3. Continuing the previous example to $q = 2$ and $n = 3$ lets us analyze the irreducible characters of $GL_3(\mathbb{F}_2)$. Recalling our labelling $\rho_1, \rho_2, \rho_3, \rho_3'$ from Example 4.6.3 of the cuspidal characters of $GL_n(\mathbb{F}_2)$ for $n = 1, 2, 3$, Zelevinsky's Theorem 3.2.3 tells us that the $GL_3(\mathbb{F}_2)$-irreducible characters should be labelled by functions $\{\rho_1, \rho_2, \rho_3, \rho_3'\} \xrightarrow{\underline{\lambda}} \mathrm{Par}$ for which

$$1 \cdot |\underline{\lambda}(\rho_1)| + 2 \cdot |\underline{\lambda}(\rho_2)| + 3 \cdot |\underline{\lambda}(\rho_3)| + 3 \cdot |\underline{\lambda}(\rho_3')| = 3.$$

We will label such an irreducible character $\chi^{\underline{\lambda}} = \chi^{(\underline{\lambda}(\rho_1), \underline{\lambda}(\rho_2), \underline{\lambda}(\rho_3), \underline{\lambda}(\rho_3'))}$.

Three of these irreducibles will be the unipotent characters, mapping under the isomorphism from Proposition 4.7.2 as follows:

- $s_{(3)} = h_3 \longmapsto \chi^{((3),\varnothing,\varnothing,\varnothing)} = \underline{1}_{GL_3}$ of dimension 1.
- $$s_{(2,1)} = \det \begin{bmatrix} h_2 & h_3 \\ 1 & h_1 \end{bmatrix} = h_2 h_1 - h_3 \longmapsto \chi^{((2,1),\varnothing,\varnothing,\varnothing)} = \mathrm{Ind}_{P_{2,1}}^{GL_3} \underline{1}_{P_{2,1}} - \underline{1}_{GL_3},$$
  of dimension $\begin{bmatrix} 3 \\ 2,1 \end{bmatrix}_q - \begin{bmatrix} 3 \\ 3 \end{bmatrix}_q = [3]_q - 1 = q^2 + q \overset{q=2}{\rightsquigarrow} 6$.
- Lastly,
  $$s_{(1,1,1)} = \det \begin{bmatrix} h_1 & h_2 & h_3 \\ 1 & h_1 & h_2 \\ 0 & 1 & h_1 \end{bmatrix} = h_1^3 - h_2 h_1 - h_1 h_2 + h_3$$
  $$\longmapsto \mathrm{St}_3 = \chi^{((1,1,1),\varnothing,\varnothing,\varnothing)} = \mathrm{Ind}_B^{GL_3} \underline{1}_B - \mathrm{Ind}_{P_{2,1}}^{GL_3} \underline{1}_{P_{2,1}} - \mathrm{Ind}_{P_{1,2}}^{GL_3} \underline{1}_{P_{1,2}} + \underline{1}_{GL_3}$$
  of dimension
  $$\begin{bmatrix} 3 \\ 1,1,1 \end{bmatrix}_q - \begin{bmatrix} 3 \\ 2,1 \end{bmatrix}_q - \begin{bmatrix} 3 \\ 1,2 \end{bmatrix}_q + \begin{bmatrix} 3 \\ 3 \end{bmatrix}_q$$
  $$= [3]!_q - [3]_q - [3]_q + 1 = q^3 \overset{q=2}{\rightsquigarrow} 8.$$

There should also be one non-unipotent, non-cuspidal character, namely

$$\chi^{((1),(1),\varnothing,\varnothing)} = \rho_1 \rho_2 = \mathrm{Ind}_{P_{1,2}}^{GL_3} \mathrm{Infl}_{GL_1 \times GL_2}^{P_{1,2}} \left(\underline{1}_{GL_1} \otimes \rho_2 \right)$$

having dimension $\begin{bmatrix} 3 \\ 1,2 \end{bmatrix}_q \cdot 1 \cdot 1 = [3]_q \overset{q=2}{\rightsquigarrow} 7$.

Finally, we expect cuspidal characters $\rho_3 = \chi^{(\varnothing,\varnothing,(1),\varnothing)}, \rho_3' = \chi^{(\varnothing,\varnothing,\varnothing,(1))}$, whose dimensions d_3, d_3' can be deduced from the equation

$$1^2 + 6^2 + 8^2 + 7^2 + d_3^2 + (d_3')^2 = |GL_3(\mathbb{F}_2)| = \left[(q^3 - q^0)(q^3 - q^1)(q^3 - q^2) \right]_{q=2} = 168.$$

This forces $d_3^2 + (d_3')^2 = 18$, whose only solution in positive integers is $d_3 = d_3' = 3$.

We can check our predictions of the dimensions for the various $GL_3(\mathbb{F}_2)$-irreducible characters since $GL_3(\mathbb{F}_2)$ is the finite simple group of order 168 (also isomorphic to $PSL_2(\mathbb{F}_7)$), with known character table (see James and Liebeck [104, p. 318]):

	centralizer order	168	8	4	3	7	7
	unipotent?/cuspidal?						
$\underline{1}_{GL_3} = \chi^{((3),\varnothing,\varnothing,\varnothing)}$	unipotent	1	1	1	1	1	1
$\chi^{((2,1),\varnothing,\varnothing,\varnothing)}$	unipotent	6	2	0	0	-1	-1
$\mathrm{St}_3 = \chi^{((1,1,1),\varnothing,\varnothing,\varnothing)}$	unipotent	8	0	0	-1	1	1
$\chi^{((1),(1),\varnothing,\varnothing)}$		7	-1	-1	1	0	0
$\rho_3 = \chi^{(\varnothing,\varnothing,(1),\varnothing)}$	cuspidal	3	-1	1	0	α	$\overline{\alpha}$
$\rho_3' = \chi^{(\varnothing,\varnothing,\varnothing,(1))}$	cuspidal	3	-1	1	0	$\overline{\alpha}$	α

Here $\alpha := -1/2 + i\sqrt{7}/2$.

Remark 4.8.4. It is known (see e.g. Bump [30, Cor. 7.4]) that, for $n \geq 2$, the dimension of any cuspidal irreducible character ρ of $GL_n(\mathbb{F}_q)$ is

$$(q^{n-1}-1)(q^{n-2}-1)\cdots(q^2-1)(q-1).$$

Note that when $q = 2$,

- for $n = 2$ this gives $2^1 - 1 = 1$ for the dimension of ρ_2, and
- for $n = 3$ it gives $(2^2-1)(2-1) = 3$ for the dimensions of ρ_3, ρ_3',

agreeing with our calculations above. Much more is known about the character table of $GL_n(\mathbb{F}_q)$; see Remark 4.9.14 below, Zelevinsky [227, Chap. 11], and Macdonald [142, Chap. IV].

4.9. **The Hall algebra.**

There is another interesting Hopf subalgebra (and quotient Hopf algebra) of $A(GL)$, related to unipotent conjugacy classes in $GL_n(\mathbb{F}_q)$.

Definition 4.9.1. Say that an element g in $GL_n(\mathbb{F}_q)$ is *unipotent* if its eigenvalues are all equal to 1. Equivalently, $g \in GL_n(\mathbb{F}_q)$ is unipotent if and only if $g - \mathrm{id}_{\mathbb{F}_q^n}$ is nilpotent. A conjugacy class in $GL_n(\mathbb{F}_q)$ is *unipotent* if its elements are unipotent.

Denote by $\mathcal{H}_n$ the $\mathbb{C}$-subspace of $R_{\mathbb{C}}(GL_n)$ consisting of those class functions which are supported only on unipotent conjugacy classes, and let $\mathcal{H} = \bigoplus_{n\geq 0} \mathcal{H}_n$ as a $\mathbb{C}$-subspace of $A_{\mathbb{C}}(GL) = \bigoplus_{n\geq 0} R_{\mathbb{C}}(GL_n)$.

Proposition 4.9.2. *The subspace $\mathcal{H}$ is a Hopf subalgebra of $A_{\mathbb{C}}(GL)$, which is graded, connected, and of finite type, and self-dual with respect to the inner product on class functions inherited from $A_{\mathbb{C}}(GL)$. It is also a quotient Hopf algebra of $A_{\mathbb{C}}(GL)$, as the $\mathbb{C}$-linear surjection $A_{\mathbb{C}}(GL) \twoheadrightarrow \mathcal{H}$ restricting class functions to unipotent classes is a Hopf algebra homomorphism. This surjection has kernel $\mathcal{H}^\perp$, which is both an ideal and a two-sided coideal.*

Proof. It is immediately clear that $\mathcal{H}^\perp$ is a graded $\mathbb{C}$-vector subspace of $A_{\mathbb{C}}(GL)$, whose n-th homogeneous component consists of those class functions on GL_n whose values on all unipotent classes are 0. (This holds no matter whether the perpendicular space is taken with respect to the Hermitian form $(\cdot,\cdot)_G$ or with respect to the bilinear form $\langle\cdot,\cdot\rangle_G$.) In other words, $\mathcal{H}^\perp$ is the kernel of the surjection $A_{\mathbb{C}}(GL) \twoheadrightarrow \mathcal{H}$ defined in the proposition.

Given two class functions χ_i, χ_j on GL_i, GL_j and g in GL_{i+j}, one has

$$(\chi_i \cdot \chi_j)(g) = \frac{1}{|P_{i,j}|} \sum_{\substack{h \in GL_{i+j}: \\ h^{-1}gh = \begin{bmatrix} g_i & * \\ 0 & g_j \end{bmatrix} \in P_{i,j}}} \chi_i(g_i)\chi_j(g_j). \tag{4.9.1}$$

Since g is unipotent if and only if $h^{-1}gh$ is unipotent if and only if both g_i, g_j are unipotent, the formula (4.9.1) shows both that $\mathcal{H}$ is a subalgebra[240] and that $\mathcal{H}^\perp$ is a two-sided ideal[241]. It also shows that the surjection $A_{\mathbb{C}}(GL) \twoheadrightarrow \mathcal{H}$ restricting every class function to unipotent classes is an algebra homomorphism[242].

[240] Indeed, if χ_i and χ_j are both supported only on unipotent classes, then the same holds for $\chi_i \cdot \chi_j$.

[241] In fact, if one of χ_i and χ_j annihilates all unipotent classes, then so does $\chi_i \cdot \chi_j$.

[242] because if g is unipotent, then the only values of χ_i and χ_j appearing on the right hand side of (4.9.1) are those on unipotent elements

Similarly, for class functions χ on GL_n and (g_i, g_j) in $GL_{i,j} = GL_i \times GL_j$, one has

$$\Delta(\chi)(g_i, g_j) = \frac{1}{q^{ij}} \sum_{k \in \mathbb{F}_q^{i \times j}} \chi \begin{bmatrix} g_i & k \\ 0 & g_j \end{bmatrix}$$

using (4.1.13). This shows both that $\mathcal{H}$ is a sub-coalgebra of $A = A_{\mathbb{C}}(GL)$ (that is, it satisfies $\Delta\mathcal{H} \subset \mathcal{H} \otimes \mathcal{H}$) and that $\mathcal{H}^\perp$ is a two-sided coideal (that is, we have $\Delta(\mathcal{H}^\perp) \subset \mathcal{H}^\perp \otimes A + A \otimes \mathcal{H}^\perp$), since it shows that if χ is supported only on unipotent classes, then $\Delta(\chi)$ vanishes on (g_1, g_2) that have either g_1 or g_2 non-unipotent. It also shows that the surjection $A_{\mathbb{C}}(GL) \twoheadrightarrow \mathcal{H}$ restricting every class function to unipotent classes is a coalgebra homomorphism. The rest follows. □

The subspace $\mathcal{H}$ is called the *Hall algebra*. It has an obvious orthogonal $\mathbb{C}$-basis, with interesting structure constants.

Definition 4.9.3. Given a partition λ of n, let J_λ denote the GL_n-conjugacy class of unipotent matrices whose *Jordan type* (that is, the list of the sizes of the Jordan blocks, in decreasing order) is given by λ. Furthermore, let $z_\lambda(q)$ denote the size of the centralizer of any element of this conjugacy class J_λ.

The indicator class functions[243] $\{\underline{1}_{J_\lambda}\}_{\lambda \in \mathrm{Par}}$ form a $\mathbb{C}$-basis for $\mathcal{H}$ whose multiplicative structure constants are called the *Hall coefficients* $g^\lambda_{\mu,\nu}(q)$:

$$\underline{1}_{J_\mu} \underline{1}_{J_\nu} = \sum_\lambda g^\lambda_{\mu,\nu}(q) \, \underline{1}_{J_\lambda}.$$

Because the dual basis to $\{\underline{1}_{J_\lambda}\}$ is $\{z_\lambda(q) \underline{1}_{J_\lambda}\}$, self-duality of $\mathcal{H}$ shows that the Hall coefficients are (essentially) also structure constants for the comultiplication:

$$\Delta \underline{1}_{J_\lambda} = \sum_{\mu,\nu} g^\lambda_{\mu,\nu}(q) \frac{z_\mu(q) z_\nu(q)}{z_\lambda(q)} \cdot \underline{1}_{J_\mu} \otimes \underline{1}_{J_\nu}.$$

The Hall coefficient $g^\lambda_{\mu,\nu}(q)$ has the following interpretation.

Proposition 4.9.4. *Fix any g in $GL_n(\mathbb{F}_q)$ acting unipotently on $\mathbb{F}_q^n$ with Jordan type λ. Then $g^\lambda_{\mu,\nu}(q)$ counts the g-stable $\mathbb{F}_q$-subspaces $V \subset \mathbb{F}_q^n$ for which the restriction $g|V$ acts with Jordan type μ, and the induced map $\bar{g}$ on the quotient space $\mathbb{F}_q^n/V$ has Jordan type ν.*

Proof. Given μ, ν partitions of i, j with $i + j = n$, taking χ_i, χ_j equal to $\underline{1}_{J_\mu}, \underline{1}_{J_\nu}$ in (4.9.1) shows that for any g in GL_n, the value of $\left(\underline{1}_{J_\mu} \cdot \underline{1}_{J_\nu}\right)(g)$ is given by

$$\frac{1}{|P_{i,j}|} \left| \left\{ h \in GL_n : h^{-1} g h = \begin{bmatrix} g_i & * \\ 0 & g_j \end{bmatrix} \text{ with } g_i \in J_\mu, g_j \in J_\nu \right\} \right|. \tag{4.9.2}$$

Let S denote the set appearing in (4.9.2), and let $\mathbb{F}_q^i$ denote the i-dimensional subspace of $\mathbb{F}_q^n$ spanned by the first i standard basis vectors. Note that the condition on an element h in S saying that $h^{-1}gh$ is in block upper-triangular form can be re-expressed by saying that the subspace $V := h(\mathbb{F}_q^i)$ is g-stable. One then sees that the map $h \overset{\varphi}{\longmapsto} V = h(\mathbb{F}_q^i)$ surjects S onto the set of i-dimensional g-stable subspaces V of $\mathbb{F}_q^n$ for which $g|V$ and $\bar{g}$ are unipotent of types μ, ν, respectively. Furthermore, for any particular such V, its fiber $\varphi^{-1}(V)$ in S is a coset of the stabilizer within GL_n of V, which is conjugate to $P_{i,j}$, and hence has cardinality $|\varphi^{-1}(V)| = |P_{i,j}|$. This proves the assertion of the proposition. □

The Hall algebra $\mathcal{H}$ will turn out to be isomorphic to the ring $\Lambda_{\mathbb{C}}$ of symmetric functions with $\mathbb{C}$ coefficients, via a composite φ of three maps

$$\Lambda_{\mathbb{C}} \longrightarrow A(GL)(\iota)_{\mathbb{C}} \longrightarrow A(GL)_{\mathbb{C}} \longrightarrow \mathcal{H}$$

in which the first map is the isomorphism from Proposition 4.7.2, the second is inclusion, and the third is the quotient map from Proposition 4.9.2.

[243] Here we use the following notation: Whenever P is a subset of a group G, we denote by $\underline{1}_P$ the map $G \to \mathbb{C}$ which sends every element of P to 1 and all remaining elements of G to 0. This is not in conflict with the notation $\underline{1}_G$ for the trivial character of G, since $\underline{1}_P = \underline{1}_G$ for $P = G$. Note that $\underline{1}_P$ is a class function when P is a union of conjugacy classes of G.

Theorem 4.9.5. *The above composite φ is a Hopf algebra isomorphism, sending*

$$\begin{array}{lll} h_n & \longmapsto & \sum_{\lambda \in \operatorname{Par}_n} \underline{1}_{J_\lambda}, \\ e_n & \longmapsto & q^{\binom{n}{2}} \underline{1}_{J_{(1^n)}}, \\ p_n & \longmapsto & \sum_{\lambda \in \operatorname{Par}_n} (q;q)_{\ell(\lambda)-1} \underline{1}_{J_\lambda} \qquad (\text{for } n > 0), \end{array}$$

where we are using the notation

$$(x;q)_m := (1-x)(1-qx)(1-q^2x)\cdots(1-q^{m-1}x) \qquad \text{for all } m \in \mathbb{N} \text{ and } x \text{ in any ring.}$$

Proof. That φ is a graded Hopf morphism follows because it is a composite of three such morphisms. We claim that once one shows the formula for the (nonzero) image $\varphi(p_n)$ given above is correct, then this will already show φ is an isomorphism, by the following argument. Note first that $\Lambda_{\mathbb{C}}$ and $\mathcal{H}$ both have dimension $|\operatorname{Par}_n|$ for their n-th homogeneous components, so it suffices to show that the graded map φ is injective. On the other hand, both $\Lambda_{\mathbb{C}}$ and $\mathcal{H}$ are (graded, connected, finite type) *self-dual* Hopf algebras (although with respect to a sesquilinear form), so Theorem 3.1.7 says that each is the symmetric algebra on its space of primitive elements. Thus it suffices to check that φ is injective when restricted to their subspaces of primitives.[244] For $\Lambda_{\mathbb{C}}$, by Corollary 3.1.8 the primitives are spanned by $\{p_1, p_2, \ldots\}$, with only one basis element in each degree $n \geq 1$. Hence φ is injective on the subspace of primitives if and only if it does not annihilate any p_n.

Thus it only remains to show the above formulas for the images of h_n, e_n, p_n under φ. This is clear for h_n, since Proposition 4.7.2 shows that it maps under the first two composites to the indicator function $\underline{1}_{GL_n}$ which then restricts to the sum of indicators $\sum_{\lambda \in \operatorname{Par}_n} \underline{1}_{J_\lambda}$ in $\mathcal{H}$. For e_n, p_n, we resort to generating functions. Let $\tilde{h}_n, \tilde{e}_n, \tilde{p}_n$ denote the three putative images in $\mathcal{H}$ of h_n, e_n, p_n, appearing on the right side in the theorem, and define generating functions

$$\tilde{H}(t) := \sum_{n \geq 0} \tilde{h}_n t^n, \quad \tilde{E}(t) := \sum_{n \geq 0} \tilde{e}_n t^n, \quad \tilde{P}(t) := \sum_{n \geq 0} \tilde{p}_{n+1} t^n \qquad \text{in } \mathcal{H}[[t]].$$

We wish to show that the map $\varphi[[t]] : \Lambda_{\mathbb{C}}[[t]] \to \mathcal{H}[[t]]$ (induced by φ) maps $H(t), E(t), P(t)$ in $\Lambda[[t]]$ to these three generating functions[245]. Since we have already shown this is correct for $H(t)$, by (2.4.3), (2.5.13), it suffices to check that in $\mathcal{H}[[t]]$ one has

$$\begin{array}{rcl} \tilde{H}(t)\tilde{E}(-t) = 1, & \text{or equivalently,} & \sum_{k=0}^{n} (-1)^k \tilde{e}_k \tilde{h}_{n-k} = \delta_{0,n}; \\ \tilde{H}'(t)\tilde{E}(-t) = \tilde{P}(t), & \text{or equivalently,} & \sum_{k=0}^{n} (-1)^k (n-k) \tilde{e}_k \tilde{h}_{n-k} = \tilde{p}_n. \end{array}$$

Thus it would be helpful to evaluate the class function $\tilde{e}_k \tilde{h}_{n-k}$. Note that a unipotent g in GL_n having ℓ Jordan blocks has an ℓ-dimensional 1-eigenspace, so that the number of k-dimensional g-stable $\mathbb{F}_q$-subspaces of $\mathbb{F}_q^n$ on which g has Jordan type (1^k) (that is, on which g acts as the identity) is the *q-binomial coefficient*

$$\begin{bmatrix} \ell \\ k \end{bmatrix}_q = \frac{(q;q)_\ell}{(q;q)_k (q;q)_{\ell-k}},$$

counting k-dimensional $\mathbb{F}_q$-subspaces V of an ℓ-dimensional $\mathbb{F}_q$-vector space; see, e.g., [206, §1.7]. Hence, for a unipotent g in GL_n having ℓ Jordan blocks, we have

$$(\tilde{e}_k \tilde{h}_{n-k})(g) = q^{\binom{k}{2}} \cdot \left(\underline{1}_{J_{(1^k)}} \cdot \tilde{h}_{n-k} \right)(g) = q^{\binom{k}{2}} \cdot \sum_{\nu \in \operatorname{Par}_{n-k}} \left(\underline{1}_{J_{(1^k)}} \cdot \underline{1}_{J_\nu} \right)(g) = q^{\binom{k}{2}} \begin{bmatrix} \ell \\ k \end{bmatrix}_q$$

(by Proposition 4.9.4). Thus one needs for $\ell \geq 1$ that

$$\sum_{k=0}^{\ell} (-1)^k q^{\binom{k}{2}} \begin{bmatrix} \ell \\ k \end{bmatrix}_q = 0, \tag{4.9.3}$$

$$\sum_{k=0}^{\ell} (-1)^k (n-k) q^{\binom{k}{2}} \begin{bmatrix} \ell \\ k \end{bmatrix}_q = (q;q)_{\ell-1}. \tag{4.9.4}$$

[244] An alternative way to see that it suffices to check this is by recalling Exercise 1.4.35(c).

[245] See (2.4.1), (2.4.2), (2.5.13) for the definitions of $H(t), E(t), P(t)$.

Identity (4.9.3) comes from setting $x = 1$ in the *q-binomial theorem* [206, Exer. 3.119]:

$$\sum_{k=0}^{\ell} (-1)^k q^{\binom{k}{2}} \begin{bmatrix} \ell \\ k \end{bmatrix}_q x^{\ell-k} = (x-1)(x-q)(x-q^2)\cdots(x-q^{\ell-1}). \tag{4.9.5}$$

Identity (4.9.4) comes from applying $\frac{d}{dx}$ to (4.9.5), then setting $x = 1$, and finally adding $(n-\ell)$ times (4.9.3). □

Exercise 4.9.6. Fix a prime power q. For any $k \in \mathbb{N}$, and any k partitions $\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(k)}$, we define a family $\left(g^{\lambda}_{\lambda^{(1)},\lambda^{(2)},\ldots,\lambda^{(k)}}(q)\right)_{\lambda\in\mathrm{Par}}$ of elements of $\mathbb{C}$ by the equation

$$\underline{1}_{J_{\lambda^{(1)}}} \underline{1}_{J_{\lambda^{(2)}}} \cdots \underline{1}_{J_{\lambda^{(k)}}} = \sum_{\lambda\in\mathrm{Par}} g^{\lambda}_{\lambda^{(1)},\lambda^{(2)},\ldots,\lambda^{(k)}}(q)\, \underline{1}_{J_\lambda}$$

in $\mathcal{H}$. This notation generalizes the notation $g^{\lambda}_{\mu,\nu}(q)$ we introduced in Definition 4.9.3. Note that $g^{\lambda}_{\mu}(q) = \delta_{\lambda,\mu}$ for any two partitions λ and μ, and that $g^{\lambda}(q) = \delta_{\lambda,\varnothing}$ for any partition λ (where $g^{\lambda}(q)$ is to be understood as $g^{\lambda}_{\lambda^{(1)},\lambda^{(2)},\ldots,\lambda^{(k)}}(q)$ for $k = 0$).

(a) Let $\lambda \in \mathrm{Par}$, and let $n = |\lambda|$. Let V be an n-dimensional $\mathbb{F}_q$-vector space, and let g be a unipotent endomorphism of V having Jordan type λ. Let $k \in \mathbb{N}$, and let $\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(k)}$ be k partitions. A $\left(\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(k)}\right)$*-compatible g-flag* will mean a sequence $0 = V_0 \subset V_1 \subset V_2 \subset \cdots \subset V_k = V$ of g-invariant $\mathbb{F}_q$-vector subspaces V_i of V such that for every $i \in \{1, 2, \ldots, k\}$, the endomorphism of V_i/V_{i-1} induced by g [246] has Jordan type $\lambda^{(i)}$.

Show that $g^{\lambda}_{\lambda^{(1)},\lambda^{(2)},\ldots,\lambda^{(k)}}(q)$ is the number of $\left(\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(k)}\right)$-compatible g-flags.[247]

(b) Let $\lambda \in \mathrm{Par}$. Let $k \in \mathbb{N}$, and let $\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(k)}$ be k partitions. Show that $g^{\lambda}_{\lambda^{(1)},\lambda^{(2)},\ldots,\lambda^{(k)}}(q) = 0$ unless $\left|\lambda^{(1)}\right| + \left|\lambda^{(2)}\right| + \cdots + \left|\lambda^{(k)}\right| = |\lambda|$ and $\lambda^{(1)} + \lambda^{(2)} + \cdots + \lambda^{(k)} \triangleright \lambda$. (Here and in the following, we are using the notations of Exercise 2.9.17).

(c) Let $\lambda \in \mathrm{Par}$, and let us write the transpose partition λ^t as $\lambda^t = \left((\lambda^t)_1, (\lambda^t)_2, \ldots, (\lambda^t)_\ell\right)$. Show that $g^{\lambda}_{\left(1^{(\lambda^t)_1}\right),\left(1^{(\lambda^t)_2}\right),\ldots,\left(1^{(\lambda^t)_\ell}\right)}(q) \neq 0$.

(d) Let $n \in \mathbb{N}$ and $\lambda \in \mathrm{Par}_n$. Show that

$$\varphi(e_\lambda) = \sum_{\mu\in\mathrm{Par}_n;\ \lambda^t \triangleright \mu} \alpha_{\lambda,\mu} \underline{1}_{J_\mu}$$

for some coefficients $\alpha_{\lambda,\mu} \in \mathbb{C}$ satisfying $\alpha_{\lambda,\lambda^t} \neq 0$.

(e) Give another proof of the fact that the map φ is injective.

[**Hint:** For (b), use Exercise 2.9.22(b).]

We next indicate, without proof, how $\mathcal{H}$ relates to the classical Hall algebra.

Definition 4.9.7. Let p be a prime. The usual *Hall algebra*, or what Schiffmann [190, §2.3] calls *Steinitz's classical Hall algebra* (see also Macdonald [142, Chap. II]), has $\mathbb{Z}$-basis elements $\{u_\lambda\}_{\lambda\in\mathrm{Par}}$, with the multiplicative structure constants $g^{\lambda}_{\mu,\nu}(p)$ in

$$u_\mu u_\nu = \sum_{\lambda} g^{\lambda}_{\mu,\nu}(p)\ u_\lambda$$

defined as follows: fix a finite abelian p-group L of *type* λ, meaning that

$$L \cong \bigoplus_{i=1}^{\ell(\lambda)} \mathbb{Z}/p^{\lambda_i}\mathbb{Z},$$

[246] This is well-defined. In fact, both V_i and V_{i-1} are g-invariant, so that g restricts to an endomorphism of V_i, which further restricts to an endomorphism of V_{i-1}, and thus gives rise to an endomorphism of V_i/V_{i-1}.

[247] This can be seen as a generalization of Proposition 4.9.4. In fact, if μ and ν are two partitions, then a (μ,ν)-compatible g-flag is a sequence $0 = V_0 \subset V_1 \subset V_2 = V$ of g-invariant $\mathbb{F}_q$-vector subspaces V_i of V such that the endomorphism of $V_1/V_0 \cong V_1$ induced by g has Jordan type μ, and the endomorphism of $V_2/V_1 \cong V/V_1$ induced by g has Jordan type ν. Choosing such a sequence amounts to choosing V_1 (since there is only one choice for each of V_0 and V_2), and the conditions on this V_1 are precisely the conditions on V in Proposition 4.9.4.

and let $g^{\lambda}_{\mu,\nu}(p)$ be the number of subgroups M of L of type μ, for which the quotient $N := L/M$ is of type ν. In other words, $g^{\lambda}_{\mu,\nu}(p)$ counts, for a fixed abelian p-group L of type λ, the number of short exact sequences $0 \to M \to L \to N \to 0$ in which M, N have types μ, ν, respectively (modulo isomorphism of short exact sequences restricting to the identity on L).

We claim that when one takes the finite field $\mathbb{F}_q$ of order $q = p$ a *prime*, the $\mathbb{Z}$-linear map

$$u_\lambda \longmapsto \underline{1}_{J_\lambda} \tag{4.9.6}$$

gives an isomorphism from this classical Hall algebra to the $\mathbb{Z}$-algebra $\mathcal{H}_{\mathbb{Z}} \subset \mathcal{H}$. The key point is *Hall's Theorem*, a non-obvious statement for which Macdonald includes two proofs in [142, Chap. II], one of them due to Zelevinsky[248]. To state it, we first recall some notions about discrete valuation rings.

Definition 4.9.8. A *discrete valuation ring* (short *DVR*) $\mathfrak{o}$ is a principal ideal domain having only one maximal ideal $\mathfrak{m} \neq 0$, with quotient $k = \mathfrak{o}/\mathfrak{m}$ called its *residue field*.

The structure theorem for finitely generated modules over a PID implies that an $\mathfrak{o}$-module L with finite composition series of composition length n must have $L \cong \bigoplus_{i=1}^{\ell(\lambda)} \mathfrak{o}/\mathfrak{m}^{\lambda_i}$ for some partition λ of n; say L has *type* λ in this situation.

Here are the two crucial examples for us.

Example 4.9.9. For any field $\mathbb{F}$, the power series ring $\mathfrak{o} = \mathbb{F}[[t]]$ is a DVR with maximal ideal $\mathfrak{m} = (t)$ and residue field $k = \mathfrak{o}/\mathfrak{m} = \mathbb{F}[[t]]/(t) \cong \mathbb{F}$. An $\mathfrak{o}$-module L of type λ is an $\mathbb{F}$-vector space together with an $\mathbb{F}$-linear transformation $T \in \operatorname{End} L$ that acts on L nilpotently (so that $g := T + 1$ acts unipotently, where $1 = \mathrm{id}_L$) with Jordan blocks of sizes given by λ: each summand $\mathfrak{o}/\mathfrak{m}^{\lambda_i} = \mathbb{F}[[t]]/(t^{\lambda_i})$ of L has an $\mathbb{F}$-basis $\{1, t, t^2, \ldots, t^{\lambda_i - 1}\}$ on which the map T that multiplies by t acts as a nilpotent Jordan block of size λ_i. Note also that, in this setting, $\mathfrak{o}$-submodules are the same as T-stable (or g-stable) $\mathbb{F}$-subspaces.

Example 4.9.10. The ring of p-adic integers $\mathfrak{o} = \mathbb{Z}_p$ is a DVR with maximal ideal $\mathfrak{m} = (p)$ and residue field $k = \mathfrak{o}/\mathfrak{m} = \mathbb{Z}_p/p\mathbb{Z}_p \cong \mathbb{Z}/p\mathbb{Z}$. An $\mathfrak{o}$-module L of type λ is an abelian p-group of type λ: for each summand, $\mathfrak{o}/\mathfrak{m}^{\lambda_i} = \mathbb{Z}_p/p^{\lambda_i}\mathbb{Z}_p \cong \mathbb{Z}/p^{\lambda_i}\mathbb{Z}$. Note also that, in this setting, $\mathfrak{o}$-submodules are the same as subgroups.

One last notation: $n(\lambda) := \sum_{i \geq 1} (i-1)\lambda_i$, for λ in Par. Hall's Theorem is as follows.

Theorem 4.9.11. *Assume $\mathfrak{o}$ is a DVR with maximal ideal $\mathfrak{m}$, and that its residue field $k = \mathfrak{o}/\mathfrak{m}$ is finite of cardinality q. Fix an $\mathfrak{o}$-module L of type λ. Then the number of $\mathfrak{o}$-submodules M of type μ for which the quotient $N = L/M$ is of type ν can be written as the specialization*

$$[g^{\lambda}_{\mu,\nu}(t)]_{t=q}$$

of a polynomial $g^{\lambda}_{\mu,\nu}(t)$ in $\mathbb{Z}[t]$, called the Hall polynomial.

Furthermore, the Hall polynomial $g^{\lambda}_{\mu,\nu}(t)$ has degree at most $n(\lambda) - (n(\mu) + n(\nu))$, and its coefficient of $t^{n(\lambda)-(n(\mu)+n(\nu))}$ is the Littlewood-Richardson coefficient $c^{\lambda}_{\mu,\nu}$.

Comparing what Hall's Theorem says in Examples 4.9.9 and 4.9.10, shows that the map (4.9.6) gives the desired isomorphism from the classical Hall algebra to $\mathcal{H}_{\mathbb{Z}}$.

We close this section with some remarks on the vast literature on Hall algebras that we will *not* discuss here.

Remark 4.9.12. Macdonald's version of Hall's Theorem [142, (4.3)] is stronger than Theorem 4.9.11, and useful for certain applications: he shows that $g^{\lambda}_{\mu,\nu}(t)$ is the zero polynomial whenever the Littlewood-Richardson coefficient $c^{\lambda}_{\mu,\nu}$ is zero.

Remark 4.9.13. In general, not all coefficients of the Hall polynomials $g^{\lambda}_{\mu,\nu}(t)$ are nonnegative (see Butler/Hales [32] for a study of when they are); it often happens that $g^{\lambda}_{\mu,\nu}(1) = 0$ despite $g^{\lambda}_{\mu,\nu}(t)$ not being the

[248] See also [190, Thm. 2.6, Prop. 2.7] for quick proofs of part of it, similar to Zelevinsky's. Another proof, based on a recent category-theoretical paradigm, can be found in [61, Theorem 3.53].

zero polynomial[249]. However, in [110, Thm. 4.2], Klein showed that the polynomial values $g^{\lambda}_{\mu,\nu}(p)$ for p prime are always positive when $c^{\lambda}_{\mu,\nu} \neq 0$. (This easily yields the same result for p a prime power.)

Remark 4.9.14. Zelevinsky in [227, Chaps 10, 11] uses the isomorphism $\Lambda_{\mathbb{C}} \to \mathcal{H}$ to derive J. Green's formula for the value of any irreducible character χ of GL_n on any unipotent class J_λ. The answer involves values of irreducible characters of $\mathfrak{S}_n$ along with *Green's polynomials* $Q^{\lambda}_{\mu}(q)$ (see Macdonald [142, §III.7]; they are denoted $Q(\lambda, \mu)$ by Zelevinsky), which express the images under the isomorphism of Theorem 4.9.5 of the symmetric function basis $\{p_\mu\}$ in terms of the basis $\{\underline{1}_{J_\lambda}\}$.

Remark 4.9.15. The Hall polynomials $g^{\lambda}_{\mu,\nu}(t)$ also essentially give the multiplicative structure constants for $\Lambda(\mathbf{x})[t]$ with respect to its basis of *Hall-Littlewood symmetric functions* $P_\lambda = P_\lambda(\mathbf{x}; t)$:

$$P_\mu P_\nu = \sum_{\lambda} t^{n(\lambda)-(n(\mu)+n(\nu))} g^{\lambda}_{\mu,\nu}(t^{-1}) P_\lambda.$$

See Macdonald [142, §III.3].

Remark 4.9.16. Schiffmann [190] discusses self-dual Hopf algebras which vastly generalize the classical Hall algebra called *Ringel-Hall algebras*, associated to abelian categories which are hereditary. Examples come from categories of nilpotent representations of quivers; the quiver having exactly one node and one arc recovers the classical Hall algebra $\mathcal{H}_{\mathbb{Z}}$ discussed above.

Remark 4.9.17. The general linear groups $GL_n(\mathbb{F}_q)$ are one of four families of so-called *classical groups*. Progress has been made on extending Zelevinsky's PSH theory to the other families:

(a) Work of Thiem and Vinroot [217] shows that the tower $\{G_*\}$ of *finite unitary groups* $U_n(\mathbb{F}_{q^2})$ give rise to another positive self-dual Hopf algebra $A = \bigoplus_{n \geq 0} R(U_n(\mathbb{F}_{q^2}))$, in which the role of Harish-Chandra induction is played by *Deligne-Lusztig induction*. In this theory, character and degree formulas for $U_n(\mathbb{F}_{q^2})$ are related to those of $GL_n(\mathbb{F}_q)$ by substituting $q \mapsto -q$, along with appropriate scalings by ±1, a phenomenon sometimes called *Ennola duality*. See also [207, §4].

(b) van Leeuwen [128] has studied $\bigoplus_{n\geq 0} R\left(Sp_{2n}\left(\mathbb{F}_q\right)\right)$, $\bigoplus_{n\geq 0} R\left(O_{2n}\left(\mathbb{F}_q\right)\right)$ and $\bigoplus_{n\geq 0} R\left(U_n\left(\mathbb{F}_{q^2}\right)\right)$ not as Hopf algebras, but rather as so-called *twisted PSH-modules* over the PSH $A(GL)$ (a "deformed" version of the older notion of Hopf modules). He classified these PSH-modules axiomatically similarly to Zelevinsky's above classification of PSH's.

(c) In a recent honors thesis [201], Shelley-Abrahamson defined yet another variation of the concept of Hopf modules, named 2-*compatible Hopf modules*, and identified $\bigoplus_{n\geq 0} R\left(Sp_{2n}\left(\mathbb{F}_q\right)\right)$ and $\bigoplus_{n\geq 0} R\left(O_{2n+1}\left(\mathbb{F}_q\right)\right)$ as such modules over $A(GL)$.

[249] Actually, Butler/Hales show in [32, proof of Prop. 2.4] that the values $g^{\lambda}_{\mu,\nu}(1)$ are the structure constants of the ring Λ with respect to its basis $(m_\lambda)_{\lambda\in\mathrm{Par}}$: we have

$$m_\mu m_\nu = \sum_{\lambda\in\mathrm{Par}} g^{\lambda}_{\mu,\nu}(1) m_\lambda$$

for all partitions μ and ν.

5. Quasisymmetric functions and P-partitions

We discuss here our next important example of a Hopf algebra arising in combinatorics: the *quasisymmetric functions* of Gessel [79], with roots in work of Stanley [203] on P-partitions. Other treatments of quasisymmetric functions can be found in [206, Section 7.19] and [187, Chapter 8] (with focus on their enumerative applications rather than on their Hopf structure) and in [153, Chapter 6] (with a focus on their representation-theoretical meaning). Quasisymmetric functions have found applications in combinatorial enumeration ([187, Chapter 8], [206, Section 7.19]), topology ([12]) and algebraic geometry ([158], [163]).

5.1. **Definitions, and Hopf structure.** The definitions of quasisymmetric functions require a totally ordered variable set. Usually we will use a variable set denoted $\mathbf{x} = (x_1, x_2, \ldots)$ with the usual ordering $x_1 < x_2 < \cdots$. However, it is good to have some flexibility in changing the ordering, which is why we make the following definition.

Definition 5.1.1. Given any totally ordered set I, create a totally ordered variable set $\{x_i\}_{i\in I}$, and then let $R(\{x_i\}_{i\in I})$ denote the power series of bounded degree in $\{x_i\}_{i\in I}$ having coefficients in $\mathbf{k}$.

The *ring of quasisymmetric functions* $\operatorname{QSym}(\{x_i\}_{i\in I})$ *over the alphabet* $\{x_i\}_{i\in I}$ will be the $\mathbf{k}$-submodule consisting of the elements f in $R(\{x_i\}_{i\in I})$ that have the same coefficient on the monomials $x_{i_1}^{\alpha_1}\cdots x_{i_\ell}^{\alpha_\ell}$ and $x_{j_1}^{\alpha_1}\cdots x_{j_\ell}^{\alpha_\ell}$ whenever both $i_1 < \cdots < i_\ell$ and $j_1 < \cdots < j_\ell$ in the total order on I. We write $\operatorname{QSym}_{\mathbf{k}}(\{x_i\}_{i\in I})$ instead of $\operatorname{QSym}(\{x_i\}_{i\in I})$ to stress the choice of base ring $\mathbf{k}$.

It immediately follows from this definition that $\operatorname{QSym}(\{x_i\}_{i\in I})$ is a free $\mathbf{k}$-submodule of $R(\{x_i\}_{i\in I})$, having as $\mathbf{k}$-basis elements the *monomial quasisymmetric functions*

$$M_\alpha(\{x_i\}_{i\in I}) := \sum_{i_1<\cdots<i_\ell \text{ in } I} x_{i_1}^{\alpha_1}\cdots x_{i_\ell}^{\alpha_\ell}$$

for all compositions[250] α satisfying $\ell(\alpha) \leq |I|$. When I is infinite, this means that the M_α for all compositions α form a basis of $\operatorname{QSym}(\{x_i\}_{i\in I})$.

Note that $\operatorname{QSym}(\{x_i\}_{i\in I}) = \bigoplus_{n\geq 0} \operatorname{QSym}_n(\{x_i\}_{i\in I})$ is a graded $\mathbf{k}$-module of finite type, where $\operatorname{QSym}_n(\{x_i\}_{i\in I})$ is the $\mathbf{k}$-submodule of quasisymmetric functions which are homogeneous of degree n. Letting Comp denote the set of all compositions α, and Comp_n the compositions α of n (that is, compositions whose parts sum to n), the subset $\{M_\alpha\}_{\alpha\in\operatorname{Comp}_n;\ \ell(\alpha)\leq|I|}$ gives a $\mathbf{k}$-basis for $\operatorname{QSym}_n(\{x_i\}_{i\in I})$.

Example 5.1.2. Taking the variable set $\mathbf{x} = (x_1 < x_2 < \cdots)$ to define $\operatorname{QSym}(\mathbf{x})$, for $n = 0, 1, 2, 3$, one has these basis elements in $\operatorname{QSym}_n(\mathbf{x})$:

$$\begin{aligned}
M_{()} = M_\varnothing &= 1, && \\
\\
M_{(1)} &= x_1 + x_2 + x_3 + \cdots && = m_{(1)} = s_{(1)} = e_1 = h_1 = p_1, \\
\\
M_{(2)} &= x_1^2 + x_2^2 + x_3^2 + \cdots && = m_{(2)} = p_2, \\
M_{(1,1)} &= x_1x_2 + x_1x_3 + x_2x_3 + \cdots && = m_{(1,1)} = e_2, \\
\\
M_{(3)} &= x_1^3 + x_2^3 + x_3^3 + \cdots && = m_{(3)} = p_3, \\
M_{(2,1)} &= x_1^2x_2 + x_1^2x_3 + x_2^2x_3 + \cdots, && \\
M_{(1,2)} &= x_1x_2^2 + x_1x_3^2 + x_2x_3^2 + \cdots, && \\
M_{(1,1,1)} &= x_1x_2x_3 + x_1x_2x_4 + x_1x_3x_4 + \cdots && = m_{(1,1,1)} = e_3.
\end{aligned}$$

It is not obvious that $\operatorname{QSym}(\mathbf{x})$ is a subalgebra of $R(\mathbf{x})$, but we will show this momentarily. For example,

$$\begin{aligned}
M_{(a)}M_{(b,c)} &= (x_1^a + x_2^a + x_3^a + \cdots)(x_1^bx_2^c + x_1^bx_3^c + x_2^bx_3^c + \cdots) \\
&= x_1^{a+b}x_2^c + \cdots + x_1^bx_3^{a+c} + \cdots + x_1^ax_2^bx_3^c + \cdots + x_1^bx_2^ax_3^c + \cdots + x_1^bx_2^cx_3^a + \cdots \\
&= M_{(a+b,c)} + M_{(b,a+c)} + M_{(a,b,c)} + M_{(b,a,c)} + M_{(b,c,a)}.
\end{aligned}$$

Proposition 5.1.3. *For any infinite totally ordered set I, one has that $\operatorname{QSym}(\{x_i\}_{i\in I})$ is a $\mathbf{k}$-subalgebra of $R(\{x_i\}_{i\in I})$, with multiplication in the $\{M_\alpha\}$-basis as follows: Fix three disjoint chain posets $(i_1 < \cdots < i_\ell)$,*

[250] Recall that compositions were defined in Definition 4.3.1, along with related concepts such as length and size.

$(j_1 < \cdots < j_m)$ and $(k_1 < k_2 < \cdots)$. Now, if $\alpha = (\alpha_1, \ldots, \alpha_\ell)$ and $\beta = (\beta_1, \ldots, \beta_m)$ are two compositions, then

$$M_\alpha M_\beta = \sum_f M_{\mathrm{wt}(f)} \tag{5.1.1}$$

in which the sum is over all $p \in \mathbb{N}$ and all maps f from the disjoint union of two chains to a chain

$$(i_1 < \cdots < i_\ell) \sqcup (j_1 < \cdots < j_m) \xrightarrow{f} (k_1 < \cdots < k_p) \tag{5.1.2}$$

which are both surjective and strictly order-preserving (that is, if x and y are two elements in the domain satisfying $x < y$, then $f(x) < f(y)$), and where the composition $\mathrm{wt}(f) := (\mathrm{wt}_1(f), \ldots, \mathrm{wt}_p(f))$ is defined by $\mathrm{wt}_s(f) := \sum_{i_u \in f^{-1}(k_s)} \alpha_u + \sum_{j_v \in f^{-1}(k_s)} \beta_v$.

Example 5.1.4. For this example, set $\alpha = (2,1)$ and $\beta = (3,4,2)$. Let us compute $M_\alpha M_\beta$ using (5.1.1). Indeed, the length of α is $\ell = 2$, and the length of β is $m = 3$, so the sum on the right hand side of (5.1.1) is a sum over all $p \in \mathbb{N}$ and all surjective strictly order-preserving maps f from the disjoint union $(i_1 < i_2) \sqcup (j_1 < j_2 < j_3)$ of two chains to the chain $(k_1 < k_2 < \cdots < k_p)$. Such maps can exist only when $p \leq 5$ (due to having to be surjective) and only for $p \geq 3$ (since, being strictly order-preserving, they have to be injective when restricted to $(j_1 < j_2 < j_3)$). Hence, enumerating them is a finite problem. The reader can check that the value obtained for $M_\alpha M_\beta$ is

$$\begin{aligned}
&M_{(2,1,3,4,2)} + M_{(2,3,1,4,2)} + M_{(2,3,4,1,2)} + M_{(2,3,4,2,1)} + M_{(3,2,1,4,2)} \\
&+ M_{(3,2,4,1,2)} + M_{(3,2,4,2,1)} + M_{(3,4,2,1,2)} + M_{(3,4,2,2,1)} + M_{(3,4,2,2,1)} \\
&+ M_{(2,3,4,3)} + M_{(2,3,5,2)} + M_{(2,4,4,2)} + M_{(3,2,4,3)} + M_{(3,2,5,2)} + M_{(3,4,2,3)} \\
&+ M_{(3,4,4,1)} + M_{(3,6,1,2)} + M_{(3,6,2,1)} + M_{(5,1,4,2)} + M_{(5,4,1,2)} + M_{(5,4,2,1)} \\
&+ M_{(5,4,3)} + M_{(5,5,2)} + M_{(3,6,3)}.
\end{aligned}$$

Here, we have listed the addends corresponding to $p = 5$ on the first two rows, the addends corresponding to $p = 4$ on the next two rows, and those corresponding to $p = 3$ on the fifth row. The reader might notice that the first two rows (i.e., the addends with $p = 5$) are basically a list of shuffles of α and β: In general, the maps (5.1.2) for $p = \ell + m$ are in bijection with the elements of $\mathrm{Sh}_{\ell,m}$ [251], and the corresponding compositions $\mathrm{wt}(f)$ are the shuffles of α and β. Therefore the name "overlapping shuffle product".

Proof of Proposition 5.1.3. It clearly suffices to prove the formula (5.1.1). Let $\alpha = (\alpha_1, \ldots, \alpha_\ell)$ and $\beta = (\beta_1, \ldots, \beta_m)$ be two compositions. Fix three disjoint chain posets $(i_1 < \cdots < i_\ell)$, $(j_1 < \cdots < j_m)$ and $(k_1 < k_2 < \cdots)$.

Thus, multiplying $M_\alpha = \sum_{u_1 < \cdots < u_\ell} x_{u_1}^{\alpha_1} \cdots x_{u_\ell}^{\alpha_\ell}$ with $M_\beta = \sum_{v_1 < \cdots < v_m} x_{v_1}^{\beta_1} \cdots x_{v_m}^{\beta_m}$, we obtain

$$\begin{aligned}
M_\alpha M_\beta &= \sum_{u_1 < \cdots < u_\ell} \sum_{v_1 < \cdots < v_m} \left(x_{u_1}^{\alpha_1} \cdots x_{u_\ell}^{\alpha_\ell}\right) \left(x_{v_1}^{\beta_1} \cdots x_{v_m}^{\beta_m}\right) \\
&= \sum_{\gamma = (\gamma_1, \ldots, \gamma_p) \in \mathrm{Comp}} \sum_{w_1 < \cdots < w_p \text{ in } I} N^\gamma_{w_1, \ldots, w_p} x_{w_1}^{\gamma_1} \cdots x_{w_p}^{\gamma_p}, \qquad (5.1.3)
\end{aligned}$$

where $N^\gamma_{w_1, \ldots, w_p}$ is the number of all pairs

$$((u_1 < \cdots < u_\ell), (v_1 < \cdots < v_m)) \in I^\ell \times I^m \tag{5.1.4}$$

of two strictly increasing tuples satisfying

$$\left(x_{u_1}^{\alpha_1} \cdots x_{u_\ell}^{\alpha_\ell}\right) \left(x_{v_1}^{\beta_1} \cdots x_{v_m}^{\beta_m}\right) = x_{w_1}^{\gamma_1} \cdots x_{w_p}^{\gamma_p}. \tag{5.1.5}$$

[252] Thus, we need to show that $N^\gamma_{w_1, \ldots, w_p}$ (for a given $\gamma = (\gamma_1, \ldots, \gamma_p) \in \mathrm{Comp}$ and a given $(w_1 < \cdots < w_p) \in I^p$) is also the number of all surjective strictly order-preserving maps

$$(i_1 < \cdots < i_\ell) \sqcup (j_1 < \cdots < j_m) \xrightarrow{f} (k_1 < \cdots < k_p) \text{ satisfying } \mathrm{wt}(f) = \gamma \tag{5.1.6}$$

[251] The bijection takes a map f to the inverse of the permutation $\sigma \in \mathfrak{S}_p$ which sends every $x \in \{1, 2, \ldots, \ell\}$ to the index y satisfying $f(i_x) = k_y$, and sends every $x \in \{\ell+1, \ell+2, \ldots, \ell+m\}$ to the index y satisfying $f(j_{x-\ell}) = k_y$.

[252] In the second equality in (5.1.3), we have used the fact that each monomial can be uniquely written in the form $x_{w_1}^{\gamma_1} \cdots x_{w_p}^{\gamma_p}$ for some composition $\gamma = (\gamma_1, \ldots, \gamma_p) \in \mathrm{Comp}$ and some strictly increasing tuple $(w_1 < \cdots < w_p) \in I^p$.

(because then, (5.1.3) will simplify to (5.1.1)).

In order to show this, it suffices to construct a bijection from the set of all pairs (5.1.4) satisfying (5.1.5) to the set of all surjective strictly order-preserving maps (5.1.6). This bijection is easy to construct: Given a pair (5.1.4) satisfying (5.1.5), the bijection sends it to the map (5.1.6) determined by:

$$i_g \overset{f}{\mapsto} k_h, \text{ where } h \text{ is chosen such that } u_g = w_h;$$
$$j_g \overset{f}{\mapsto} k_h, \text{ where } h \text{ is chosen such that } v_g = w_h.$$

Proving that this bijection is well-defined and bijective is straightforward[253]. □

The multiplication rule (5.1.1) shows that the $\mathbf{k}$-algebra $\mathrm{QSym}(\{x_i\}_{i\in I})$ does not depend much on I, as long as I is infinite. More precisely, all such $\mathbf{k}$-algebras are mutually isomorphic. We can use this to define a $\mathbf{k}$-algebra of quasisymmetric functions without any reference to I:

Definition 5.1.5. Let QSym be the $\mathbf{k}$-algebra defined as having $\mathbf{k}$-basis $\{M_\alpha\}_{\alpha\in\mathrm{Comp}}$ and with multiplication defined $\mathbf{k}$-linearly by (5.1.1). This is called the *$\mathbf{k}$-algebra of quasisymmetric functions*. We write $\mathrm{QSym}_{\mathbf{k}}$ instead of QSym to stress the choice of base ring $\mathbf{k}$.

The $\mathbf{k}$-algebra QSym is graded, and its n-th graded component QSym_n has $\mathbf{k}$-basis $\{M_\alpha\}_{\alpha\in\mathrm{Comp}_n}$.

For every infinite totally ordered set I, the $\mathbf{k}$-algebra QSym is isomorphic to the $\mathbf{k}$-algebra $\mathrm{QSym}(\{x_i\}_{i\in I})$. The isomorphism sends $M_\alpha \longmapsto M_\alpha(\{x_i\}_{i\in I})$.

In particular, we obtain the isomorphism $\mathrm{QSym} \cong \mathrm{QSym}(\mathbf{x})$ for $\mathbf{x}$ being the infinite chain $(x_1 < x_2 < x_3 < \cdots)$. We will identify QSym with $\mathrm{QSym}(\mathbf{x})$ along this isomorphism. This allows us to regard quasisymmetric functions either as power series in a specific set of variables ("alphabet"), or as formal linear combinations of M_α's, whatever is more convenient.

For any infinite alphabet $\{x_i\}_{i\in I}$ and any $f \in \mathrm{QSym}$, we denote by $f\left(\{x_i\}_{i\in I}\right)$ the image of f under the algebra isomorphism $\mathrm{QSym} \to \mathrm{QSym}\left(\{x_i\}_{i\in I}\right)$ defined in Definition 5.1.5.

The comultiplication of QSym will extend the one that we defined for Λ, but we need to take care about the order of the variables this time. We consider the linear order from (2.3.2) on two sets of variables $(\mathbf{x},\mathbf{y}) = (x_1 < x_2 < \cdots < y_1 < y_2 < \cdots)$, and we embed the $\mathbf{k}$-algebra $\mathrm{QSym}(\mathbf{x}) \otimes \mathrm{QSym}(\mathbf{y})$ into the $\mathbf{k}$-algebra $R(\mathbf{x},\mathbf{y})$ by identifying every $f \otimes g \in \mathrm{QSym}(\mathbf{x}) \otimes \mathrm{QSym}(\mathbf{y})$ with $fg \in R(\mathbf{x},\mathbf{y})$ (this embedding is indeed injective[254]). It can then be seen that

$$\mathrm{QSym}(\mathbf{x},\mathbf{y}) \subset \mathrm{QSym}(\mathbf{x}) \otimes \mathrm{QSym}(\mathbf{y})$$

(where the right hand side is viewed as $\mathbf{k}$-subalgebra of $R(\mathbf{x},\mathbf{y})$ via said embedding)[255], so that one can define $\mathrm{QSym} \overset{\Delta}{\longrightarrow} \mathrm{QSym} \otimes \mathrm{QSym}$ as the composite of the maps in the bottom row here:

(5.1.7)

$$\begin{tikzcd}
 & R(\mathbf{x},\mathbf{y}) \arrow[r, phantom, "="] & R(\mathbf{x},\mathbf{y}) & \\
\mathrm{QSym} \arrow[r, phantom, "\cong"] & \mathrm{QSym}(\mathbf{x},\mathbf{y}) \arrow[u, phantom, "\cup"] \arrow[r, hook] & \mathrm{QSym}(\mathbf{x}) \otimes \mathrm{QSym}(\mathbf{y}) \arrow[u, phantom, "\cup"] \arrow[r, phantom, "\cong"] & \mathrm{QSym} \otimes \mathrm{QSym}, \\
f \arrow[r, mapsto] & f(\mathbf{x},\mathbf{y}) = f(x_1, x_2, \ldots, y_1, y_2, \ldots). & &
\end{tikzcd}$$

(Recall that $f(\mathbf{x},\mathbf{y})$ is formally defined as the image of f under the algebra isomorphism $\mathrm{QSym} \to \mathrm{QSym}(\mathbf{x},\mathbf{y})$ defined in Definition 5.1.5.)

[253] The inverse of this bijection sends each map (5.1.6) to the pair (5.1.4) determined by

$$u_g = w_h, \text{ where } h \text{ is chosen such that } f(i_g) = k_h;$$
$$v_g = w_h, \text{ where } h \text{ is chosen such that } f(j_g) = k_h.$$

[254] This is because it sends the basis elements $M_\beta(\mathbf{x}) \otimes M_\gamma(\mathbf{y})$ of the former $\mathbf{k}$-algebra to the linearly independent power series $M_\beta(\mathbf{x}) M_\gamma(\mathbf{y})$.

[255] This is not completely obvious, but can be easily checked by verifying that $M_\alpha(\mathbf{x},\mathbf{y}) = \sum_{\substack{(\beta,\gamma): \\ \beta\cdot\gamma=\alpha}} M_\beta(\mathbf{x}) \otimes M_\gamma(\mathbf{y})$ for every composition α (see the proof of Proposition 5.1.7 for why this holds).

Example 5.1.6. For example,

$$\begin{aligned}\Delta M_{(a,b,c)} &= M_{(a,b,c)}(x_1, x_2, \ldots, y_1, y_2, \ldots)\\ &= x_1^a x_2^b x_3^c + x_1^a x_2^b x_4^c + \cdots\\ &\quad + x_1^a x_2^b \cdot y_1^c + x_1^a x_2^b \cdot y_2^c + \cdots\\ &\quad + x_1^a \cdot y_1^b y_2^c + x_1^a \cdot y_1^b y_3^c + \cdots\\ &\quad + y_1^a y_2^b y_3^c + y_1^a y_2^b y_4^c + \cdots\\ &= M_{(a,b,c)}(\mathbf{x}) + M_{(a,b)}(\mathbf{x}) M_{(c)}(\mathbf{y}) + M_{(a)}(\mathbf{x}) M_{(b,c)}(\mathbf{y}) + M_{(a,b,c)}(\mathbf{y})\\ &= M_{(a,b,c)} \otimes 1 + M_{(a,b)} \otimes M_{(c)} + M_{(a)} \otimes M_{(b,c)} + 1 \otimes M_{(a,b,c)}.\end{aligned}$$

Defining the *concatenation* $\beta \cdot \gamma$ of two compositions $\beta = (\beta_1, \ldots, \beta_r), \gamma = (\gamma_1, \ldots, \gamma_s)$ to be the composition $(\beta_1, \ldots, \beta_r, \gamma_1, \ldots, \gamma_s)$, one has the following description of the coproduct in the $\{M_\alpha\}$ basis.

Proposition 5.1.7. *For a composition $\alpha = (\alpha_1, \ldots, \alpha_\ell)$, one has*

$$\Delta M_\alpha = \sum_{k=0}^{\ell} M_{(\alpha_1,\ldots,\alpha_k)} \otimes M_{(\alpha_{k+1},\ldots,\alpha_\ell)} = \sum_{\substack{(\beta,\gamma):\\ \beta\cdot\gamma=\alpha}} M_\beta \otimes M_\gamma.$$

Proof. We work with the infinite totally ordered set $I = \{1 < 2 < 3 < \cdots\}$. The definition of Δ yields

$$\Delta M_\alpha = M_\alpha(\mathbf{x}, \mathbf{y}) = \sum_{p_1 < p_2 < \cdots < p_\ell \text{ in } (\mathbf{x},\mathbf{y})} p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_\ell^{\alpha_\ell}, \tag{5.1.8}$$

where the sum runs over strictly increasing ℓ-tuples $(p_1 < p_2 < \cdots < p_\ell)$ of variables in the variable set $(\mathbf{x}, \mathbf{y})$. But every such ℓ-tuple $(p_1 < p_2 < \cdots < p_\ell)$ can be expressed uniquely in the form $\left(x_{i_1}, \ldots, x_{i_k}, y_{j_1}, \ldots, y_{j_{\ell-k}}\right)$ for some $k \in \{0, 1, \ldots, \ell\}$ and some subscripts $i_1 < \cdots < i_k$ and $j_1 < \cdots < j_{\ell-k}$ in I. The corresponding monomial $p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_\ell^{\alpha_\ell}$ then rewrites as $x_{i_1}^{\alpha_1} \cdots x_{i_k}^{\alpha_k} \cdot y_{j_1}^{\alpha_{k+1}} \cdots y_{j_{\ell-k}}^{\alpha_\ell}$. Thus, the sum on the right hand side of (5.1.8) rewrites as

$$\begin{aligned}&\sum_{k=0}^{\ell} \sum_{i_1<\cdots<i_k} \sum_{j_1<\cdots<j_{\ell-k}} x_{i_1}^{\alpha_1} \cdots x_{i_k}^{\alpha_k} \cdot y_{j_1}^{\alpha_{k+1}} \cdots y_{j_{\ell-k}}^{\alpha_\ell}\\ &= \sum_{k=0}^{\ell} \underbrace{\left(\sum_{i_1<\cdots<i_k} x_{i_1}^{\alpha_1} \cdots x_{i_k}^{\alpha_k}\right)}_{=M_{(\alpha_1,\ldots,\alpha_k)}(\mathbf{x})} \cdot \underbrace{\left(\sum_{j_1<\cdots<j_{\ell-k}} y_{j_1}^{\alpha_{k+1}} \cdots y_{j_{\ell-k}}^{\alpha_\ell}\right)}_{=M_{(\alpha_{k+1},\ldots,\alpha_\ell)}(\mathbf{y})}\\ &= \sum_{k=0}^{\ell} M_{(\alpha_1,\ldots,\alpha_k)}(\mathbf{x}) M_{(\alpha_{k+1},\ldots,\alpha_\ell)}(\mathbf{y}).\end{aligned}$$

Thus, (5.1.8) becomes

$$\begin{aligned}\Delta M_\alpha = \sum_{p_1<p_2<\cdots<p_\ell \text{ in } (\mathbf{x},\mathbf{y})} p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_\ell^{\alpha_\ell} &= \sum_{k=0}^{\ell} M_{(\alpha_1,\ldots,\alpha_k)}(\mathbf{x}) M_{(\alpha_{k+1},\ldots,\alpha_\ell)}(\mathbf{y})\\ &= \sum_{k=0}^{\ell} M_{(\alpha_1,\ldots,\alpha_k)} \otimes M_{(\alpha_{k+1},\ldots,\alpha_\ell)} = \sum_{\substack{(\beta,\gamma):\\ \beta\cdot\gamma=\alpha}} M_\beta \otimes M_\gamma.\end{aligned}$$

□

Proposition 5.1.8. *The quasisymmetric functions* QSym *form a connected graded Hopf algebra of finite type, which is commutative, and contains the symmetric functions Λ as a Hopf subalgebra.*

Proof. To prove coassociativity of Δ, we need to be slightly careful. It seems reasonable to argue by $(\Delta \otimes \mathrm{id}) \circ \Delta f = f(\mathbf{x}, \mathbf{y}, \mathbf{z}) = (\mathrm{id} \otimes \Delta) \circ \Delta f$ as in the case of Λ, but this would now require further justification, as terms like $f(\mathbf{x},\mathbf{y})$ and $f(\mathbf{x},\mathbf{y},\mathbf{z})$ are no longer directly defined as evaluations of f on some sequences (but rather are defined as images of f under certain homomorphisms). However, it is very easy to see that Δ is coassociative by checking $(\Delta \otimes \mathrm{id}) \circ \Delta = (\mathrm{id} \otimes \Delta) \circ \Delta$ on the $\{M_\alpha\}$ basis: Proposition 5.1.7 yields

$$\begin{aligned}((\Delta \otimes \mathrm{id}) \circ \Delta)\, M_\alpha &= \sum_{k=0}^{\ell} \Delta(M_{(\alpha_1,\ldots,\alpha_k)}) \otimes M_{(\alpha_{k+1},\ldots,\alpha_\ell)} \\ &= \sum_{k=0}^{\ell} \left(\sum_{i=0}^{k} M_{(\alpha_1,\ldots,\alpha_i)} \otimes M_{(\alpha_{i+1},\ldots,\alpha_k)} \right) \otimes M_{(\alpha_{k+1},\ldots,\alpha_\ell)} \\ &= \sum_{k=0}^{\ell} \sum_{i=0}^{k} M_{(\alpha_1,\ldots,\alpha_i)} \otimes M_{(\alpha_{i+1},\ldots,\alpha_k)} \otimes M_{(\alpha_{k+1},\ldots,\alpha_\ell)}\end{aligned}$$

and the same expression for $((\mathrm{id} \otimes \Delta) \circ \Delta)\, M_\alpha$.

The coproduct Δ of QSym is an algebra morphism because it is defined as a composite of algebra morphisms in the bottom row of (5.1.7). To prove that the restriction of Δ to the subring Λ of QSym is the comultiplication of Λ, it thus is enough to check that it sends the elementary symmetric function e_n to $\sum_{i=0}^{n} e_i \otimes e_{n-i}$ for every $n \in \mathbb{N}$. This again follows from Proposition 5.1.7, since $e_n = M_{(1,1,\ldots,1)}$ (with n times 1).

The counit is as usual for a connected graded coalgebra, and just as in the case of Λ, sends a quasisymmetric function $f(\mathbf{x})$ to its constant term $f(0,0,\ldots)$. This is an evaluation, and hence an algebra morphism. Hence QSym forms a bialgebra, and as it is graded and connected, also a Hopf algebra by Proposition 1.4.16. It is clearly of finite type and contains Λ as a Hopf subalgebra. □

We will identify the antipode in QSym shortly, but we first deal with another slightly subtle issue. In addition to the counit evaluation $\epsilon(f) = f(0,0,\ldots)$, starting in Section 7.1, we will want to specialize elements in $\mathrm{QSym}(\mathbf{x})$ by making other variable substitutions, in which all but a finite list of variables are set to zero. We justify this here.

Proposition 5.1.9. *Fix a totally ordered set I, a commutative $\mathbf{k}$-algebra A, a finite list of variables $x_{i_1}, \ldots, x_{i_m}$, say with $i_1 < \cdots < i_m$ in I, and an ordered list of elements $(a_1,\ldots,a_m) \in A^m$.*

Then there is a well-defined evaluation homomorphism

$$\begin{array}{ccc} \mathrm{QSym}(\{x_i\}_{i\in I}) & \longrightarrow & A, \\ f & \longmapsto & [f]_{\substack{x_{i_1}=a_1,\ldots,x_{i_m}=a_m \\ x_j=0 \text{ for } j\notin\{i_1,\ldots,i_m\}}} . \end{array}$$

Furthermore, this homomorphism depends only upon the list $(a_1,\ldots,a_m)$, as it coincides with the following:

$$\begin{array}{rccc} \mathrm{QSym}(\{x_i\}_{i\in I}) & \cong \mathrm{QSym}(x_1,x_2,\ldots) & \longrightarrow & A, \\ & f(x_1,x_2,\ldots) & \longmapsto & f(a_1,\ldots,a_m,0,0\ldots). \end{array}$$

(This latter statement is stated for the case when I is infinite; otherwise, read "$x_1,x_2,\ldots,x_{|I|}$" for "$x_1,x_2,\ldots$", and interpret $(a_1,\ldots,a_m,0,0\ldots)$ as an $|I|$-tuple.)

Proof. One already can make sense of evaluating $x_{i_1} = a_1, \ldots, x_{i_m} = a_m$ and $x_j = 0$ for $j \notin \{i_1,\ldots,i_m\}$ in the ambient ring $R(\{x_i\}_{i\in I})$ containing $\mathrm{QSym}(\{x_i\}_{i\in I})$, since a power series f of bounded degree will have finitely many monomials that only involve the variables $x_{i_1},\ldots,x_{i_m}$. The last assertion follows from quasisymmetry of f, and is perhaps checked most easily when $f = M_\alpha(\{x_i\}_{i\in I})$ for some α. □

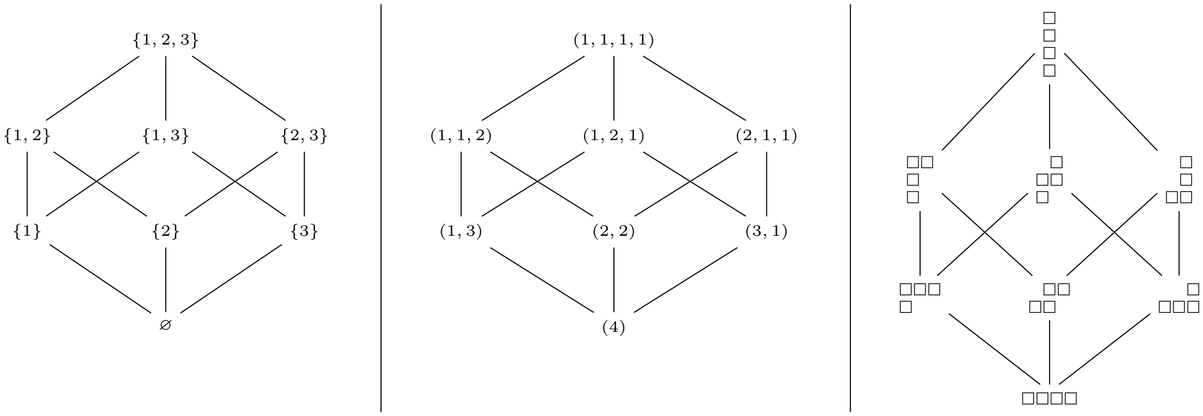
The antipode in QSym has a reasonably simple expression in the $\{M_\alpha\}$ basis, but requiring a definition.

Definition 5.1.10. For α, β in Comp_n, say that α *refines* β or β *coarsens* α if, informally, one can obtain β from α by combining some of its adjacent parts. Alternatively, this can be defined as follows: One has a bijection $\mathrm{Comp}_n \to 2^{[n-1]}$ where $[n-1] := \{1,2,\ldots,n-1\}$ which sends $\alpha = (\alpha_1,\ldots,\alpha_\ell)$ having length $\ell(\alpha) = \ell$ to its subset of partial sums

$$D(\alpha) := \{\alpha_1, \alpha_1+\alpha_2, \ldots, \alpha_1+\cdots+\alpha_{\ell-1}\},$$

and this sends the refinement ordering to the inclusion ordering on the Boolean algebra $2^{[n-1]}$ (to be more precise: a composition $\alpha \in \mathrm{Comp}_n$ refines a composition $\beta \in \mathrm{Comp}_n$ if and only if $D(\alpha) \supset D(\beta)$).

There is also a bijection sending every composition α to its *ribbon* diagram $\operatorname{Rib}(\alpha)$: the skew diagram λ/μ having rows of sizes $\alpha_1, \ldots, \alpha_\ell$ read from bottom to top with exactly one column of overlap between adjacent rows. These bijections and the refinement partial order are illustrated here for $n = 4$:



(where we have drawn each ribbon diagram with its boxes spaced out).

Given $\alpha = (\alpha_1, \ldots, \alpha_\ell)$, its *reverse* composition is $\operatorname{rev}(\alpha) = (\alpha_\ell, \alpha_{\ell-1}, \ldots, \alpha_2, \alpha_1)$. Note that $\alpha \mapsto \operatorname{rev}(\alpha)$ is a poset automorphism of Comp_n for the refinement ordering.

Theorem 5.1.11. *For any composition α in* Comp,

$$S(M_\alpha) = (-1)^{\ell(\alpha)} \sum_{\substack{\gamma \in \mathrm{Comp}: \\ \gamma \textit{ coarsens } \operatorname{rev}(\alpha)}} M_\gamma .$$

For example,

$$S(M_{(a,b,c)}) = -\left(M_{(c,b,a)} + M_{(b+c,a)} + M_{(c,a+b)} + M_{(a+b+c)}\right).$$

Proof. We give Ehrenborg's proof[256] [64, Prop. 3.4] via induction on $\ell = \ell(\alpha)$. One has easy base cases when $\ell(\alpha) = 0$, where $S(M_\varnothing) = S(1) = 1 = (-1)^0 M_{\operatorname{rev}(\varnothing)}$, and when $\ell(\alpha) = 1$, where $M_{(n)}$ is primitive by Proposition 5.1.7, so Proposition 1.4.17 shows $S(M_{(n)}) = -M_{(n)} = (-1)^1 M_{\operatorname{rev}((n))}$.

For the inductive step, apply the inductive definition of S from the proof of Proposition 1.4.16:

$$\begin{aligned} S(M_{(\alpha_1,\ldots,\alpha_\ell)}) &= -\sum_{i=0}^{\ell-1} S(M_{(\alpha_1,\ldots,\alpha_i)}) M_{(\alpha_{i+1},\ldots,\alpha_\ell)} \\ &= \sum_{i=0}^{\ell-1} \sum_{\substack{\beta \text{ coarsening} \\ (\alpha_i,\alpha_{i-1},\ldots,\alpha_1)}} (-1)^{i+1} M_\beta M_{(\alpha_{i+1},\ldots,\alpha_\ell)}. \end{aligned}$$

The idea will be to cancel terms of opposite sign that appear in the expansions of the products $M_\beta M_{(\alpha_{i+1},\ldots,\alpha_\ell)}$. Note that each composition β appearing above has first part β_1 of the form $\alpha_i + \alpha_{i-1} + \cdots + \alpha_h$ for some $h \leq i$ (unless $\beta = \varnothing$), and hence each term M_γ in the expansion of the product $M_\beta M_{(\alpha_{i+1},\ldots,\alpha_\ell)}$ has γ_1 (that is, the first entry of γ) a sum that can take one of these three forms:

- $\alpha_i + \alpha_{i-1} + \cdots + \alpha_h$,
- $\alpha_{i+1} + (\alpha_i + \alpha_{i-1} + \cdots + \alpha_h)$,
- α_{i+1}.

Say that the *type* of γ is i in the first case, and $i + 1$ in the second two cases[257]; in other words, the type is the largest subscript k on a part α_k which was combined in the sum γ_1. It is not hard to see that a given γ for which the type k is strictly smaller than ℓ arises from exactly two pairs $(\beta, \gamma), (\beta', \gamma)$, having opposite

[256] A different proof was given by Malvenuto and Reutenauer [146, Cor. 2.3], and is sketched in Remark 5.4.4 below.

[257] We imagine that we label the terms obtained by expanding $M_\beta M_{(\alpha_{i+1},\ldots,\alpha_\ell)}$ by distinct labels, so that each term knows how exactly it was created (i.e., which i, which β and which map f as in (5.1.2) gave rise to it). Strictly speaking, it is these triples (i, β, f) that we should be assigning types to, not terms.

signs $(-1)^k$ and $(-1)^{k+1}$ in the above sum[258]. For example, if $\alpha = (\alpha_1, \ldots, \alpha_8)$, then the composition $\gamma = (\alpha_6 + \alpha_5 + \alpha_4, \alpha_3, \alpha_7, \alpha_8 + \alpha_2 + \alpha_1)$ of type 6 can arise from either of

$$\beta = (\alpha_6 + \alpha_5 + \alpha_4, \alpha_3, \alpha_2 + \alpha_1) \text{ with } i = 6 \text{ and sign } (-1)^7,$$
$$\beta' = (\alpha_5 + \alpha_4, \alpha_3, \alpha_2 + \alpha_1) \text{ with } i = 5 \text{ and sign } (-1)^6.$$

Similarly, $\gamma = (\alpha_6, \alpha_5 + \alpha_4, \alpha_3, \alpha_7, \alpha_8 + \alpha_2 + \alpha_1)$ can arise from either of

$$\beta = (\alpha_6, \alpha_5 + \alpha_4, \alpha_3, \alpha_2 + \alpha_1) \text{ with } i = 6 \text{ and sign } (-1)^7,$$
$$\beta' = (\alpha_5 + \alpha_4, \alpha_3, \alpha_2 + \alpha_1) \text{ with } i = 5 \text{ and sign } (-1)^6.$$

Thus one can cancel almost all the terms, excepting those with γ of type ℓ among the terms M_γ in the expansion of the last ($i = \ell - 1$) summand $M_\beta M_{(\alpha_\ell)}$. A bit of thought shows that these are the γ coarsening $\operatorname{rev}(\alpha)$, and all have sign $(-1)^\ell$. □

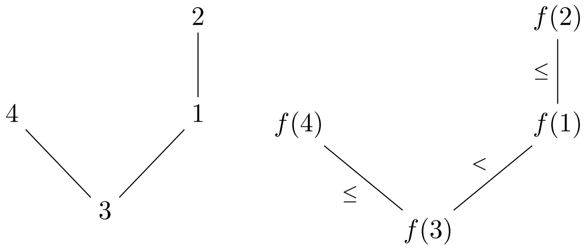
5.2. **The fundamental basis and P-partitions.** There is a second important basis for QSym which arose originally in Stanley's P-partition theory [203].[259]

Definition 5.2.1. A *labelled poset* will here mean a partially ordered set P whose underlying set is some finite subset of the integers. A *P-partition* is a function $P \xrightarrow{f} \{1, 2, \ldots\}$ with the following two properties:

- If $i \in P$ and $j \in P$ satisfy $i <_P j$ and $i <_{\mathbb{Z}} j$, then $f(i) \leq f(j)$.
- If $i \in P$ and $j \in P$ satisfy $i <_P j$ and $i >_{\mathbb{Z}} j$, then $f(i) < f(j)$.

Denote by $\mathcal{A}(P)$ the set of all P-partitions f, and let $F_P(\mathbf{x}) := \sum_{f \in \mathcal{A}(P)} \mathbf{x}_f$ where $\mathbf{x}_f := \prod_{i \in P} x_{f(i)}$. This $F_P(\mathbf{x})$ is an element of $\mathbf{k}[[\mathbf{x}]] := \mathbf{k}[[x_1, x_2, \ldots]]$.

Example 5.2.2. Depicted is a labelled poset P, along with the relations among the four values $f = (f(1), f(2), f(3), f(4))$ that define its P-partitions f:



Remark 5.2.3. Stanley's treatment of P-partitions in [206, §3.15 and §7.19] uses a language different from ours. First, Stanley works not with labelled posets P, but with pairs (P, ω) of a poset P and a bijective labelling $\omega : P \to [n]$. Thus, the relation $<_{\mathbb{Z}}$ is not given on P a priori, but has to be pulled back from $[n]$ using ω (and it depends on ω, whence Stanley speaks of "(P, ω)-partitions"). Furthermore, what we call "P-partition" is called a "reverse P-partition" in [206]. Finally, Stanley uses the notations F_P and $F_{P,\omega}$ for something different from what we denote by F_P, whereas what we call F_P is dubbed $K_{P,\omega}$ in [206, §7.19].

The so-called *fundamental quasisymmetric functions* are an important special case of the $F_P(\mathbf{x})$. We shall first define them directly and then see how they are obtained as P-partition enumerators $F_P(\mathbf{x})$ for some special labelled posets P.

Definition 5.2.4. Let $n \in \mathbb{N}$ and $\alpha \in \mathrm{Comp}_n$. We define the *fundamental quasisymmetric function* $L_\alpha = L_\alpha(\mathbf{x}) \in \mathrm{QSym}$ by

$$L_\alpha := \sum_{\substack{\beta \in \mathrm{Comp}_n : \\ \beta \text{ refines } \alpha}} M_\beta. \tag{5.2.1}$$

[258] Strictly speaking, this means that we have an involution on the set of our (i, β, f) triples having type smaller than ℓ, and this involution switches the sign of $(-1)^i M_{\mathrm{wt}(f)}$.

[259] See [80] for a history of P-partitions; our notations, however, strongly differ from those in [80].

Example 5.2.5. The extreme cases for α in Comp_n give quasisymmetric functions L_α which are symmetric:

$$L_{(1^n)} = M_{(1^n)} = e_n,$$
$$L_{(n)} = \sum_{\alpha \in \operatorname{Comp}_n} M_\alpha = h_n.$$

Before studying the L_α in earnest, we recall a basic fact about finite sets, which is sometimes known as the "principle of inclusion and exclusion" (although it is more general than the formula for the size of a union of sets that commonly goes by this name):

Lemma 5.2.6. *Let G be a finite set. Let V be a $\mathbf{k}$-module. For each subset A of G, we let f_A and g_A be two elements of V.*

(a) *If*

$$\text{every } A \subset G \text{ satisfies } g_A = \sum_{B \subset A} f_B,$$

then

$$\text{every } A \subset G \text{ satisfies } f_A = \sum_{B \subset A} (-1)^{|A \setminus B|} g_B.$$

(b) *If*

$$\text{every } A \subset G \text{ satisfies } g_A = \sum_{B \subset G;\ B \supset A} f_B,$$

then

$$\text{every } A \subset G \text{ satisfies } f_A = \sum_{B \subset G;\ B \supset A} (-1)^{|B \setminus A|} g_B.$$

Proof. This can be proven by elementary arguments (easy exercise). Alternatively, Lemma 5.2.6 can be viewed as a particular case of the Möbius inversion principle (see, e.g., [206, Propositions 3.7.1 and 3.7.2]) applied to the Boolean lattice 2^G (whose Möbius function is very simple: see [206, Example 3.8.3]). (This is spelled out in [138, Example 4.52], for example.) □

Lemma 5.2.6 can be translated into the language of compositions:

Lemma 5.2.7. *Let $n \in \mathbb{N}$. Let V be a $\mathbf{k}$-module. For each $\alpha \in \operatorname{Comp}_n$, we let f_α and g_α be two elements of V.*

(a) *If*

$$\text{every } \alpha \in \operatorname{Comp}_n \text{ satisfies } g_\alpha = \sum_{\beta \text{ coarsens } \alpha} f_\beta,$$

then

$$\text{every } \alpha \in \operatorname{Comp}_n \text{ satisfies } f_\alpha = \sum_{\beta \text{ coarsens } \alpha} (-1)^{\ell(\alpha)-\ell(\beta)} g_\beta.$$

(b) *If*

$$\text{every } \alpha \in \operatorname{Comp}_n \text{ satisfies } g_\alpha = \sum_{\beta \text{ refines } \alpha} f_\beta,$$

then

$$\text{every } \alpha \in \operatorname{Comp}_n \text{ satisfies } f_\alpha = \sum_{\beta \text{ refines } \alpha} (-1)^{\ell(\beta)-\ell(\alpha)} g_\beta.$$

Proof. Set $[n-1] = \{1, 2, \ldots, n-1\}$. Recall (from Definition 5.1.10) that there is a bijection $D : \operatorname{Comp}_n \to 2^{[n-1]}$ that sends each $\alpha \in \operatorname{Comp}_n$ to $D(\alpha) \subset [n-1]$. This bijection D has the properties that:

- a composition β refines a composition α if and only if $D(\beta) \supset D(\alpha)$;
- a composition β coarsens a composition α if and only if $D(\beta) \subset D(\alpha)$;
- any composition $\alpha \in \operatorname{Comp}_n$ satisfies $|D(\alpha)| = \ell(\alpha) - 1$ (unless $n = 0$), and thus
- any compositions α and β in Comp_n satisfy $|D(\alpha)| - |D(\beta)| = \ell(\alpha) - \ell(\beta)$.

This creates a dictionary between compositions in Comp_n and subsets of $[n-1]$. Now, apply Lemma 5.2.6 to $G = [n-1]$, $f_A = f_{D^{-1}(A)}$ and $g_A = g_{D^{-1}(A)}$, and translate using the dictionary. □

Now, we can see the following about the fundamental quasisymmetric functions:

Proposition 5.2.8. *The family $\{L_\alpha\}_{\alpha \in \operatorname{Comp}}$ is a $\mathbf{k}$-basis for* QSym, *and each $n \in \mathbb{N}$ and $\alpha \in \operatorname{Comp}_n$ satisfy*

$$M_\alpha = \sum_{\substack{\beta \in \operatorname{Comp}_n : \\ \beta \text{ refines } \alpha}} (-1)^{\ell(\beta)-\ell(\alpha)} L_\beta. \tag{5.2.2}$$

Proof. Fix $n \in \mathbb{N}$. Recall the equality (5.2.1). Thus, Lemma 5.2.7(b) (applied to $V = \mathrm{QSym}$, $f_\alpha = M_\alpha$ and $g_\alpha = L_\alpha$) yields (5.2.2).

Recall that the family $(M_\alpha)_{\alpha \in \operatorname{Comp}_n}$ is a basis of the $\mathbf{k}$-module QSym_n. The equality (5.2.1) shows that the family $(L_\alpha)_{\alpha \in \operatorname{Comp}_n}$ expands invertibly triangularly[260] with respect to the family $(M_\alpha)_{\alpha \in \operatorname{Comp}_n}$ (where Comp_n is equipped with the refinement order).[261] Thus, Corollary 11.1.19(e) (applied to QSym_n, Comp_n, $(M_\alpha)_{\alpha \in \operatorname{Comp}_n}$ and $(L_\alpha)_{\alpha \in \operatorname{Comp}_n}$ instead of M, S, $(e_s)_{s \in S}$ and $(f_s)_{s \in S}$) shows that the family $(L_\alpha)_{\alpha \in \operatorname{Comp}_n}$ is a basis of the $\mathbf{k}$-module QSym_n. Combining this fact for all $n \in \mathbb{N}$, we conclude that the family $(L_\alpha)_{\alpha \in \operatorname{Comp}}$ is a basis of the $\mathbf{k}$-module QSym. This completes the proof of Proposition 5.2.8. □

Proposition 5.2.9. *Let $n \in \mathbb{N}$. Let α be a composition of n. Let I be an infinite totally ordered set. Then,*

$$L_\alpha\left(\{x_i\}_{i \in I}\right) = \sum_{\substack{i_1 \leq i_2 \leq \cdots \leq i_n \text{ in } I; \\ i_j < i_{j+1} \text{ if } j \in D(\alpha)}} x_{i_1} x_{i_2} \cdots x_{i_n},$$

where $L_\alpha\left(\{x_i\}_{i \in I}\right)$ is defined as the image of L_α under the isomorphism $\mathrm{QSym} \to \mathrm{QSym}\left(\{x_i\}_{i \in I}\right)$ *obtained in Definition 5.1.5. In particular, for the standard (totally ordered) variable set $\mathbf{x} = (x_1 < x_2 < \cdots)$, we obtain*

$$L_\alpha = L_\alpha(\mathbf{x}) = \sum_{\substack{(1 \leq) i_1 \leq i_2 \leq \cdots \leq i_n; \\ i_j < i_{j+1} \text{ if } j \in D(\alpha)}} x_{i_1} x_{i_2} \cdots x_{i_n}. \tag{5.2.3}$$

Proof. Every composition $\beta = (\beta_1, \ldots, \beta_\ell)$ of n satisfies

$$M_\beta\left(\{x_i\}_{i \in I}\right) = \sum_{k_1 < \cdots < k_\ell \text{ in } I} x_{k_1}^{\beta_1} \cdots x_{k_\ell}^{\beta_\ell} = \sum_{\substack{i_1 \leq i_2 \leq \cdots \leq i_n \text{ in } I; \\ i_j < i_{j+1} \text{ if and only if } j \in D(\beta)}} x_{i_1} x_{i_2} \cdots x_{i_n}. \tag{5.2.4}$$

Applying the ring homomorphism $\mathrm{QSym} \to \mathrm{QSym}\left(\{x_i\}_{i \in I}\right)$ to (5.2.1), we obtain

$$\begin{aligned}
L_\alpha\left(\{x_i\}_{i \in I}\right) &= \sum_{\substack{\beta \in \operatorname{Comp}_n : \\ \beta \text{ refines } \alpha}} M_\beta\left(\{x_i\}_{i \in I}\right) \overset{(5.2.4)}{=} \sum_{\substack{\beta \in \operatorname{Comp}_n : \\ \beta \text{ refines } \alpha}} \sum_{\substack{i_1 \leq i_2 \leq \cdots \leq i_n \text{ in } I; \\ i_j < i_{j+1} \text{ if and only if } j \in D(\beta)}} x_{i_1} x_{i_2} \cdots x_{i_n} \\
&= \sum_{\substack{\beta \in \operatorname{Comp}_n : \\ D(\alpha) \subset D(\beta)}} \sum_{\substack{i_1 \leq i_2 \leq \cdots \leq i_n \text{ in } I; \\ i_j < i_{j+1} \text{ if and only if } j \in D(\beta)}} x_{i_1} x_{i_2} \cdots x_{i_n} \\
&= \sum_{\substack{Z \subset [n-1] : \\ D(\alpha) \subset Z}} \sum_{\substack{i_1 \leq i_2 \leq \cdots \leq i_n \text{ in } I; \\ i_j < i_{j+1} \text{ if and only if } j \in Z}} x_{i_1} x_{i_2} \cdots x_{i_n} = \sum_{\substack{i_1 \leq i_2 \leq \cdots \leq i_n \text{ in } I; \\ i_j < i_{j+1} \text{ if } j \in D(\alpha)}} x_{i_1} x_{i_2} \cdots x_{i_n}.
\end{aligned}$$

□

Proposition 5.2.10. *Assume that the labelled poset P is a total or linear order $w = (w_1 < \cdots < w_n)$ (that is, $P = \{w_1, w_2, \ldots, w_n\}$ as sets, and the order $<_P$ is given by $w_1 <_P w_2 <_P \cdots <_P w_n$). Let* $\operatorname{Des}(w)$ *be the* descent set *of w, defined by*

$$\operatorname{Des}(w) := \{i : w_i >_{\mathbb{Z}} w_{i+1}\} \subset \{1, 2, \ldots, n-1\}.$$

[260] See Section 11.1 for a definition of this concept.

[261] In fact, it expands unitriangularly with respect to the latter family.

Let $\alpha \in \operatorname{Comp}_n$ be the unique composition in Comp_n having partial sums $D(\alpha) = \operatorname{Des}(w)$. Then, the generating function $F_w(\mathbf{x})$ equals the fundamental quasisymmetric function L_α. In particular, $F_w(\mathbf{x})$ depends only upon the descent set $\operatorname{Des}(w)$.

E.g., total order $w = 35142$ has $\operatorname{Des}(w) = \{2,4\}$ and composition $\alpha = (2,2,1)$, so

$$\begin{aligned} F_{35142}(\mathbf{x}) &= \sum_{f(3)\leq f(5)<f(1)\leq f(4)<f(2)} x_{f(3)}x_{f(5)}x_{f(1)}x_{f(4)}x_{f(2)} \\ &= \sum_{i_1\leq i_2<i_3\leq i_4<i_5} x_{i_1}x_{i_2}x_{i_3}x_{i_4}x_{i_5} \\ &= L_{(2,2,1)} = M_{(2,2,1)} + M_{(2,1,1,1)} + M_{(1,1,2,1)} + M_{(1,1,1,1,1)}. \end{aligned}$$

Proof of Proposition 5.2.10. Write $F_w(\mathbf{x})$ as a sum of monomials $x_{f(w_1)} \cdots x_{f(w_n)}$ over all w-partitions f. These w-partitions are exactly the maps $f : w \to \{1,2,3,\ldots\}$ satisfying $f(w_1) \leq \cdots \leq f(w_n)$ and having strict inequalities $f(w_i) < f(w_{i+1})$ whenever i is in $\operatorname{Des}(w)$ (because if two elements w_a and w_b of w satisfy $w_a <_w w_b$ and $w_a >_{\mathbb{Z}} w_b$, then they must satisfy $a < b$ and $i \in \operatorname{Des}(w)$ for some $i \in \{a, a+1, \ldots, b-1\}$; thus, the conditions "$f(w_1) \leq \cdots \leq f(w_n)$" and "$f(w_i) < f(w_{i+1})$ whenever i is in $\operatorname{Des}(w)$" ensure that $f(w_a) < f(w_b)$ in this case). Therefore, they are in bijection with the weakly increasing sequences $(i_1 \leq i_2 \leq \cdots \leq i_n)$ of positive integers having strict inequalities $i_j < i_{j+1}$ whenever $i \in \operatorname{Des}(w)$ (namely, the bijection sends any w-partition f to the sequence $(f(w_1) \leq f(w_2) \leq \cdots \leq f(w_n))$). Hence,

$$F_w(\mathbf{x}) = \sum_{f\in\mathcal{A}(w)} \mathbf{x}_f = \sum_{\substack{(1\leq)i_1\leq i_2\leq\cdots\leq i_n;\\ i_j<i_{j+1} \text{ if } j\in\operatorname{Des}(w)}} x_{i_1}x_{i_2}\cdots x_{i_n} = \sum_{\substack{(1\leq)i_1\leq i_2\leq\cdots\leq i_n;\\ i_j<i_{j+1} \text{ if } j\in D(\alpha)}} x_{i_1}x_{i_2}\cdots x_{i_n}$$

(since $\operatorname{Des}(w) = D(\alpha)$). Comparing this with (5.2.3), we conclude that $F_w(\mathbf{x}) = L_\alpha$. □

The next proposition ([206, Cor. 7.19.5], [140, Cor. 3.3.24]) is an algebraic shadow of Stanley's main lemma [206, Thm. 7.19.4] in P-partition theory. It expands any $F_P(\mathbf{x})$ in the $\{L_\alpha\}$ basis, as a sum over the set $\mathcal{L}(P)$ of all *linear extensions* w of P [262]. E.g., the poset P from Example 5.2.2 has $\mathcal{L}(P) = \{3124, 3142, 3412\}$.

Theorem 5.2.11. *For any labelled poset P,*

$$F_P(\mathbf{x}) = \sum_{w\in\mathcal{L}(P)} F_w(\mathbf{x}).$$

Proof. We give Gessel's proof [79, Thm. 1], via induction on the number of pairs i, j which are incomparable in P. When this quantity is 0, then P is itself a linear order w, so that $\mathcal{L}(P) = \{w\}$ and there is nothing to prove.

In the inductive step, let i, j be incomparable elements. Consider the two posets $P_{i<j}$ and $P_{j<i}$ which are obtained from P by adding in an order relation between i and j, and then taking the transitive closure; it is not hard to see that these transitive closures cannot contain a cycle, so that these really do define two posets. The result then follows by induction applied to $P_{i<j}, P_{j<i}$, once one notices that $\mathcal{L}(P) = \mathcal{L}(P_{i<j}) \sqcup \mathcal{L}(P_{j<i})$ since every linear extension w of P either has i before j or vice-versa, and $\mathcal{A}(P) = \mathcal{A}(P_{i<j}) \sqcup \mathcal{A}(P_{j<i})$ since, assuming that $i <_{\mathbb{Z}} j$ without loss of generality, every f in $\mathcal{A}(P)$ either satisfies $f(i) \leq f(j)$ or $f(i) > f(j)$. □

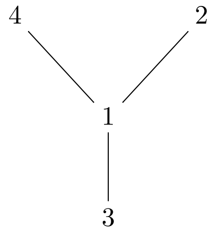
[262] Let us explain what we mean by linear extensions and how we represent them.

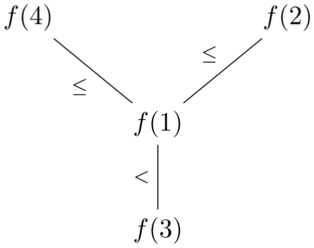
If $\mathbf{P}$ is a finite poset, then a *linear extension* of $\mathbf{P}$ denotes a total order w on the set $\mathbf{P}$ having the property that every two elements i and j of $\mathbf{P}$ satisfying $i <_{\mathbf{P}} j$ satisfy $i <_w j$. (In other words, it is a linear order on the ground set $\mathbf{P}$ which extends $\mathbf{P}$ as a poset; therefore the name.) We identify such a total order w with the list $(\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n)$ containing all elements of $\mathbf{P}$ in w-increasing order (that is, $\mathbf{p}_1 <_w \mathbf{p}_2 <_w \cdots <_w \mathbf{p}_n$).

(Stanley, in [206, §3.5], defines linear extensions in a slightly different way: For him, a linear extension of a finite poset $\mathbf{P}$ is an order-preserving bijection from $\mathbf{P}$ to the subposet $\{1, 2, \ldots, |\mathbf{P}|\}$ of $\mathbb{Z}$. But this is equivalent to our definition, since a bijection like this can be used to transport the order relation of $\{1, 2, \ldots, |\mathbf{P}|\}$ back to $\mathbf{P}$, thus resulting in a total order on $\mathbf{P}$ which is a linear extension of $\mathbf{P}$ in our sense.)

Example 5.2.12. To illustrate the induction in the above proof, consider the poset P from Example 5.2.2, having $\mathcal{L}(P) = \{3124, 3142, 3412\}$. Then choosing as incomparable pair $(i, j) = (1, 4)$, one has

$P_{i<j} =$

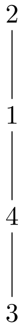


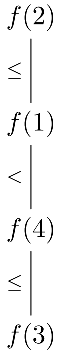


, thus $\mathcal{L}(P_{i<j}) = \{3124, 3142\}$

and

$P_{j<i} =$





, thus $\mathcal{L}(P_{j<i}) = \{3412\}$.

Exercise 5.2.13. Give an alternative proof for Theorem 5.2.11.

[**Hint:** For every $f : P \to \{1, 2, 3, \ldots\}$, we can define a binary relation $\prec_f$ on the set P by letting $i \prec_f j$ hold if and only if

$$(f(i) < f(j) \text{ or } (f(i) = f(j) \text{ and } i <_{\mathbb{Z}} j)).$$

Show that this binary relation $\prec_f$ is (the smaller relation of) a total order. When f is a P-partition, then endowing the set P with this total order yields a linear extension of P. Use this to show that the set $\mathcal{A}(P)$ is the union of its disjoint subsets $\mathcal{A}(w)$ with $w \in \mathcal{L}(P)$.]

Various other properties of the quasisymmetric functions $F_P(\mathbf{x})$ are studied, e.g., in [152].

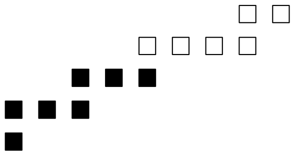
We next wish to describe the structure maps for the Hopf algebra QSym in the basis $\{L_\alpha\}$ of fundamental quasisymmetric functions. For this purpose, two more definitions are useful.

Definition 5.2.14. Given two nonempty compositions $\alpha = (\alpha_1, \ldots, \alpha_\ell)$ and $\beta = (\beta_1, \ldots, \beta_m)$, their *near-concatenation* is

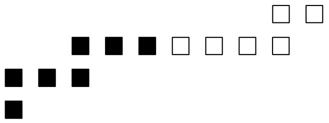
$$\alpha \odot \beta := (\alpha_1, \ldots, \alpha_{\ell-1}, \alpha_\ell + \beta_1, \beta_2, \ldots, \beta_m).$$

For example, the figure below depicts for $\alpha = (1, 3, 3)$ (black squares) and $\beta = (4, 2)$ (white squares) the concatenation and near-concatenation as ribbons:[263]

$\mathrm{Rib}(\alpha \cdot \beta) =$



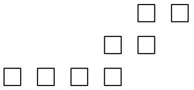
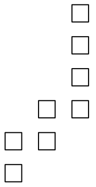
$\mathrm{Rib}(\alpha \odot \beta) =$



Lastly, given α in Comp_n, let $\omega(\alpha)$ be the unique composition in Comp_n whose partial sums $D(\omega(\alpha))$ form the complementary set within $[n-1]$ to the partial sums $D(\mathrm{rev}(\alpha))$; alternatively, one can check this means that the ribbon for $\omega(\alpha)$ is obtained from that of α by conjugation or transposing, that is, if $\mathrm{Rib}(\alpha) = \lambda/\mu$

[263] The ribbons are drawn with their boxes spaced out in order to facilitate counting.

then $\mathrm{Rib}\,(\omega(\alpha)) = \lambda^t/\mu^t$. E.g. if $\alpha = (4,2,2)$ so that $n = 8$, then $\mathrm{rev}(\alpha) = (2,2,4)$ has $D(\mathrm{rev}(\alpha)) = \{2,4\} \subset [7]$, complementary to the set $\{1,3,5,6,7\}$ which are the partial sums for $\omega(\alpha) = (1,2,2,1,1,1)$, and the ribbon diagrams of α and $\omega(\alpha)$ are

$\mathrm{Rib}\,(\alpha) =$  and $\mathrm{Rib}\,(\omega(\alpha)) =$ 

Proposition 5.2.15. *The structure maps for the Hopf algebra* QSym *in the basis $\{L_\alpha\}$ of fundamental quasisymmetric functions are as follows:*

$$\Delta L_\alpha = \sum_{\substack{(\beta,\gamma):\\ \beta\cdot\gamma=\alpha \text{ or } \beta\odot\gamma=\alpha}} L_\beta \otimes L_\gamma, \tag{5.2.5}$$

$$L_\alpha L_\beta = \sum_{w \in w_\alpha \sqcup\!\sqcup\, w_\beta} L_{\gamma(w)}, \tag{5.2.6}$$

$$S(L_\alpha) = (-1)^{|\alpha|} L_{\omega(\alpha)}. \tag{5.2.7}$$

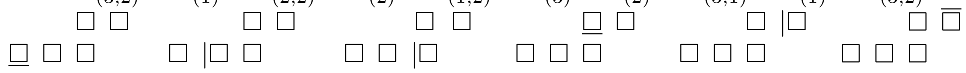
Here we are making use of the following notations in (5.2.6) *(recall also Definition 1.6.2):*

- *A* labelled linear order *will mean a labelled poset P whose order $<_P$ is a total order. We will identify any labelled linear order P with the word (over the alphabet $\mathbb{Z}$) obtained by writing down the elements of P in increasing order (with respect to the total order $<_P$). This way, every word (over the alphabet $\mathbb{Z}$) which has no two equal letters becomes identified with a labelled linear order.*
- *w_α is any labelled linear order with underlying set $\{1,2,\ldots,|\alpha|\}$ such that* $\mathrm{Des}\,(w_\alpha) = D\,(\alpha)$.
- *w_β is any labelled linear order with underlying set $\{|\alpha|+1, |\alpha|+2, \ldots, |\alpha|+|\beta|\}$ such that* $\mathrm{Des}\,(w_\beta) = D\,(\beta)$.
- *$\gamma(w)$ is the unique composition of $|\alpha|+|\beta|$ with* $D(\gamma(w)) = \mathrm{Des}(w)$.

(The right hand side of (5.2.6) *is to be read as a sum over all w, for a fixed choice of w_α and w_β.)*

At first glance the formula (5.2.5) for ΔL_α might seem more complicated than the formula of Proposition 5.1.7 for ΔM_α. However, it is equally simple when viewed in terms of ribbon diagrams: it cuts the ribbon diagram $\mathrm{Rib}\,(\alpha)$ into two smaller ribbons $\mathrm{Rib}\,(\beta)$ and $\mathrm{Rib}\,(\gamma)$, in all $|\alpha|+1$ possible ways, via *horizontal* cuts ($\beta\cdot\gamma = \alpha$) or *vertical* cuts ($\beta \odot \gamma = \alpha$). For example,

$$\begin{aligned} &\Delta L_{(3,2)} \\ &= 1 \otimes L_{(3,2)} + L_{(1)} \otimes L_{(2,2)} + L_{(2)} \otimes L_{(1,2)} + L_{(3)} \otimes L_{(2)} + L_{(3,1)} \otimes L_{(1)} + L_{(3,2)} \otimes 1. \end{aligned}$$



Example 5.2.16. To multiply $L_{(1,1)}L_{(2)}$, one could pick $w_\alpha = 21$ and $w_\beta = 34$, and then

$$\begin{aligned} L_{(1,1)}L_{(2)} = \sum_{w \in 21 \sqcup\!\sqcup\, 34} L_{\gamma(w)} &= L_{\gamma(2134)} + L_{\gamma(2314)} + L_{\gamma(3214)} + L_{\gamma(2341)} + L_{\gamma(3241)} + L_{\gamma(3421)} \\ &= L_{(1,3)} + L_{(2,2)} + L_{(1,1,2)} + L_{(3,1)} + L_{(1,2,1)} + L_{(2,1,1)}. \end{aligned}$$

Before we prove Proposition 5.2.15, we state a simple lemma:

Lemma 5.2.17. *Let Q and R be two labelled posets whose underlying sets are disjoint. Let $Q \sqcup R$ be the disjoint union of these posets Q and R; this is again a labelled poset. Then,*

$$F_Q(\mathbf{x})\, F_R(\mathbf{x}) = F_{Q\sqcup R}(\mathbf{x}).$$

Proof. We identify the underlying set of $Q \sqcup R$ with $Q \cup R$ (since the sets Q and R are already disjoint). If $f : Q \sqcup R \to \{1,2,3,\ldots\}$ is a $Q \sqcup R$-partition, then its restrictions $f\mid_Q$ and $f\mid_R$ are a Q-partition and an R-partition, respectively. Conversely, any pair of a Q-partition and an R-partition can be combined to form a $Q \sqcup R$-partition. Thus, there is a bijective correspondence between the addends in the expanded sum $F_Q(\mathbf{x})\, F_R(\mathbf{x})$ and the addends in $F_{Q\sqcup R}(\mathbf{x})$. $\square$

Proof of Proposition 5.2.15. To prove formula (5.2.5) for α in Comp_n, note that

$$\Delta L_\alpha = L_\alpha(\mathbf{x},\mathbf{y}) = \sum_{k=0}^{n} \sum_{\substack{1\leq i_1\leq\cdots\leq i_k,\\ 1\leq i_{k+1}\leq\cdots\leq i_n:\\ i_r<i_{r+1} \text{ for } r\in D(\alpha)\setminus\{k\}}} x_{i_1}\cdots x_{i_k}\cdot y_{i_{k+1}}\cdots y_{i_n} \tag{5.2.8}$$

by Proposition 5.2.9 (where we identify $\mathrm{QSym}\otimes\mathrm{QSym}$ with a $\mathbf{k}$-subalgebra of $R(\mathbf{x},\mathbf{y})$ by means of the embedding $\mathrm{QSym}\otimes\mathrm{QSym}\overset{\cong}{\to}\mathrm{QSym}(\mathbf{x})\otimes\mathrm{QSym}(\mathbf{y})\hookrightarrow R(\mathbf{x},\mathbf{y})$ as in the definition of the comultiplication on QSym). One then realizes that the inner sums corresponding to values of k that lie (resp. do not lie) in $D(\alpha)\cup\{0,n\}$ correspond to the terms $L_\beta(\mathbf{x})L_\gamma(\mathbf{y})$ for pairs (β,γ) in which $\beta\cdot\gamma=\alpha$ (resp. $\beta\odot\gamma=\alpha$).

For formula (5.2.6), let P be the labelled poset which is the disjoint union of linear orders w_α, w_β. Then

$$L_\alpha L_\beta = F_{w_\alpha}(\mathbf{x})F_{w_\beta}(\mathbf{x}) = F_P(\mathbf{x}) = \sum_{w\in\mathcal{L}(P)} F_w(\mathbf{x}) = \sum_{w\in w_\alpha \sqcup\!\sqcup\, w_\beta} L_{\gamma(w)}$$

where the first equality used Proposition 5.2.10, the second equality comes from Lemma 5.2.17, the third equality from Theorem 5.2.11, and the fourth from the equality $\mathcal{L}(P) = w_\alpha \sqcup\!\sqcup\, w_\beta$.

To prove formula (5.2.7), compute using Theorem 5.1.11 that

$$S(L_\alpha) = \sum_{\beta \text{ refining } \alpha} S(M_\beta) = \sum_{\substack{(\beta,\gamma):\\ \beta \text{ refines } \alpha,\\ \gamma \text{ coarsens } \operatorname{rev}(\beta)}} (-1)^{\ell(\beta)}M_\gamma = \sum_\gamma M_\gamma \sum_\beta (-1)^{\ell(\beta)}$$

in which the last inner sum is over β for which

$$D(\beta)\supset D(\alpha)\cup D(\operatorname{rev}(\gamma)).$$

The alternating signs make such inner sums vanish unless they have only the single term where $D(\beta)=[n-1]$ (that is, $\beta=(1^n)$). This happens exactly when $D(\operatorname{rev}(\gamma))\cup D(\alpha)=[n-1]$ or equivalently, when $D(\operatorname{rev}(\gamma))$ contains the complement of $D(\alpha)$, that is, when $D(\gamma)$ contains the complement of $D(\operatorname{rev}(\alpha))$, that is, when γ refines $\omega(\alpha)$. Thus

$$S(L_\alpha) = \sum_{\substack{\gamma\in\operatorname{Comp}_n:\\ \gamma \text{ refines } \omega(\alpha)}} M_\gamma\cdot(-1)^n = (-1)^{|\alpha|}L_{\omega(\alpha)}.$$

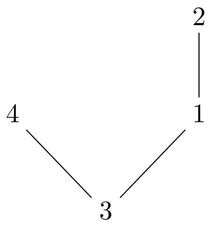
□

The antipode formula (5.2.7) for L_α leads to a general interpretation for the antipode of QSym acting on P-partition enumerators $F_P(\mathbf{x})$.

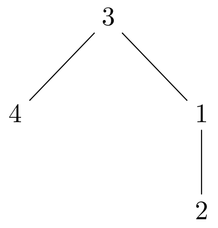
Definition 5.2.18. Given a labelled poset P on $\{1,2,\ldots,n\}$, let the *opposite* or *dual* labelled poset P^{opp} be the labelled poset on $\{1,2,\ldots,n\}$ that has $i<_{P^{\mathrm{opp}}} j$ if and only if $j<_P i$.

For example,

$P =$



$P^{\mathrm{opp}} =$

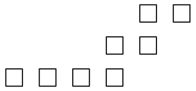
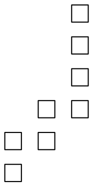


The following observation is straightforward.

Proposition 5.2.19. *When P is a linear order corresponding to some permutation $w=(w_1,\ldots,w_n)$ in $\mathfrak{S}_n$, then $w^{\mathrm{opp}}=ww_0$ where $w_0\in\mathfrak{S}_n$ is the permutation that swaps $i\leftrightarrow n+1-i$ (this is the so-called* longest permutation, *thus named due to it having the highest "Coxeter length" among all permutations in $\mathfrak{S}_n$). Furthermore, in this situation one has $F_w(\mathbf{x})=L_\alpha$, that is,* $\operatorname{Des}(w)=D(\alpha)$ *if and only if* $\operatorname{Des}(w^{\mathrm{opp}})=D(\omega(\alpha))$*, that is $F_{w^{\mathrm{opp}}}(\mathbf{x})=L_{\omega(\alpha)}$. Thus,*

$$S(F_w(\mathbf{x}))=(-1)^nF_{w^{\mathrm{opp}}}(\mathbf{x}).$$

For example, given the compositions considered earlier:

$\alpha = (4,2,2) =$  and $\omega(\alpha) = (1,2,2,1,1,1) =$ 

if one picks $w = 1235 \cdot 47 \cdot 68$ (with descent positions marked by dots) having $\mathrm{Des}(w) = \{4,6\} = D(\alpha)$, then $w^{\mathrm{opp}} = ww_0 = 8 \cdot 67 \cdot 45 \cdot 3 \cdot 2 \cdot 1$ has $\mathrm{Des}(w^{\mathrm{opp}}) = \{1,3,5,6,7\} = D(\omega(\alpha))$.

Corollary 5.2.20. *For any labelled poset P on $\{1,2,\ldots,n\}$, one has*

$$S\left(F_P(\mathbf{x})\right) = (-1)^n F_{P^{\mathrm{opp}}}(\mathbf{x}).$$

Proof. Since S is linear, one can apply Theorem 5.2.11 and Proposition 5.2.19, obtaining

$$S\left(F_P(\mathbf{x})\right) = \sum_{w \in \mathcal{L}(P)} S(F_w(\mathbf{x})) = \sum_{w \in \mathcal{L}(P)} (-1)^n F_{w^{\mathrm{opp}}}(\mathbf{x}) = (-1)^n F_{P^{\mathrm{opp}}}(\mathbf{x}),$$

as $\mathcal{L}(P^{\mathrm{opp}}) = \{w^{\mathrm{opp}} : w \in \mathcal{L}(P)\}$. □

Remark 5.2.21. Malvenuto and Reutenauer, in [147, Theorem 3.1], prove an even more general antipode formula, which encompasses our Corollary 5.2.20, Proposition 5.2.19, Theorem 5.1.11 and (5.2.7). See [85, Theorem 4.2] for a restatement and a self-contained proof of this theorem (and [85, Theorem 4.7] for an even further generalization).

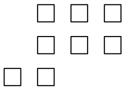
We remark on a special case of Corollary 5.2.20 to which we alluded earlier, related to skew Schur functions.

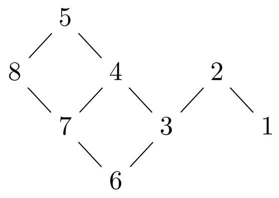
Corollary 5.2.22. *In Λ, the action of ω and the antipode S on skew Schur functions $s_{\lambda/\mu}$ are as follows:*

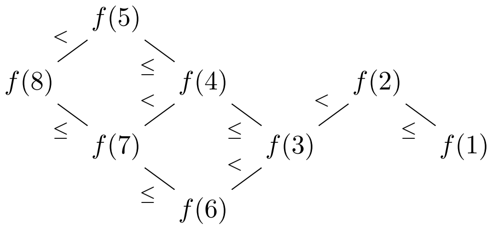
$$\omega(s_{\lambda/\mu}) = s_{\lambda^t/\mu^t}, \tag{5.2.9}$$

$$S(s_{\lambda/\mu}) = (-1)^{|\lambda/\mu|} s_{\lambda^t/\mu^t}. \tag{5.2.10}$$

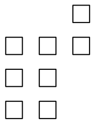
Proof. Given a skew shape λ/μ, one can always create a labelled poset P which is its *skew Ferrers poset*, together with one of many *column-strict labellings*, in such a way that $F_P(\mathbf{x}) = s_{\lambda/\mu}(\mathbf{x})$. An example is shown here for $\lambda/\mu = (4,4,2)/(1,1,0)$:

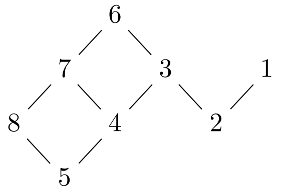
$\lambda/\mu =$ 

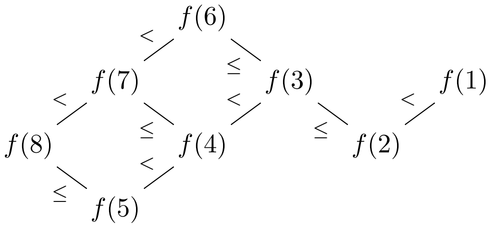
$P =$ 



The general definition is as follows: Let P be the set of all boxes of the skew diagram λ/μ. Label these boxes by the numbers $1,2,\ldots,n$ (where $n = |\lambda/\mu|$) row by row from bottom to top (reading every row from left to right), and then define an order relation $<_P$ on P by requiring that every box be smaller (in P) than its right neighbor and smaller (in P) than its lower neighbor. It is not hard to see that in this situation, $F_{P^{\mathrm{opp}}}(\mathbf{x}) = \sum_T \mathbf{x}^{\mathrm{cont}(T)}$ as T ranges over all *reverse semistandard tableaux* or *column-strict plane partitions* of λ^t/μ^t:

$\lambda^t/\mu^t =$ 

$P^{\mathrm{opp}} =$ 



But this means that $F_{P^{\mathrm{opp}}}(\mathbf{x}) = s_{\lambda^t/\mu^t}(\mathbf{x})$, since the fact that skew Schur functions lie in Λ implies that they can be defined either as generating functions for column-strict tableaux or reverse semistandard tableaux; see Remark 2.2.5 above, or [206, Prop. 7.10.4].

Thus we have

$$F_P(\mathbf{x}) = s_{\lambda/\mu}(\mathbf{x}),$$
$$F_{P^{\mathrm{opp}}}(\mathbf{x}) = s_{\lambda^t/\mu^t}(\mathbf{x}).$$

Corollary 1.4.27 tells us that the antipode for QSym must specialize to the antipode for Λ (see also Remark 5.4.11 below), so (5.2.10) is a special case of Corollary 5.2.20. Then (5.2.9) follows from the relation (2.4.11) that $S(f) = (-1)^n \omega(f)$ for f in Λ_n. □

Remark 5.2.23. Before leaving P-partitions temporarily, we mention two open questions about them.

The first is a conjecture of Stanley from his thesis [203]. As mentioned in the proof of Corollary 5.2.22, each skew Schur function $s_{\lambda/\mu}(\mathbf{x})$ is a special instance of P-partition enumerator $F_P(\mathbf{x})$.

Conjecture 5.2.24. *A labelled poset P has $F_P(\mathbf{x})$ symmetric, and not just quasisymmetric,* if and only if *P is a column-strict labelling of some skew Ferrers poset λ/μ.*

A somewhat weaker result in this direction was proven by Malvenuto in her thesis [145, Thm. 6.4], showing that if a labelled poset P has the stronger property that its set of linear extensions $\mathcal{L}(P)$ is a union of *plactic* or *Knuth equivalence classes*, then P must be a column-strict labelling of a skew Ferrers poset.

The next question is due to P. McNamara, and is suggested by the obvious factorizations of P-partition enumerators $F_{P_1 \sqcup P_2}(\mathbf{x}) = F_{P_1}(\mathbf{x}) F_{P_2}(\mathbf{x})$ (Lemma 5.2.17).

Question 5.2.25. If $\mathbf{k}$ is a field, does a *connected* labelled poset P always have $F_P(\mathbf{x})$ *irreducible* within the ring QSym?

The phrasing of this question requires further comment. It is assumed here that $\mathbf{x} = (x_1, x_2, \ldots)$ is infinite; for example when P is a 2-element chain labelled "against the grain" (i.e., the bigger element of the chain has the smaller label), then $F_P(\mathbf{x}) = e_2(\mathbf{x})$ is irreducible, but its specialization to two variables $\mathbf{x} = (x_1, x_2)$ is $e_2(x_1, x_2) = x_1 x_2$, which is reducible. If one wishes to work in finitely many variables $\mathbf{x} = (x_1, \ldots, x_m)$ one can perhaps assume that m is at least $|P| + 1$.

When working in QSym = QSym($\mathbf{x}$) in infinitely many variables, it is perhaps not so clear where factorizations occur. For example, if f lies in QSym and factors $f = g \cdot h$ with g, h in $R(\mathbf{x})$, does this imply that g, h also lie in QSym? The answer is "Yes" (for $\mathbf{k} = \mathbb{Z}$), but this is not obvious, and was proven by P. Pylyavskyy in [175, Chap. 11].

One also might wonder whether $\mathrm{QSym}_{\mathbb{Z}}$ is a unique factorization domain, but this follows from the result of M. Hazewinkel ([89] and [93, Thm. 6.7.5], and Theorem 6.4.3 further below) who proved a conjecture of Ditters that $\mathrm{QSym}_{\mathbb{Z}}$ is a polynomial algebra; earlier Malvenuto and Reutenauer [146, Cor. 2.2] had shown that $\mathrm{QSym}_{\mathbb{Q}}$ is a polynomial algebra. In fact, one can find polynomial generators $\{P_\alpha\}$ for $\mathrm{QSym}_{\mathbb{Q}}$ as a subset of the dual basis to the $\mathbb{Q}$-basis $\{\xi_\alpha\}$ for $\mathrm{NSym}_{\mathbb{Q}}$ which comes from taking products $\xi_\alpha := \xi_{\alpha_1} \cdots \xi_{\alpha_\ell}$ of the elements $\{\xi_n\}$ defined in Remark 5.4.4 below. Specifically, one takes those P_α for which the composition α is a *Lyndon composition*; see the First proof of Proposition 6.4.4 for a mild variation on this construction.

Hazewinkel's proof [93, Thm. 6.7.5] of the polynomiality of $\mathrm{QSym}_{\mathbb{Z}}$ also shows that QSym is a polynomial ring over Λ (see Corollary 6.5.33); in particular, this yields that QSym is a free Λ-module.[264]

An affirmative answer to Question 5.2.25 is known at least in the special case where P is a connected column-strict labelling of a skew Ferrers diagram, that is, when $F_P(\mathbf{x}) = s_{\lambda/\mu}(\mathbf{x})$ for some connected skew diagram λ/μ; see [13].

5.3. **Standardization of n-tuples and the fundamental basis.** Another equivalent description of the fundamental quasisymmetric functions L_α (Lemma 5.3.6 below) relies on the concept of words and of their standardizations. We shall study words in detail in Chapter 6; at this point, we merely introduce the few notions that we will need:

Definition 5.3.1. We fix a totally ordered set $\mathfrak{A}$, which we call the *alphabet*.

We recall that a *word over* $\mathfrak{A}$ is just a (finite) tuple of elements of $\mathfrak{A}$. A word $(w_1, w_2, \ldots, w_n)$ can be written as $w_1 w_2 \cdots w_n$ when this incurs no ambiguity.

If $w \in \mathfrak{A}^n$ is a word and $i \in \{1, 2, \ldots, n\}$, then the *$i$-th letter* of w means the i-th entry of the n-tuple w. This i-th letter will be denoted by w_i.

Our next definition relies on a simple fact about permutations and words:[265]

Proposition 5.3.2. *Let $w = (w_1, w_2, \ldots, w_n) \in \mathfrak{A}^n$ be any word. Then, there exists a unique permutation $\sigma \in \mathfrak{S}_n$ such that for every two elements a and b of $\{1, 2, \ldots, n\}$ satisfying $a < b$, we have ($\sigma(a) < \sigma(b)$ if and only if $w_a \leq w_b$).*

Definition 5.3.3. Let $w \in \mathfrak{A}^n$ be any word. The unique permutation $\sigma \in \mathfrak{S}_n$ defined in Proposition 5.3.2 is called the *standardization* of w, and is denoted by $\operatorname{std} w$.

Example 5.3.4. If $\mathfrak{A}$ is the alphabet $\{1 < 2 < 3 < \cdots\}$, then $\operatorname{std}(41211424)$ is the permutation which is written (in one-line notation) as 61423758.

A simple method to compute the standardization of a word $w \in \mathfrak{A}^n$ is the following: Replace all occurrences of the smallest letter appearing in w by the numbers $1, 2, \ldots, m_1$ (where m_1 is the number of these occurrences); then replace all occurrences of the second-smallest letter appearing in w by the numbers $m_1 + 1, m_1 + 2, \ldots, m_1 + m_2$ (where m_2 is the number of these occurrences), and so on, until all letters are replaced by numbers.[266] The result is the standardization of w, in one-line notation.

Another method to compute the standardization $\operatorname{std} w$ of a word $w = (w_1, w_2, \ldots, w_n) \in \mathfrak{A}^n$ is based on sorting. Namely, consider the total order on the set $\mathfrak{A} \times \mathbb{Z}$ given by

$$(a, i) \leq (b, j) \text{ if and only if } (\text{either } a < b \text{ or } (a = b \text{ and } i \leq j)).$$

(In other words, two pairs in $\mathfrak{A} \times \mathbb{Z}$ are compared by first comparing their first entries, and then, in the case of a tie, using the second entries as tiebreakers.) Now, in order to compute $\operatorname{std} w$, we sort the n-tuple $((w_1, 1), (w_2, 2), \ldots, (w_n, n)) \in (\mathfrak{A} \times \mathbb{Z})^n$ into increasing order (with respect to the total order just described), thus obtaining a new n-tuple of the form $\left(\left(w_{\tau(1)}, \tau(1)\right), \left(w_{\tau(2)}, \tau(2)\right), \ldots, \left(w_{\tau(n)}, \tau(n)\right)\right)$ for some $\tau \in \mathfrak{S}_n$; the standardization $\operatorname{std} w$ is then τ^{-1}.

Definition 5.3.5. Let $n \in \mathbb{N}$. Let $\sigma \in \mathfrak{S}_n$. Define a subset $\operatorname{Des} \sigma$ of $\{1, 2, \ldots, n-1\}$ by

$$\operatorname{Des} \sigma = \{i \in \{1, 2, \ldots, n-1\} \mid \sigma(i) > \sigma(i+1)\}.$$

(This is a particular case of the definition of $\operatorname{Des} w$ in Exercise 2.9.11, if we identify σ with the n-tuple $(\sigma(1), \sigma(2), \ldots, \sigma(n))$. It is also a particular case of the definition of $\operatorname{Des} w$ in Proposition 5.2.10, if we identify σ with the total order $(\sigma(1) < \sigma(2) < \cdots < \sigma(n))$ on the set $\{1, 2, \ldots, n\}$.)

There is a unique composition α of n satisfying $D(\alpha) = \operatorname{Des} \sigma$ (where $D(\alpha)$ is defined as in Definition 5.1.10). This composition will be denoted by $\gamma(\sigma)$.

[264] The latter statement has an analogue in finitely many indeterminates, proven by Lauve and Mason in [125, Corollary 13]: The quasisymmetric functions $\mathrm{QSym}\left(\{x_i\}_{i \in I}\right)$ are free as a $\Lambda\left(\{x_i\}_{i \in I}\right)$-module for any totally ordered set I, infinite or not. In the case of finite I, this cannot be derived by Hazewinkel's arguments, as the ring $\mathrm{QSym}\left(\{x_i\}_{i \in I}\right)$ is not in general a polynomial ring (e.g., when $\mathbf{k} = \mathbb{Q}$ and $I = \{1, 2\}$, this ring is not even a UFD, as witnessed by $\left(x_1^2 x_2\right) \cdot \left(x_1 x_2^2\right) = (x_1 x_2)^3$).

[265] See Exercise 5.3.7 below for a proof of Proposition 5.3.2.

[266] Here, a number is not considered to be a letter; thus, a number that replaces a letter will always be left in peace afterwards.

The following lemma (equivalent to [182, Lemma 9.39]) yields another description of the fundamental quasisymmetric functions:

Lemma 5.3.6. *Let $\mathfrak{A}$ denote the totally ordered set $\{1 < 2 < 3 < \cdots\}$ of positive integers. For each word $w = (w_1, w_2, \ldots, w_n) \in \mathfrak{A}^n$, we define a monomial $\mathbf{x}_w$ in $\mathbf{k}[[\mathbf{x}]]$ by $\mathbf{x}_w = x_{w_1} x_{w_2} \cdots x_{w_n}$.*

Let $n \in \mathbb{N}$ and $\sigma \in \mathfrak{S}_n$. Then,

$$L_{\gamma(\sigma)} = \sum_{\substack{w \in \mathfrak{A}^n; \\ \operatorname{std} w = \sigma^{-1}}} \mathbf{x}_w.$$

Exercise 5.3.7. Prove Proposition 5.3.2 and Lemma 5.3.6.

5.4. The Hopf algebra NSym dual to QSym.

We introduce here the (graded) dual Hopf algebra to QSym. This is well-defined, as QSym is connected graded of finite type.

Definition 5.4.1. Let $\operatorname{NSym} := \operatorname{QSym}^o$, with dual pairing $\operatorname{NSym} \otimes \operatorname{QSym} \xrightarrow{(\cdot,\cdot)} \mathbf{k}$. Let $\{H_\alpha\}$ be the $\mathbf{k}$-basis of NSym dual to the $\mathbf{k}$-basis $\{M_\alpha\}$ of QSym, so that

$$(H_\alpha, M_\beta) = \delta_{\alpha,\beta}.$$

When the base ring $\mathbf{k}$ is not clear from the context, we write $\operatorname{NSym}_{\mathbf{k}}$ in lieu of NSym.

The Hopf algebra NSym is known as the *Hopf algebra of noncommutative symmetric functions*. Its study goes back to [77].

Theorem 5.4.2. *Letting $H_n := H_{(n)}$ for $n = 0, 1, 2, \ldots$, with $H_0 = 1$, one has that*

$$\operatorname{NSym} \cong \mathbf{k}\langle H_1, H_2, \ldots \rangle, \tag{5.4.1}$$

the free associative (but not commutative) algebra on generators $\{H_1, H_2, \ldots\}$ with coproduct determined by[267]

$$\Delta H_n = \sum_{i+j=n} H_i \otimes H_j. \tag{5.4.2}$$

Proof. Since Proposition 5.1.7 asserts that $\Delta M_\alpha = \sum_{(\beta,\gamma):\beta\cdot\gamma=\alpha} M_\beta \otimes M_\gamma$, and since $\{H_\alpha\}$ are dual to $\{M_\alpha\}$, one concludes that for any compositions β, γ, one has

$$H_\beta H_\gamma = H_{\beta\cdot\gamma}.$$

Iterating this gives

$$H_\alpha = H_{(\alpha_1,\ldots,\alpha_\ell)} = H_{\alpha_1} \cdots H_{\alpha_\ell}. \tag{5.4.3}$$

Since the H_α are a $\mathbf{k}$-basis for NSym, this shows $\operatorname{NSym} \cong \mathbf{k}\langle H_1, H_2, \ldots \rangle$.

Note that $H_n = H_{(n)}$ is dual to $M_{(n)}$, so to understand ΔH_n, one should understand how $M_{(n)}$ can appear as a term in the product $M_\alpha M_\beta$. By (5.1.1) this occurs only if $\alpha = (i), \beta = (j)$ where $i + j = n$, where

$$M_{(i)} M_{(j)} = M_{(i+j)} + M_{(i,j)} + M_{(j,i)}$$

(where the $M_{(i,j)}$ and $M_{(j,i)}$ addends have to be disregarded if one of i and j is 0). By duality, this implies the formula (5.4.2). □

Corollary 5.4.3. *The algebra homomorphism defined by*

$$\begin{array}{rcl} \operatorname{NSym} & \xrightarrow{\pi} & \Lambda, \\ H_n & \longmapsto & h_n \end{array}$$

is a Hopf algebra surjection, and adjoint to the inclusion $\Lambda \overset{i}{\hookrightarrow} \operatorname{QSym}$ (with respect to the dual pairing $\operatorname{NSym} \otimes \operatorname{QSym} \xrightarrow{(\cdot,\cdot)} \mathbf{k}$).

[267] The abbreviated summation indexing $\sum_{i+j=n} t_{i,j}$ used here is intended to mean

$$\sum_{\substack{(i,j) \in \mathbb{N}^2; \\ i+j=n}} t_{i,j}.$$

Proof. As an algebra morphism, π may be identified with the surjection $T(V) \to \operatorname{Sym}(V)$ from the tensor algebra on a graded free $\mathbf{k}$-module V with basis $\{H_1, H_2, \ldots\}$ to the symmetric algebra on V, since

$$\begin{aligned} \operatorname{NSym} &\cong \mathbf{k}\langle H_1, H_2, \ldots \rangle, \\ \Lambda &\cong \mathbf{k}[h_1, h_2, \ldots]. \end{aligned}$$

As (5.4.2) and Proposition 2.3.6(iii) assert that

$$\begin{aligned} \Delta H_n &= \sum_{i+j=n} H_i \otimes H_j, \\ \Delta h_n &= \sum_{i+j=n} h_i \otimes h_j, \end{aligned}$$

this map π is also a bialgebra morphism, and hence a Hopf morphism by Corollary 1.4.27.

To check π is adjoint to i, let $\lambda(\alpha)$ denote the partition which is the weakly decreasing rearrangement of the composition α, and note that the bases $\{H_\alpha\}$ of NSym and $\{m_\lambda\}$ of Λ satisfy

$$(\pi(H_\alpha), m_\lambda) = (h_{\lambda(\alpha)}, m_\lambda) = \left\{ \begin{matrix} 1 & \text{if } \lambda(\alpha) = \lambda \\ 0 & \text{otherwise} \end{matrix} \right\} = \left(H_\alpha, \sum_{\beta : \lambda(\beta) = \lambda} M_\beta \right) = (H_\alpha, i(m_\lambda)).$$

□

Remark 5.4.4. For those who prefer generating functions to sign-reversing involutions, we sketch here Malvenuto and Reutenauer's elegant proof [146, Cor. 2.3] of the antipode formula (Theorem 5.1.11). One needs to know that when $\mathbb{Q}$ is a subring of $\mathbf{k}$, and A is a $\mathbf{k}$-algebra (possibly noncommutative), in the ring of power series $A[[t]]$ where t commutes with all of A, one still has familiar facts, such as

$$a(t) = \log b(t) \quad \text{if and only if} \quad b(t) = \exp a(t)$$

and whenever $a(t), b(t)$ commute in $A[[t]]$, one has

$$\exp\left(a(t) + b(t)\right) = \exp a(t) \exp b(t), \tag{5.4.4}$$

$$\log\left(a(t) b(t)\right) = \log a(t) + \log b(t). \tag{5.4.5}$$

Start by assuming WLOG that $\mathbf{k} = \mathbb{Z}$ (as $\operatorname{NSym}_{\mathbf{k}} = \operatorname{NSym}_{\mathbb{Z}} \otimes_{\mathbb{Z}} \mathbf{k}$ in the general case). Now, define in $\operatorname{NSym}_{\mathbb{Q}} = \operatorname{NSym} \otimes_{\mathbb{Z}} \mathbb{Q}$ the elements $\{\xi_1, \xi_2, \ldots\}$ via generating functions in $\operatorname{NSym}_{\mathbb{Q}}[[t]]$:

$$\begin{aligned} \widetilde{H}(t) &:= \sum_{n \geq 0} H_n t^n, \\ \xi(t) &:= \sum_{n \geq 1} \xi_n t^n = \log \widetilde{H}(t). \end{aligned} \tag{5.4.6}$$

One first checks that this makes each ξ_n primitive, via a computation in the ring $(\operatorname{NSym}_{\mathbb{Q}} \otimes \operatorname{NSym}_{\mathbb{Q}})[[t]]$ (into which we "embed" the ring $(\operatorname{NSym}_{\mathbb{Q}}[[t]]) \otimes_{\mathbb{Q}[[t]]} (\operatorname{NSym}_{\mathbb{Q}}[[t]])$ via the canonical ring homomorphism from the latter into the former [268]):

$$\begin{aligned} \Delta \xi(t) &= \Delta \left(\log \sum_{n \geq 0} H_n t^n \right) = \log \sum_{n \geq 0} \Delta(H_n) t^n = \log \sum_{n \geq 0} \left(\sum_{i+j=n} H_i \otimes H_j \right) t^n \\ &= \log \left(\left(\sum_{i \geq 0} H_i t^i \right) \otimes \left(\sum_{j \geq 0} H_j t^j \right) \right) = \log \left(\left(\sum_{i \geq 0} H_i t^i \otimes 1 \right) \left(1 \otimes \sum_{j \geq 0} H_j t^j \right) \right) \\ &\overset{(5.4.5)}{=} \log \widetilde{H}(t) \otimes 1 + 1 \otimes \log \widetilde{H}(t) = \xi(t) \otimes 1 + 1 \otimes \xi(t). \end{aligned}$$

[268] This ring homomorphism might fail to be injective, whence the "embed" stands in quotation marks. This does not need to worry us, since we will not draw any conclusions in $(\operatorname{NSym}_{\mathbb{Q}}[[t]]) \otimes_{\mathbb{Q}[[t]]} (\operatorname{NSym}_{\mathbb{Q}}[[t]])$ from our computation.

We are also somewhat cavalier with the notation Δ: we use it both for the comultiplication $\Delta : \operatorname{NSym}_{\mathbb{Q}} \to \operatorname{NSym}_{\mathbb{Q}} \otimes \operatorname{NSym}_{\mathbb{Q}}$ of the Hopf algebra $\operatorname{NSym}_{\mathbb{Q}}$ and for the continuous $\mathbf{k}$-algebra homomorphism $\operatorname{NSym}_{\mathbb{Q}}[[t]] \to \left(\operatorname{NSym}_{\mathbb{Q}} \otimes \operatorname{NSym}_{\mathbb{Q}}\right)[[t]]$ it induces.

Comparing coefficients in this equality yields $\Delta(\xi_n) = \xi_n \otimes 1 + 1 \otimes \xi_n$. Thus $S(\xi_n) = -\xi_n$, by Proposition 1.4.17. This allows one to determine $S(H_n)$ and $S(H_\alpha)$, after one first inverts the relation (5.4.6) to get that $\widetilde{H}(t) = \exp \xi(t)$, and hence

$$S(\widetilde{H}(t)) = S(\exp \xi(t)) = \exp S(\xi(t)) = \exp(-\xi(t)) \overset{(5.4.4)}{=} (\exp \xi(t))^{-1}$$
$$= \widetilde{H}(t)^{-1} = \left(1 + H_1 t + H_2 t^2 + \cdots\right)^{-1}.$$

Upon expanding the right side, and comparing coefficients of t^n, this gives

$$S(H_n) = \sum_{\beta \in \mathrm{Comp}_n} (-1)^{\ell(\beta)} H_\beta$$

and hence

$$S(H_\alpha) = S(H_{\alpha_\ell}) \cdots S(H_{\alpha_2}) S(H_{\alpha_1}) = \sum_{\substack{\gamma: \\ \gamma \text{ refines } \operatorname{rev}(\alpha)}} (-1)^{\ell(\gamma)} H_\gamma = \sum_{\substack{\gamma: \\ \operatorname{rev}(\gamma) \text{ refines } \alpha}} (-1)^{\ell(\gamma)} H_\gamma$$

(because if μ and ν are two compositions, then μ refines ν if and only if $\operatorname{rev}(\mu)$ refines $\operatorname{rev}(\nu)$). As $S_{\mathrm{NSym}}, S_{\mathrm{QSym}}$ are adjoint, and $\{H_\alpha\}, \{M_\alpha\}$ are dual bases, this is equivalent to saying that

$$S(M_\alpha) = (-1)^{\ell(\alpha)} \sum_{\substack{\gamma: \\ \operatorname{rev}(\alpha) \text{ refines } \gamma}} M_\gamma \qquad \text{for all } \alpha \in \mathrm{Comp}.$$

But this is precisely the claim of Theorem 5.1.11. Thus, Theorem 5.1.11 is proven once again.

Let us say a bit more about the elements ξ_n defined in (5.4.6) above. The elements $n\xi_n$ are noncommutative analogues of the power sum symmetric functions p_n (and, indeed, are lifts of the latter to NSym, as Exercise 5.4.5 below shows). They are called the *noncommutative power sums of the second kind* in [77][269], and their products form a basis of NSym. They are furthermore useful in studying the so-called *Eulerian idempotent* of a cocommutative Hopf algebra, as shown in Exercise 5.4.6 below.

Exercise 5.4.5. Assume that $\mathbb{Q}$ is a subring of $\mathbf{k}$. Define a sequence of elements $\xi_1, \xi_2, \xi_3, \ldots$ of $\mathrm{NSym} = \mathrm{NSym}_{\mathbf{k}}$ by (5.4.6).

(a) For every $n \geq 1$, show that ξ_n is a primitive homogeneous element of NSym of degree n.
(b) For every $n \geq 1$, show that $\pi(n\xi_n)$ is the n-th power sum symmetric function $p_n \in \Lambda$.
(c) For every $n \geq 1$, show that

$$\xi_n = \sum_{\alpha \in \mathrm{Comp}_n} (-1)^{\ell(\alpha)-1} \frac{1}{\ell(\alpha)} H_\alpha. \tag{5.4.7}$$

(d) For every composition α, define an element ξ_α of NSym by $\xi_\alpha = \xi_{\alpha_1} \xi_{\alpha_2} \cdots \xi_{\alpha_\ell}$, where α is written in the form $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_\ell)$ with $\ell = \ell(\alpha)$. Show that

$$H_n = \sum_{\alpha \in \mathrm{Comp}_n} \frac{1}{\ell(\alpha)!} \xi_\alpha \tag{5.4.8}$$

for every $n \in \mathbb{N}$.

Use this to prove that $(\xi_\alpha)_{\alpha \in \mathrm{Comp}_n}$ is a $\mathbf{k}$-basis of NSym_n for every $n \in \mathbb{N}$.

Exercise 5.4.6. Assume that $\mathbb{Q}$ is a subring of $\mathbf{k}$. Let A be a cocommutative connected graded $\mathbf{k}$-bialgebra. Let $A = \bigoplus_{n \geq 0} A_n$ be the decomposition of A into homogeneous components. If f is any $\mathbf{k}$-linear map $A \to A$ annihilating A_0, then f is locally $\star$-nilpotent[270], and so the sum $\log^\star(f + u\epsilon) := \sum_{n \geq 1} (-1)^{n-1} \frac{1}{n} f^{\star n}$ is a well-defined endomorphism of A [271]. Let $\mathfrak{e}$ denote the endomorphism $\log^\star(\mathrm{id}_A)$ of A (obtained by setting $f = \mathrm{id}_A - u\epsilon : A \to A$). Show that $\mathfrak{e}$ is a projection from A to the $\mathbf{k}$-submodule $\mathfrak{p}$ of all primitive elements of A (and thus, in particular, is idempotent).

[269] See Exercise 5.4.12 for the ones of the first kind.

[270] See the proof of Proposition 1.4.24 for what this means.

[271] This definition of $\log^\star(f + u\epsilon)$ is actually a particular case of Definition 1.7.17. This can be seen as follows:

We have $f(A_0) = 0$. Thus, Proposition 1.7.11(h) (applied to $C = A$) yields $f \in \mathfrak{n}(A, A)$ (where $\mathfrak{n}(A, A)$ is defined as in Section 1.7), so that $(f + u\epsilon) - u\epsilon = f \in \mathfrak{n}(A, A)$. Therefore, Definition 1.7.17 defines a map $\log^\star(f + u\epsilon) \in \mathfrak{n}(A, A)$. This map is identical to the map $\log^\star(f + u\epsilon) := \sum_{n \geq 1} (-1)^{n-1} \frac{1}{n} f^{\star n}$ we have just defined, because Proposition 1.7.18(f) (applied

Hint: For every $n \geq 0$, let $\pi_n : A \to A$ be the projection onto the n-th homogeneous component A_n. Since NSym is the free $\mathbf{k}$-algebra with generators $H_1, H_2, H_3, \ldots$, we can define a $\mathbf{k}$-algebra homomorphism $\mathfrak{W} : \mathrm{NSym} \to (\mathrm{End}\, A, \star)$ by sending H_n to π_n. Show that:

(a) The map $\mathfrak{e} : A \to A$ is graded. For every $n \geq 0$, we will denote the map $\pi_n \circ \mathfrak{e} = \mathfrak{e} \circ \pi_n : A \to A$ by $\mathfrak{e}_n$.

(b) We have $\mathfrak{W}(\xi_n) = \mathfrak{e}_n$ for all $n \geq 1$, where ξ_n is defined as in Exercise 5.4.5.

(c) If w is an element of NSym, and if we write $\Delta(w) = \sum_{(w)} w_1 \otimes w_2$ using the Sweedler notation, then $\Delta \circ (\mathfrak{W}(w)) = \left(\sum_{(w)} \mathfrak{W}(w_1) \otimes \mathfrak{W}(w_2)\right) \circ \Delta$.

(d) We have $\mathfrak{e}_n(A) \subset \mathfrak{p}$ for every $n \geq 0$.

(e) We have $\mathfrak{e}(A) \subset \mathfrak{p}$.

(f) The map $\mathfrak{e}$ fixes any element of $\mathfrak{p}$.

Remark 5.4.7. The endomorphism $\mathfrak{e}$ of Exercise 5.4.6 is known as the *Eulerian idempotent* of A, and can be contrasted with the Dynkin idempotent of Remark 1.5.15. It has been studied in [166], [169], [31] and [60], and relates to the Hochschild cohomology of commutative algebras [134, §4.5.2].

Exercise 5.4.8. Assume that $\mathbb{Q}$ is a subring of $\mathbf{k}$. Let A, A_n and $\mathfrak{e}$ be as in Exercise 5.4.6.

(a) Show that $\mathfrak{e}^{\star n} \circ \mathfrak{e}^{\star m} = n!\delta_{n,m}\mathfrak{e}^{\star n}$ for all $n \in \mathbb{N}$ and $m \in \mathbb{N}$.

(b) Show that $\mathfrak{e}^{\star n} \circ \mathrm{id}_A^{\star m} = \mathrm{id}_A^{\star m} \circ \mathfrak{e}^{\star n} = m^n \mathfrak{e}^{\star n}$ for all $n \in \mathbb{N}$ and $m \in \mathbb{N}$.

We next explore the basis for NSym dual to the $\{L_\alpha\}$ in QSym.

Definition 5.4.9. Define the *noncommutative ribbon functions* $\{R_\alpha\}_{\alpha \in \mathrm{Comp}}$ to be the $\mathbf{k}$-basis of NSym dual to the fundamental basis $\{L_\alpha\}_{\alpha \in \mathrm{Comp}}$ of QSym, so that

$$(R_\alpha, L_\beta) = \delta_{\alpha,\beta} \qquad \text{for all } \alpha, \beta \in \mathrm{Comp}.$$

Theorem 5.4.10. (a) *One has that*

$$H_\alpha = \sum_{\beta \text{ coarsens } \alpha} R_\beta; \tag{5.4.9}$$

$$R_\alpha = \sum_{\beta \text{ coarsens } \alpha} (-1)^{\ell(\beta)-\ell(\alpha)} H_\beta. \tag{5.4.10}$$

(b) *The surjection* NSym $\xrightarrow{\pi} \Lambda$ *sends* $R_\alpha \longmapsto s_{\mathrm{Rib}(\alpha)}$, *the skew Schur function associated to the ribbon* $\mathrm{Rib}(\alpha)$.

(c) *Furthermore,*

$$R_\alpha R_\beta = R_{\alpha \cdot \beta} + R_{\alpha \odot \beta} \qquad \textit{if } \alpha \textit{ and } \beta \textit{ are nonempty}; \tag{5.4.11}$$

$$S(R_\alpha) = (-1)^{|\alpha|} R_{\omega(\alpha)}. \tag{5.4.12}$$

Finally, $R_\varnothing$ *is the multiplicative identity of* NSym.

Proof. (a) For (5.4.9), note that

$$H_\alpha = \sum_\beta (H_\alpha, L_\beta) R_\beta = \sum_\beta \left(H_\alpha, \sum_{\substack{\gamma: \\ \gamma \text{ refines } \beta}} M_\gamma \right) R_\beta = \sum_{\substack{\beta: \\ \beta \text{ coarsens } \alpha}} R_\beta.$$

The equality (5.4.10) follows from (5.4.9) by Lemma 5.2.7(a).

(b) Write α as $(\alpha_1, \ldots, \alpha_\ell)$. To show that $\pi(R_\alpha) = s_{\mathrm{Rib}(\alpha)}$, we instead examine $\pi(H_\alpha)$:

$$\pi(H_\alpha) = \pi(H_{\alpha_1} \cdots H_{\alpha_\ell}) = h_{\alpha_1} \cdots h_{\alpha_\ell} = s_{(\alpha_1)} \cdots s_{(\alpha_\ell)} = s_{(\alpha_1) \oplus \cdots \oplus (\alpha_\ell)}$$

to $C = A$) shows that the map $\log^\star(f + u\epsilon)$ defined using Definition 1.7.17 satisfies

$$\log^\star(f + u\epsilon) = \sum_{n \geq 1} \frac{(-1)^{n-1}}{n} f^{\star n} = \sum_{n \geq 1} (-1)^{n-1} \frac{1}{n} f^{\star n}.$$

where $(\alpha_1) \oplus \cdots \oplus (\alpha_\ell)$ is some skew shape which is a horizontal strip having rows of lengths $\alpha_1, \ldots, \alpha_\ell$ from bottom to top. We claim

$$s_{(\alpha_1)\oplus\cdots\oplus(\alpha_\ell)} = \sum_{\substack{\beta: \\ \beta \text{ coarsens } \alpha}} s_{\mathrm{Rib}(\beta)},$$

because column-strict tableaux T of shape $(\alpha_1) \oplus \cdots \oplus (\alpha_\ell)$ biject to column-strict tableaux T' of some ribbon $\mathrm{Rib}\,(\beta)$ with β coarsening α, as follows: Let a_i, b_i denote the leftmost, rightmost entries of the i-th row from the bottom in T, of length α_i, and

- if $b_i \leq a_{i+1}$, merge parts α_i, α_{i+1} in β, and concatenate the rows of length α_i, α_{i+1} in T', or
- if $b_i > a_{i+1}$, do not merge parts α_i, α_{i+1} in β, and let these two rows overlap in one column in T'.

E.g., if $\alpha = (3, 3, 2, 3, 2)$, then

$$\text{the tableau } T = \begin{array}{ccccccccccccc} &&&&&&&&&&& 3 & 4 \\ &&&&&&&& 4 & 4 & 5 && \\ &&&&&& 4 & 4 &&&&& \\ &&& 2 & 2 & 3 &&&&&&& \\ 1 & 1 & 3 &&&&&&&&&& \end{array} \quad \text{of shape } (\alpha_1) \oplus \cdots \oplus (\alpha_\ell)$$

$$\text{maps to the tableau } T' = \begin{array}{ccccccccccc} &&&&&&&&& 3 & 4 \\ && 2 & 2 & 3 & 4 & 4 & 4 & 4 & 5 & \\ 1 & 1 & 3 &&&&&&&& \end{array} \quad \text{of shape } \mathrm{Rib}\,(\beta) \text{ for } \beta = (3, 8, 2).$$

The reverse bijection breaks the rows of T' into the rows of T of lengths dictated by the parts of α. Having shown $\pi(H_\alpha) = \sum_{\beta:\beta \text{ coarsens } \alpha} s_{\mathrm{Rib}(\beta)}$, we can now apply Lemma 5.2.7(a) to obtain

$$s_{\mathrm{Rib}(\alpha)} = \sum_{\beta:\beta \text{ coarsens } \alpha} (-1)^{\ell(\alpha)-\ell(\beta)} \pi(H_\beta) = \pi(R_\alpha) \qquad (\text{by } (5.4.10))\,;$$

thus, $\pi(R_\alpha) = s_{\mathrm{Rib}(\alpha)}$ is proven.

(c) Finally, (5.4.11) and (5.4.12) follow from (5.2.5) and (5.2.7) by duality. □

Remark 5.4.11. Since the maps

$$\begin{tikzcd} \mathrm{NSym} \arrow[dr, "\pi"'] & & \mathrm{QSym} \\ & \Lambda \arrow[ur, hook, "i"'] & \end{tikzcd}$$

are Hopf morphisms, they must respect the antipodes $S_\Lambda, S_{\mathrm{QSym}}, S_{\mathrm{NSym}}$, but it is interesting to compare them explicitly using the fundamental basis for QSym and the ribbon basis for NSym.

On one hand (5.2.7) shows that $S_{\mathrm{QSym}}(L_\alpha) = (-1)^{|\alpha|} L_{\omega(\alpha)}$ extends the map S_Λ since $L_{(1^n)} = e_n$ and $L_{(n)} = h_n$, as observed in Example 5.2.5, and $\omega((n)) = (1^n)$.

On the other hand, (5.4.12) shows that $S_{\mathrm{NSym}}(R_\alpha) = (-1)^{|\alpha|} R_{\omega(\alpha)}$ lifts the map S_Λ to S_{NSym}: Theorem 5.4.10(b) showed that R_α lifts the skew Schur function $s_{\mathrm{Rib}(\alpha)}$, while (2.4.15) asserted that $S(s_{\lambda/\mu}) = (-1)^{|\lambda/\mu|} s_{\lambda^t/\mu^t}$, and a ribbon $\mathrm{Rib}\,(\alpha) = \lambda/\mu$ has $\mathrm{Rib}\,(\omega(\alpha)) = \lambda^t/\mu^t$.

Exercise 5.4.12. (a) Show that any integers n and i with $0 \leq i < n$ satisfy

$$R_{(1^i, n-i)} = \sum_{j=0}^{i} (-1)^{i-j} R_{(1^j)} H_{n-j}.$$

(Here, as usual, 1^i stands for the number 1 repeated i times.)

(b) Show that any integers n and i with $0 \leq i < n$ satisfy

$$(-1)^i R_{(1^i, n-i)} = \sum_{j=0}^{i} S(H_j) H_{n-j}.$$

(c) For every positive integer n, define an element Ψ_n of NSym by

$$\Psi_n = \sum_{i=0}^{n-1} (-1)^i R_{(1^i,n-i)}.$$

Show that $\Psi_n = (S \star E)(H_n)$, where the map $E : \operatorname{NSym} \to \operatorname{NSym}$ is defined as in Exercise 1.5.14 (for $A = \operatorname{NSym}$). Conclude that Ψ_n is primitive.

(d) Prove that

$$\sum_{k=0}^{n-1} H_k \Psi_{n-k} = nH_n$$

for every $n \in \mathbb{N}$.

(e) Define two power series $\psi(t)$ and $\widetilde{H}(t)$ in $\operatorname{NSym}[[t]]$ by

$$\psi(t) = \sum_{n\geq 1} \Psi_n t^{n-1};$$
$$\widetilde{H}(t) = \sum_{n\geq 0} H_n t^n.$$

Show that[272] $\dfrac{d}{dt}\widetilde{H}(t) = \widetilde{H}(t) \cdot \psi(t)$.

(The functions Ψ_n are called *noncommutative power sums of the first kind*; they are studied in [77]. The power sums of the second kind are the $n\xi_n$ in Remark 5.4.4.)

(f) Show that $\pi(\Psi_n)$ equals the power sum symmetric function p_n for every positive integer n.

(g) Show that every positive integer n satisfies

$$p_n = \sum_{i=0}^{n-1} (-1)^i s_{(n-i,1^i)} \qquad \text{in } \Lambda.$$

(h) For every nonempty composition α, define a positive integer $\operatorname{lp}(\alpha)$ by $\operatorname{lp}(\alpha) = \alpha_\ell$, where α is written in the form $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_\ell)$ with $\ell = \ell(\alpha)$. (Thus, $\operatorname{lp}(\alpha)$ is the last part of α.) Show that every positive integer n satisfies

$$\Psi_n = \sum_{\alpha\in\operatorname{Comp}_n} (-1)^{\ell(\alpha)-1} \operatorname{lp}(\alpha) H_\alpha. \tag{5.4.13}$$

(i) Assume that $\mathbb{Q}$ is a subring of $\mathbf{k}$. For every composition α, define an element Ψ_α of NSym by $\Psi_\alpha = \Psi_{\alpha_1}\Psi_{\alpha_2}\cdots\Psi_{\alpha_\ell}$, where α is written in the form $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_\ell)$ with $\ell = \ell(\alpha)$. For every composition α, define $\pi_u(\alpha)$ to be the positive integer $\alpha_1(\alpha_1+\alpha_2)\cdots(\alpha_1+\alpha_2+\cdots+\alpha_\ell)$, where α is written in the form $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_\ell)$ with $\ell = \ell(\alpha)$. Show that

$$H_n = \sum_{\alpha\in\operatorname{Comp}_n} \frac{1}{\pi_u(\alpha)} \Psi_\alpha \tag{5.4.14}$$

for every $n \in \mathbb{N}$.

Use this to prove that $(\Psi_\alpha)_{\alpha\in\operatorname{Comp}_n}$ is a $\mathbf{k}$-basis of NSym_n for every $n \in \mathbb{N}$.

(j) Assume that $\mathbb{Q}$ is a subring of $\mathbf{k}$. Let V be the free $\mathbf{k}$-module with basis $(\mathfrak{b}_n)_{n\in\{1,2,3,\ldots\}}$. Define a $\mathbf{k}$-module homomorphism $f : V \to \operatorname{NSym}$ by requiring that $f(\mathfrak{b}_n) = \Psi_n$ for every $n \in \{1, 2, 3, \ldots\}$. Let F be the $\mathbf{k}$-algebra homomorphism $T(V) \to \operatorname{NSym}$ induced by this f (using the universal property of the tensor algebra $T(V)$). Show that F is a Hopf algebra isomorphism (where the Hopf algebra structure on $T(V)$ is as in Example 1.4.18).

(k) Assume that $\mathbb{Q}$ is a subring of $\mathbf{k}$. Let V be as in Exercise 5.4.12(j). Show that QSym is isomorphic to the shuffle algebra $\operatorname{Sh}(V)$ (defined as in Proposition 1.6.7) as Hopf algebras.

(l) Solve parts (a) and (b) of Exercise 2.9.14 again using the ribbon basis functions R_α.

[272] The derivative $\frac{d}{dt}Q(t)$ of a power series $Q(t) \in R[[t]]$ over a noncommutative ring R is defined just as in the case of R commutative: by setting $\frac{d}{dt}Q(t) = \sum_{i\geq 1} iq_i t^{i-1}$, where $Q(t)$ is written in the form $Q(t) = \sum_{i\geq 0} q_i t^i$.

One might wonder whether the Frobenius endomorphisms of Λ (defined in Exercise 2.9.9) and the Verschiebung endomorphisms of Λ (defined in Exercise 2.9.10) generalize to analogous operators on either QSym or NSym. The next two exercises (whose claims mostly come from [90, §13]) answer this question: The Frobenius endomorphisms extend to QSym, and the Verschiebung ones lift to NSym.

Exercise 5.4.13. For every $n \in \{1, 2, 3, \ldots\}$, define a map $\mathbf{F}_n : \operatorname{QSym} \to \operatorname{QSym}$ by setting

$$\mathbf{F}_n(a) = a\left(x_1^n, x_2^n, x_3^n, \ldots\right) \qquad \text{for every } a \in \operatorname{QSym}.$$

(So what $\mathbf{F}_n$ does to a quasi-symmetric function is replacing all variables $x_1, x_2, x_3, \ldots$ by their n-th powers.)

(a) Show that $\mathbf{F}_n : \operatorname{QSym} \to \operatorname{QSym}$ is a $\mathbf{k}$-algebra homomorphism for every $n \in \{1, 2, 3, \ldots\}$.
(b) Show that $\mathbf{F}_n \circ \mathbf{F}_m = \mathbf{F}_{nm}$ for any two positive integers n and m.
(c) Show that $\mathbf{F}_1 = \operatorname{id}$.
(d) Prove that $\mathbf{F}_n\left(M_{(\beta_1,\beta_2,\ldots,\beta_s)}\right) = M_{(n\beta_1,n\beta_2,\ldots,n\beta_s)}$ for every $n \in \{1, 2, 3, \ldots\}$ and $(\beta_1, \beta_2, \ldots, \beta_s) \in \operatorname{Comp}$.
(e) Prove that $\mathbf{F}_n : \operatorname{QSym} \to \operatorname{QSym}$ is a Hopf algebra homomorphism for every $n \in \{1, 2, 3, \ldots\}$.
(f) Consider the maps $\mathbf{f}_n : \Lambda \to \Lambda$ defined in Exercise 2.9.9. Show that $\mathbf{F}_n \mid_{\Lambda} = \mathbf{f}_n$ for every $n \in \{1, 2, 3, \ldots\}$.
(g) Assume that $\mathbf{k} = \mathbb{Z}$. Prove that $\mathbf{f}_p(a) \equiv a^p \bmod p \operatorname{QSym}$ for every $a \in \operatorname{QSym}$ and every prime number p.
(h) Give a new solution to Exercise 2.9.9(d).

Exercise 5.4.14. For every $n \in \{1, 2, 3, \ldots\}$, define a $\mathbf{k}$-algebra homomorphism $\mathbf{V}_n : \operatorname{NSym} \to \operatorname{NSym}$ by

$$\mathbf{V}_n(H_m) = \begin{cases} H_{m/n}, & \text{if } n \mid m; \\ 0, & \text{if } n \nmid m \end{cases} \qquad \text{for every positive integer } m$$

[273].

(a) Show that any positive integers n and m satisfy

$$\mathbf{V}_n(\Psi_m) = \begin{cases} n\Psi_{m/n}, & \text{if } n \mid m; \\ 0, & \text{if } n \nmid m \end{cases},$$

where the elements Ψ_m and $\Psi_{m/n}$ of NSym are as defined in Exercise 5.4.12(c).
(b) Show that if $\mathbb{Q}$ is a subring of $\mathbf{k}$, then any positive integers n and m satisfy

$$\mathbf{V}_n(\xi_m) = \begin{cases} \xi_{m/n}, & \text{if } n \mid m; \\ 0, & \text{if } n \nmid m \end{cases},$$

where the elements ξ_m and $\xi_{m/n}$ of NSym are as defined in Exercise 5.4.5.
(c) Prove that $\mathbf{V}_n \circ \mathbf{V}_m = \mathbf{V}_{nm}$ for any two positive integers n and m.
(d) Prove that $\mathbf{V}_1 = \operatorname{id}$.
(e) Prove that $\mathbf{V}_n : \operatorname{NSym} \to \operatorname{NSym}$ is a Hopf algebra homomorphism for every $n \in \{1, 2, 3, \ldots\}$.

Now, consider also the maps $\mathbf{F}_n : \operatorname{QSym} \to \operatorname{QSym}$ defined in Exercise 2.9.9. Fix a positive integer n.

(f) Prove that the maps $\mathbf{F}_n : \operatorname{QSym} \to \operatorname{QSym}$ and $\mathbf{V}_n : \operatorname{NSym} \to \operatorname{NSym}$ are adjoint with respect to the dual pairing $\operatorname{NSym} \otimes \operatorname{QSym} \xrightarrow{(\cdot,\cdot)} \mathbf{k}$.
(g) Consider the maps $\mathbf{v}_n : \Lambda \to \Lambda$ defined in Exercise 2.9.10. Show that the surjection $\pi : \operatorname{NSym} \to \Lambda$ satisfies $\mathbf{v}_n \circ \pi = \pi \circ \mathbf{V}_n$ for every $n \in \{1, 2, 3, \ldots\}$.
(h) Give a new solution to Exercise 2.9.10(f).

[273] This is well-defined, since NSym is (isomorphic to) the free associative algebra with generators $H_1, H_2, H_3, \ldots$ (according to (5.4.1)).

6. Polynomial generators for QSym and Lyndon words

In this chapter, we shall construct an algebraically independent generating set for QSym as a $\mathbf{k}$-algebra, thus showing that QSym is a polynomial ring over $\mathbf{k}$. This has been done by Malvenuto [145, Cor. 4.19] when $\mathbf{k}$ is a field of characteristic 0, and by Hazewinkel [89] in the general case. We will begin by introducing the notion of *Lyndon words* (Section 6.1), on which both of these constructions rely; we will then (Section 6.2) elucidate the connection of Lyndon words with shuffles, and afterwards (Section 6.3) apply it to prove *Radford's theorem* stating that the shuffle algebra of a free $\mathbf{k}$-module over a commutative $\mathbb{Q}$-algebra is a polynomial algebra (Theorem 6.3.4). The shuffle algebra is not yet QSym, but Radford's theorem on the shuffle algebra serves as a natural stepping stone for the study of the more complicated algebra QSym. We will prove – in two ways – that QSym is a polynomial algebra when $\mathbb{Q}$ is a subring of $\mathbf{k}$ in Section 6.4, and then we will finally prove the general case in Section 6.5. In Section 6.6, we will explore a different aspect of the combinatorics of words: the notion of necklaces (which are in bijection with Lyndon words, as Exercise 6.1.34 will show) and the *Gessel-Reutenauer bijection*, which help us define and understand the *Gessel-Reutenauer symmetric functions*. This will rely on Section 6.1, but not on any of the other sections of Chapter 6.

Strictly speaking, this whole Chapter 6 is a digression, as it involves almost no coalgebraic or Hopf-algebraic structures, and its results will not be used in further chapters (which means it can be skipped if so desired). However, it sheds additional light on both quasisymmetric and symmetric functions, and serves as an excuse to study Lyndon words, which are a combinatorial object of independent interest (and are involved in the study of free algebras and Hopf algebras, apart from QSym – see [177] and [182][274]).

We will take a scenic route to the proof of Hazewinkel's theorem. A reader only interested in the proof proper can restrict themselves to reading only the following:

- from Section 6.1, everything up to Corollary 6.1.6, then from Definition 6.1.13 up to Proposition 6.1.18, then from Definition 6.1.25 up to Lemma 6.1.28, and finally Theorem 6.1.30. (Proposition 6.1.19 and Theorem 6.1.20 are also relevant if one wants to use a different definition of Lyndon words, as they prove the equivalence of most such definitions.)
- from Section 6.2, everything except for Exercise 6.2.25.
- from Section 6.3, Definition 6.3.1, Lemma 6.3.7, and Lemma 6.3.10.
- from Section 6.4, Definition 6.4.1, Theorem 6.4.3, then from Proposition 6.4.5 up to Definition 6.4.9, and Lemma 6.4.11.
- all of Section 6.5.

Likewise, Section 6.6 can be read immediately after Section 6.1.

6.1. Lyndon words.

Lyndon words have been independently defined by Shirshov [202], Lyndon [141], Radford [177, §2] and de Bruijn/Klarner [29] (though using different and sometimes incompatible notations). They have since been surfacing in various places in noncommutative algebra (particularly the study of free Lie algebras); expositions of their theory can be found in [139, §5], [182, §5.1] and [124, §1] (in German). We will follow our own approach to the properties of Lyndon words that we need.

Definition 6.1.1. We fix a totally ordered set $\mathfrak{A}$, which we call the *alphabet*. Throughout Section 6.1 and Section 6.2, we will understand "word" to mean a word over $\mathfrak{A}$.

We recall that a *word* is just a (finite) tuple of elements of $\mathfrak{A}$. In other words, a word is an element of the set $\bigsqcup_{n\geq 0} \mathfrak{A}^n$. We denote this set by $\mathfrak{A}^*$.

The *empty word* is the unique tuple with 0 elements. It is denoted by $\varnothing$. If $w \in \mathfrak{A}^n$ is a word and $i \in \{1, 2, \ldots, n\}$, then the *i-th letter* of w means the i-th entry of the n-tuple w. This i-th letter will be denoted by w_i.

The *length* $\ell(w)$ of a word $w \in \bigsqcup_{n\geq 0} \mathfrak{A}^n$ is defined to be the $n \in \mathbb{N}$ satisfying $w \in \mathfrak{A}^n$. Thus, $w = (w_1, w_2, \ldots, w_{\ell(w)})$ for every word w.

Given two words u and v, we say that u is *longer* than v (or, equivalently, v is *shorter* than u) if and only if $\ell(u) > \ell(v)$.

The *concatenation* of two words u and v is defined to be the word $(u_1, u_2, \ldots, u_{\ell(u)}, v_1, v_2, \ldots, v_{\ell(v)})$. This concatenation is denoted by uv or $u \cdot v$. The set $\mathfrak{A}^*$ of all words is a monoid with respect to concatenation,

[274] They also are involved in indexing basis elements of combinatorial Hopf algebras other than QSym. See Bergeron/Zabrocki [18].

with neutral element $\varnothing$. It is precisely the free monoid on generators $\mathfrak{A}$. If u is a word and $i \in \mathbb{N}$, we will understand u^i to mean the i-th power of u in this monoid (that is, the word $\underbrace{uu \cdots u}_{i \text{ times}}$).

The elements of $\mathfrak{A}$ are called *letters*, and will be identified with elements of $\mathfrak{A}^1 \subset \bigsqcup_{n \geq 0} \mathfrak{A}^n = \mathfrak{A}^*$. This identification equates every letter $u \in \mathfrak{A}$ with the one-letter word $(u) \in \mathfrak{A}^1$. Thus, every word $(u_1, u_2, \ldots, u_n) \in \mathfrak{A}^*$ equals the concatenation $u_1 u_2 \cdots u_n$ of letters, hence allowing us to use $u_1 u_2 \cdots u_n$ as a brief notation for the word $(u_1, u_2, \ldots, u_n)$.

If w is a word, then:

- a *prefix* of w means a word of the form $(w_1, w_2, \ldots, w_i)$ for some $i \in \{0, 1, \ldots, \ell(w)\}$;
- a *suffix* of w means a word of the form $\left(w_{i+1}, w_{i+2}, \ldots, w_{\ell(w)}\right)$ for some $i \in \{0, 1, \ldots, \ell(w)\}$;
- a *proper suffix* of w means a word of the form $\left(w_{i+1}, w_{i+2}, \ldots, w_{\ell(w)}\right)$ for some $i \in \{1, 2, \ldots, \ell(w)\}$.

In other words,

- a *prefix* of $w \in \mathfrak{A}^*$ is a word $u \in \mathfrak{A}^*$ such that there exists a $v \in \mathfrak{A}^*$ satisfying $w = uv$;
- a *suffix* of $w \in \mathfrak{A}^*$ is a word $v \in \mathfrak{A}^*$ such that there exists a $u \in \mathfrak{A}^*$ satisfying $w = uv$;
- a *proper suffix* of $w \in \mathfrak{A}^*$ is a word $v \in \mathfrak{A}^*$ such that there exists a nonempty $u \in \mathfrak{A}^*$ satisfying $w = uv$.

Clearly, any proper suffix of $w \in \mathfrak{A}^*$ is a suffix of w. Moreover, if $w \in \mathfrak{A}^*$ is any word, then a proper suffix of w is the same thing as a suffix of w distinct from w.

We define a relation $\leq$ on the set $\mathfrak{A}^*$ as follows: For two words $u \in \mathfrak{A}^*$ and $v \in \mathfrak{A}^*$, we set $u \leq v$ to hold if and only if

$$\begin{aligned} &\textbf{either } \text{there exists an } i \in \{1, 2, \ldots, \min\{\ell(u), \ell(v)\}\} \\ &\qquad \text{such that } (u_i < v_i, \text{ and every } j \in \{1, 2, \ldots, i-1\} \text{ satisfies } u_j = v_j), \\ &\textbf{or } \text{the word } u \text{ is a prefix of } v. \end{aligned}$$

This order relation (taken as the smaller-or-equal relation) makes $\mathfrak{A}^*$ into a poset (by Proposition 6.1.2(a) below), and we will always be regarding $\mathfrak{A}^*$ as endowed with this poset structure (thus, notations such as $<$, $\leq$, $>$ and $\geq$ will be referring to this poset structure). This poset is actually totally ordered (see Proposition 6.1.2(a)).

Here are some examples of words compared by the relation $\leq$:

$$\begin{array}{cccc} 113 \leq 114, & 113 \leq 132, & 19 \leq 195, & 41 \leq 412, \\ 41 \leq 421, & 539 \leq 54, & \varnothing \leq 21, & \varnothing \leq \varnothing \end{array}$$

(where $\mathfrak{A}$ is the alphabet $\{1 < 2 < 3 < \cdots\}$).

Notice that if u and v are two words of the same length (i.e., we have $u, v \in \mathfrak{A}^n$ for one and the same n), then $u \leq v$ holds if and only if u is lexicographically smaller-or-equal to v. In other words, the relation $\leq$ is an extension of the lexicographic order on every $\mathfrak{A}^n$ to $\mathfrak{A}^*$. This is the reason why this relation $\leq$ is usually called the *lexicographic order* on $\mathfrak{A}^*$. In particular, we will be using this name.[275] However, unlike the lexicographic order on $\mathfrak{A}^n$, it does not always respect concatenation from the right: It can happen that $u, v, w \in \mathfrak{A}^*$ satisfy $u \leq v$ but not $uw \leq vw$. (For example, $u = 1$, $v = 13$ and $w = 4$, again with $\mathfrak{A} = \{1 < 2 < 3 < \cdots\}$.) We will see in Proposition 6.1.2 that this is rather an exception than the rule and the relation $\leq$ still behaves mostly predictably with respect to concatenation.

Some basic properties of the order relation $\leq$ just defined are collected in the following proposition:

Proposition 6.1.2. (a) *The order relation $\leq$ is (the smaller-or-equal relation of) a total order on the set $\mathfrak{A}^*$.*

(b) *If $a, c, d \in \mathfrak{A}^*$ satisfy $c \leq d$, then $ac \leq ad$.*

(c) *If $a, c, d \in \mathfrak{A}^*$ satisfy $ac \leq ad$, then $c \leq d$.*

(d) *If $a, b, c, d \in \mathfrak{A}^*$ satisfy $a \leq c$, then either we have $ab \leq cd$ or the word a is a prefix of c.*

(e) *If $a, b, c, d \in \mathfrak{A}^*$ satisfy $ab \leq cd$, then either we have $a \leq c$ or the word c is a prefix of a.*

(f) *If $a, b, c, d \in \mathfrak{A}^*$ satisfy $ab \leq cd$ and $\ell(a) \leq \ell(c)$, then $a \leq c$.*

[275] The relation $\leq$ is also known as the *dictionary order*, due to the fact that it is the order in which words appear in a dictionary.

(g) *If $a, b, c \in \mathfrak{A}^*$ satisfy $a \leq b \leq ac$, then a is a prefix of b.*
(h) *If $a \in \mathfrak{A}^*$ is a prefix of $b \in \mathfrak{A}^*$, then $a \leq b$.*
(i) *If a and b are two prefixes of $c \in \mathfrak{A}^*$, then either a is a prefix of b, or b is a prefix of a.*
(j) *If $a, b, c \in \mathfrak{A}^*$ are such that $a \leq b$ and $\ell(a) \geq \ell(b)$, then $ac \leq bc$.*
(k) *If $a \in \mathfrak{A}^*$ and $b \in \mathfrak{A}^*$ are such that b is nonempty, then $a < ab$.*

Exercise 6.1.3. Prove Proposition 6.1.2.

[**Hint:** No part of Proposition 6.1.2 requires more than straightforward case analysis. However, the proof of (a) can be simplified by identifying the order relation $\leq$ on $\mathfrak{A}^*$ as a restriction of the lexicographic order on the set $\mathfrak{B}^\infty$, where $\mathfrak{B}$ is a suitable extension of the alphabet $\mathfrak{A}$. What is this extension, and how to embed $\mathfrak{A}^*$ into $\mathfrak{B}^\infty$?]

Proposition 6.1.2 provides a set of tools for working with the lexicographic order without having to refer to its definition; we shall use it extensively. Proposition 6.1.2(h) (and its equivalent form stating that $a \leq ac$ for every $a \in \mathfrak{A}^*$ and $c \in \mathfrak{A}^*$) and Proposition 6.1.2(k) will often be used without explicit mention.

Before we define Lyndon words, let us show two more facts about words which will be used later. First, when do words commute?

Proposition 6.1.4. *Let $u, v \in \mathfrak{A}^*$ satisfy $uv = vu$. Then, there exist a $t \in \mathfrak{A}^*$ and two nonnegative integers n and m such that $u = t^n$ and $v = t^m$.*

Proof. We prove this by strong induction on $\ell(u) + \ell(v)$. We assume WLOG that $\ell(u)$ and $\ell(v)$ are positive (because otherwise, one of u and v is the empty word, and everything is trivial). It is easy to see that either u is a prefix of v, or v is a prefix of u [276]. We assume WLOG that u is a prefix of v (since our situation is symmetric). Thus, we can write v in the form $v = uw$ for some $w \in \mathfrak{A}^*$. Consider this w. Clearly, $\ell(u) + \ell(w) = \ell\left(\underbrace{uw}_{=v}\right) = \ell(v) < \ell(u) + \ell(v)$ (since $\ell(v)$ is positive). Since $v = uw$, the equality $uv = vu$ becomes $uuw = uwu$. Cancelling u from this equality, we obtain $uw = wu$. Now, we can apply Proposition 6.1.4 to w instead of v (by the induction assumption, since $\ell(u) + \ell(w) < \ell(u) + \ell(v)$), and obtain that there exist a $t \in \mathfrak{A}^*$ and two nonnegative integers n and m such that $u = t^n$ and $w = t^m$. Consider this t and these n and m. Of course, $u = t^n$ and $v = \underbrace{u}_{=t^n}\underbrace{w}_{=t^m} = t^n t^m = t^{n+m}$. So the induction step is complete, and Proposition 6.1.4 is proven. □

Proposition 6.1.5. *Let $u, v, w \in \mathfrak{A}^*$ be nonempty words satisfying $uv \geq vu$, $vw \geq wv$ and $wu \geq uw$. Then, there exist a $t \in \mathfrak{A}^*$ and three nonnegative integers n, m and p such that $u = t^n$, $v = t^m$ and $w = t^p$.*

Proof. We prove this by strong induction on $\ell(u) + \ell(v) + \ell(w)$. Clearly, $\ell(u)$, $\ell(v)$ and $\ell(w)$ are positive (since u, v and w are nonempty). We assume WLOG that $\ell(u) = \min\{\ell(u), \ell(v), \ell(w)\}$ (because there is a cyclic symmetry in our situation). Thus, $\ell(u) \leq \ell(v)$ and $\ell(u) \leq \ell(w)$. But $vu \leq uv$. Hence, Proposition 6.1.2(e) (applied to $a = v$, $b = u$, $c = u$ and $d = v$) yields that either we have $v \leq u$ or the word u is a prefix of v. But Proposition 6.1.2(f) (applied to $a = u$, $b = w$, $c = w$ and $d = u$) yields $u \leq w$ (since $uw \leq wu$ and $\ell(u) \leq \ell(w)$). Furthermore, $wv \leq vw$. Hence, Proposition 6.1.2(e) (applied to $a = w$, $b = v$, $c = v$ and $d = w$) yields that either we have $w \leq v$ or the word v is a prefix of w.

From what we have found so far, it is easy to see that u is a prefix of v [277]. In other words, there exists a $v' \in \mathfrak{A}^*$ such that $v = uv'$. Consider this v'.

If the word v' is empty, then the statement of Proposition 6.1.5 can be easily deduced from Proposition 6.1.4[278]. Thus, we assume WLOG that this is not the case. Hence, v' is nonempty.

[276] *Proof.* The word u is a prefix of uv. But the word v is also a prefix of uv (since $uv = vu$). Hence, Proposition 6.1.2(i) (applied to $a = u$, $b = v$ and $c = uv$) yields that either u is a prefix of v, or v is a prefix of u, qed.

[277] *Proof.* Assume the contrary. Then, u is not a prefix of v. Hence, we must have $v \leq u$ (since either we have $v \leq u$ or the word u is a prefix of v), and in fact $v < u$ (because $v = u$ would contradict to u not being a prefix of v). Thus, $v < u \leq w$. But recall that either we have $w \leq v$ or the word v is a prefix of w. Thus, v must be a prefix of w (because $v < w$ rules out $w \leq v$). In other words, there exists a $q \in \mathfrak{A}^*$ such that $w = vq$. Consider this q. We have $v < u \leq w = vq$. Thus, Proposition 6.1.2(g) (applied to $a = v$, $b = u$ and $c = q$) yields that v is a prefix of u. In light of $\ell(u) \leq \ell(v)$, this is only possible if $v = u$, but this contradicts $v < u$. This contradiction completes this proof.

[278] *Proof.* Assume that the word v' is empty. Then, $v = uv'$ becomes $v = u$. Therefore, $vw \geq wv$ becomes $uw \geq wu$. Combined with $wu \geq uw$, this yields $uw = wu$. Hence, Proposition 6.1.4 (applied to w instead of v) yields that there exist a

Using $v = uv'$, we can rewrite $uv \geq vu$ as $uuv' \geq uv'u$. That is, $uv'u \leq uuv'$, so that $v'u \leq uv'$ (by Proposition 6.1.2(c), applied to $a = u$, $c = v'u$ and $d = uv'$). That is, $uv' \geq v'u$. But $\ell(uw) = \ell(u) + \ell(w) = \ell(w) + \ell(u) = \ell(wu) \geq \ell(wu)$. Hence, Proposition 6.1.2(i) (applied to $a = uw$, $b = wu$ and $c = v'$) yields $uwv' \leq wuv'$ (since $uw \leq wu$). Now, $\underbrace{uv'}_{=v} w = vw \geq w \underbrace{v}_{=uv'} = wuv' \geq uwv'$ (since $uwv' \leq wuv'$), so that $uwv' \leq uv'w$. Hence, $wv' \leq v'w$ (by Proposition 6.1.2(c), applied to $a = u$, $c = wv'$ and $d = v'w$), so that $v'w \geq wv'$. Now, we can apply Proposition 6.1.5 to v' instead of v (by the induction hypothesis, because $\underbrace{\ell(u) + \ell(v')}_{\substack{=\ell(uv')=\ell(v) \\ (\text{since } uv'=v)}} + \ell(w) = \ell(v) + \ell(w) < \ell(u) + \ell(v) + \ell(w)$). As a result, we see that there exist a $t \in \mathfrak{A}^*$ and three nonnegative integers n, m and p such that $u = t^n$, $v' = t^m$ and $w = t^p$. Clearly, this t and these n, m, p satisfy $v = \underbrace{u}_{=t^n} \underbrace{v'}_{=t^m} = t^n t^m = t^{n+m}$, and so the statement of Proposition 6.1.5 is satisfied. The induction step is thus complete. □

Corollary 6.1.6. *Let $u, v, w \in \mathfrak{A}^*$ be words satisfying $uv \geq vu$ and $vw \geq wv$. Assume that v is nonempty. Then, $uw \geq wu$.*

Proof. Assume the contrary. Thus, $uw < wu$, so that $wu \geq uw$.

If u or w is empty, then everything is obvious. We thus WLOG assume that u and w are nonempty. Thus, Proposition 6.1.5 shows that there exist a $t \in \mathfrak{A}^*$ and three nonnegative integers n, m and p such that $u = t^n$, $v = t^m$ and $w = t^p$. But this yields $wu = t^p t^n = t^{p+n} = t^{n+p} = \underbrace{t^n}_{=u} \underbrace{t^p}_{=w} = uw$, contradicting $uw < wu$. This contradiction finishes the proof. □

Exercise 6.1.7. Find an alternative proof of Corollary 6.1.6 which does not use Proposition 6.1.5.

The above results have a curious consequence, which we are not going to use:

Corollary 6.1.8. *We can define a preorder on the set $\mathfrak{A}^* \setminus \{\varnothing\}$ of all nonempty words by defining a nonempty word u to be greater-or-equal to a nonempty word v (with respect to this preorder) if and only if $uv \geq vu$. Two nonempty words u, v are equivalent with respect to the equivalence relation induced by this preorder if and only if there exist a $t \in \mathfrak{A}^*$ and two nonnegative integers n and m such that $u = t^n$ and $v = t^m$.*

Proof. The alleged preorder is transitive (by Corollary 6.1.6) and reflexive (obviously), and hence is really a preorder. The claim in the second sentence follows from Proposition 6.1.4. □

As another consequence of Proposition 6.1.5, we obtain a classical property of words [139, Proposition 1.3.1]:

Exercise 6.1.9. Let u and v be words and n and m be positive integers such that $u^n = v^m$. Prove that there exists a word t and positive integers i and j such that $u = t^i$ and $v = t^j$.

Here is another application of Corollary 6.1.6:

Exercise 6.1.10. Let n and m be positive integers. Let $u \in \mathfrak{A}^*$ and $v \in \mathfrak{A}^*$ be two words. Prove that $uv \geq vu$ holds if and only if $u^n v^m \geq v^m u^n$ holds.

Exercise 6.1.11. Let n and m be positive integers. Let $u \in \mathfrak{A}^*$ and $v \in \mathfrak{A}^*$ be two words satisfying $n\ell(u) = m\ell(v)$. Prove that $uv \geq vu$ holds if and only if $u^n \geq v^m$ holds.

We can also generalize Propositions 6.1.4 and 6.1.5:

Exercise 6.1.12. Let $u_1, u_2, \ldots, u_k$ be nonempty words such that every $i \in \{1, 2, \ldots, k\}$ satisfies $u_i u_{i+1} \geq u_{i+1} u_i$, where u_{k+1} means u_1. Show that there exist a word t and nonnegative integers $n_1, n_2, \ldots, n_k$ such that $u_1 = t^{n_1}$, $u_2 = t^{n_2}$, $\ldots$, $u_k = t^{n_k}$.

Now, we define the notion of a Lyndon word. There are several definitions in literature, some of which will be proven equivalent in Theorem 6.1.20.

$t \in \mathfrak{A}^*$ and two nonnegative integers n and m such that $u = t^n$ and $w = t^m$. Clearly, $v = u = t^n$ as well, and so the statement of Proposition 6.1.5 is true.

Definition 6.1.13. A word $w \in \mathfrak{A}^*$ is said to be *Lyndon* if it is nonempty and satisfies the following property: Every nonempty proper suffix v of w satisfies $v > w$.

For example, the word 113 is Lyndon (because its nonempty proper suffixes are 13 and 3, and these are both > 113), and the word 242427 is Lyndon (its nonempty proper suffixes are 42427, 2427, 427, 27 and 7, and again these are each > 242427). The words 2424 and 35346 are not Lyndon (the word 2424 has a nonempty proper suffix $24 \leq 2424$, and the word 35346 has a nonempty proper suffix $346 \leq 35346$). Every word of length 1 is Lyndon (since it has no nonempty proper suffixes). A word $w = (w_1, w_2)$ with two letters is Lyndon if and only if $w_1 < w_2$. A word $w = (w_1, w_2, w_3)$ of length 3 is Lyndon if and only if $w_1 < w_3$ and $w_1 \leq w_2$. A four-letter word $w = (w_1, w_2, w_3, w_4)$ is Lyndon if and only if $w_1 < w_4$, $w_1 \leq w_3$, $w_1 \leq w_2$ and (if $w_1 = w_3$ then $w_2 < w_4$). (These rules only get more complicated as the words grow longer.)

We will show several properties of Lyndon words now. We begin with trivialities which will make some arguments a bit shorter:

Proposition 6.1.14. *Let w be a Lyndon word. Let u and v be words such that $w = uv$.*

(a) *If v is nonempty, then $v \geq w$.*
(b) *If v is nonempty, then $v > u$.*
(c) *If u and v are nonempty, then $vu > uv$.*
(d) *We have $vu \geq uv$.*

Proof. (a) Assume that v is nonempty. Clearly, v is a suffix of w (since $w = uv$). If v is a proper suffix of w, then the definition of a Lyndon word yields that $v > w$ (since w is a Lyndon word); otherwise, v must be w itself. In either case, we have $v \geq w$. Hence, Proposition 6.1.14(a) is proven.

(b) Assume that v is nonempty. From Proposition 6.1.14(a), we obtain $v \geq w = uv > u$ (since v is nonempty). This proves Proposition 6.1.14(b).

(c) Assume that u and v are nonempty. Since u is nonempty, we have $vu > v \geq w$ (by Proposition 6.1.14(a)). Since $w = uv$, this becomes $vu > uv$. This proves Proposition 6.1.14(c).

(d) We need to prove that $vu \geq uv$. If either u or v is empty, vu and uv are obviously equal, and thus $vu \geq uv$ is true in this case. Hence, we can WLOG assume that u and v are nonempty. Assume this. Then, $vu \geq uv$ follows from Proposition 6.1.14(c). This proves Proposition 6.1.14(d). □

Corollary 6.1.15. *Let w be a Lyndon word. Let v be a nonempty suffix of w. Then, $v \geq w$.*

Proof. Since v is a nonempty suffix of w, there exists $u \in \mathfrak{A}^*$ such that $w = uv$. Thus, $v \geq w$ follows from Proposition 6.1.14(a). □

Our next proposition is [93, Lemma 6.5.4]; its part (a) is also [182, (5.1.2)]:

Proposition 6.1.16. *Let u and v be two Lyndon words such that $u < v$. Then:*

(a) *The word uv is Lyndon.*
(b) *We have $uv < v$.*

Proof. (b) The word u is Lyndon and thus nonempty. Hence, $uv \neq v$ [279]. If $uv \leq v\varnothing$, then Proposition 6.1.16(b) easily follows[280]. Hence, for the rest of this proof, we can WLOG assume that we don't have $uv \leq v\varnothing$. Assume this.

We have $u < v$. Hence, Proposition 6.1.2(d) (applied to $a = u$, $b = v$, $c = v$ and $d = \varnothing$) yields that either we have $uv \leq v\varnothing$ or the word u is a prefix of v. Since we don't have $uv \leq v\varnothing$, we thus see that the word u is a prefix of v. In other words, there exists a $t \in \mathfrak{A}^*$ satisfying $v = ut$. Consider this t. Then, t is nonempty (else we would have $v = u \underbrace{t}_{=\varnothing} = u$ in contradiction to $u < v$).

Now, $v = ut$. Hence, t is a proper suffix of v (proper because u is nonempty). Thus, t is a nonempty proper suffix of v. Since every nonempty proper suffix of v is $> v$ (because v is Lyndon), this shows that

[279] *Proof.* Assume the contrary. Then, $uv = v$. Thus, $uv = v = \varnothing v$. Cancelling v from this equation, we obtain $u = \varnothing$. That is, u is empty. This contradicts the fact that u is nonempty. This contradiction proves that our assumption was wrong, qed.

[280] *Proof.* Assume that $uv \leq v\varnothing$. Thus, $uv \leq v\varnothing = v$. Since $uv \neq v$, this becomes $uv < v$, so that Proposition 6.1.16(b) is proven.

$t > v$. Hence, $v \leq t$. Thus, Proposition 6.1.2(b) (applied to $a = u$, $c = v$ and $d = t$) yields $uv \leq ut = v$. Combined with $uv \neq v$, this yields $uv < v$. Hence, Proposition 6.1.16(b) is proven.

(a) The word v is nonempty (since it is Lyndon). Hence, uv is nonempty. It thus remains to check that every nonempty proper suffix p of uv satisfies $p > uv$.

So let p be a nonempty proper suffix of uv. We must show that $p > uv$. Since p is a nonempty proper suffix of uv, we must be in one of the following two cases (depending on whether this suffix begins before the suffix v of uv begins or afterwards):

Case 1: The word p is a nonempty suffix of v. (Note that $p = v$ is allowed.)

Case 2: The word p has the form qv where q is a nonempty proper suffix of u.

Let us first handle Case 1. In this case, p is a nonempty suffix of v. Since v is Lyndon, this yields that $p \geq v$ (by Corollary 6.1.15, applied to v and p instead of w and v). But Proposition 6.1.16(b) yields $uv < v$, thus $v > uv$. Hence, $p \geq v > uv$. We thus have proven $p > uv$ in Case 1.

Let us now consider Case 2. In this case, p has the form qv where q is a nonempty proper suffix of u. Consider this q. Clearly, $q > u$ (since u is Lyndon and since q is a nonempty proper suffix of u), so that $u \leq q$. Thus, Proposition 6.1.2(d) (applied to $a = u$, $b = v$, $c = q$ and $d = v$) yields that either we have $uv \leq qv$ or the word u is a prefix of q. Since u being a prefix of q is impossible (in fact, q is a proper suffix of u, thus shorter than u), we thus must have $uv \leq qv$. Since $uv \neq qv$ (because otherwise we would have $uv = qv$, thus $u = q$ (because we can cancel v from the equality $uv = qv$), contradicting $q > u$), this can be strengthened to $uv < qv = p$. Thus, $p > uv$ is proven in Case 2 as well.

Now that $p > uv$ is shown to hold in both cases, we conclude that $p > uv$ always holds.

Now, let us forget that we fixed p. We have thus shown that every nonempty proper suffix p of uv satisfies $p > uv$. Since uv is nonempty, this yields that uv is Lyndon (by the definition of a Lyndon word). Thus, the proof of Proposition 6.1.16(a) is complete. □

Proposition 6.1.16(b), combined with Corollary 6.1.6, leads to a technical result which we will find good use for later:

Corollary 6.1.17. *Let u and v be two Lyndon words such that $u < v$. Let z be a word such that $zv \geq vz$ and $uz \geq zu$. Then, z is the empty word.*

Proof. Assume the contrary. Then, z is nonempty. Thus, Corollary 6.1.6 (applied to z and v instead of v and w) yields $uv \geq vu$. But Proposition 6.1.16(b) yields $uv < v \leq vu$, contradicting $uv \geq vu$. This contradiction completes our proof. □

We notice that the preorder of Corollary 6.1.8 becomes particularly simple on Lyndon words:

Proposition 6.1.18. *Let u and v be two Lyndon words. Then, $u \geq v$ if and only if $uv \geq vu$.*

Proof. We distinguish between three cases:

Case 1: We have $u < v$.

Case 2: We have $u = v$.

Case 3: We have $u > v$.

Let us consider Case 1. In this case, we have $u < v$. Thus,

$$\begin{aligned} uv &< v \qquad\qquad (\text{by Proposition 6.1.16(b)}) \\ &\leq vu. \end{aligned}$$

Hence, we have neither $u \geq v$ nor $uv \geq vu$ (because we have $u < v$ and $uv < vu$). Thus, Proposition 6.1.18 is proven in Case 1.

In Case 2, we have $u = v$. Therefore, in Case 2, both inequalities $u \geq v$ and $uv \geq vu$ hold (and actually are equalities). Thus, Proposition 6.1.18 is proven in Case 2 as well.

Let us finally consider Case 3. In this case, we have $u > v$. In other words, $v < u$. Thus,

$$\begin{aligned} vu &< u \qquad\qquad (\text{by Proposition 6.1.16(b), applied to } v \text{ and } u \text{ instead of } u \text{ and } v) \\ &\leq uv. \end{aligned}$$

Hence, we have both $u \geq v$ and $uv \geq vu$ (because we have $v < u$ and $vu < uv$). Thus, Proposition 6.1.18 is proven in Case 3.

Proposition 6.1.18 is now proven in all three possible cases. □

Proposition 6.1.19. *Let w be a nonempty word. Let v be the (lexicographically) smallest nonempty suffix of w. Then:*

(a) *The word v is a Lyndon word.*

(b) *Assume that w is not a Lyndon word. Then there exists a nonempty $u \in \mathfrak{A}^*$ such that $w = uv$, $u \geq v$ and $uv \geq vu$.*

Proof. (a) Every nonempty proper suffix of v is $\geq v$ (since every nonempty proper suffix of v is a nonempty suffix of w, but v is the smallest such suffix) and therefore $> v$ (since a proper suffix of v cannot be $= v$). Combined with the fact that v is nonempty, this yields that v is Lyndon. Proposition 6.1.19(a) is proven.

(b) Assume that w is not a Lyndon word. Then, $w \neq v$ (since v is Lyndon (by Proposition 6.1.19(a)) while w is not). Now, v is a suffix of w. Thus, there exists an $u \in \mathfrak{A}^*$ such that $w = uv$. Consider this u. Clearly, u is nonempty (since $uv = w \neq v$). Assume (for the sake of contradiction) that $u < v$. Let v' be the (lexicographically) smallest nonempty suffix of u. Then, v' is a Lyndon word (by Proposition 6.1.19(a), applied to u and v' instead of w and v) and satisfies $v' \leq u$ (since u is a nonempty suffix of u, whereas v' is the smallest such suffix). Thus, v' and v are Lyndon words such that $v' \leq u < v$. Proposition 6.1.16(a) (applied to v' instead of u) now yields that the word $v'v$ is Lyndon. Hence, every nonempty proper suffix of $v'v$ is $> v'v$. Since v is a nonempty proper suffix of $v'v$, this yields that $v > v'v$.

But v' is a nonempty suffix of u, so that $v'v$ is a nonempty suffix of $uv = w$. Since v is the smallest such suffix, this yields that $v'v \geq v$. This contradicts $v > v'v$. Our assumption (that $u < v$) therefore falls. We conclude that $u \geq v$.

It remains to prove that $uv \geq vu$. Assume the contrary. Then, $uv < vu$. Thus, there exists at least one suffix t of u such that $tv < vt$ (namely, $t = u$). Let p be the **minimum-length** such suffix. Then, $pv < vp$. Thus, p is nonempty.

Since p is a suffix of u, it is clear that pv is a suffix of $uv = w$. So we know that pv is a nonempty suffix of w. Since v is the smallest such suffix, this yields that $v \leq pv < vp$. Thus, Proposition 6.1.2(g) (applied to $a = v$, $b = pv$ and $c = p$) yields that v is a prefix of pv. In other words, there exists a $q \in \mathfrak{A}^*$ such that $pv = vq$. Consider this q. This q is nonempty (because otherwise we would have $pv = v \underbrace{q}_{=\varnothing} = v$, contradicting the fact that p is nonempty). From $vq = pv < vp$, we obtain $q \leq p$ (by Proposition 6.1.2(c), applied to $a = v$, $c = q$ and $d = p$).

We know that q is a suffix of pv (since $vq = pv$), whereas pv is a suffix of w. Thus, q is a suffix of w. So q is a nonempty suffix of w. Since v is the smallest such suffix, this yields that $v \leq q$. We now have $v \leq q \leq p \leq pv < vp$. Hence, v is a prefix of p (by Proposition 6.1.2(g), applied to $a = v$, $b = p$ and $c = p$). In other words, there exists an $r \in \mathfrak{A}^*$ such that $p = vr$. Consider this r. Clearly, r is a suffix of p, while p is a suffix of u; therefore, r is a suffix of u. Also, $pv < vp$ rewrites as $vrv < vvr$ (because $p = vr$). Thus, Proposition 6.1.2(c) (applied to $a = v$, $c = rv$ and $d = vr$) yields $rv \leq vr$. Since $rv \neq vr$ (because otherwise, we would have $rv = vr$, thus $v \underbrace{rv}_{=vr} = vvr$, contradicting $vrv < vvr$), this becomes $rv < vr$.

Now, r is a suffix of u such that $rv < vr$. Since p is the minimum-length such suffix, this yields $\ell(r) \geq \ell(p)$. But this contradicts the fact that $\ell\left(\underbrace{p}_{=vr}\right) = \ell(vr) = \underbrace{\ell(v)}_{>0} + \ell(r) > \ell(r)$. This contradiction proves our assumption wrong; thus, we have shown that $uv \geq vu$. Proposition 6.1.19(b) is proven. □

Theorem 6.1.20. *Let w be a nonempty word. The following four assertions are equivalent:*

- *Assertion $\mathcal{A}$: The word w is Lyndon.*
- *Assertion $\mathcal{B}$: Any nonempty words u and v satisfying $w = uv$ satisfy $v > w$.*
- *Assertion $\mathcal{C}$: Any nonempty words u and v satisfying $w = uv$ satisfy $v > u$.*
- *Assertion $\mathcal{D}$: Any nonempty words u and v satisfying $w = uv$ satisfy $vu > uv$.*

Proof. Proof of the implication $\mathcal{A} \Longrightarrow \mathcal{B}$: If Assertion $\mathcal{A}$ holds, then Assertion $\mathcal{B}$ clearly holds (in fact, whenever u and v are nonempty words satisfying $w = uv$, then v is a nonempty proper suffix of w, and therefore $> w$ by the definition of a Lyndon word).

Proof of the implication $\mathcal{A} \Longrightarrow \mathcal{C}$: This implication follows from Proposition 6.1.14(b).

Proof of the implication $\mathcal{A} \Longrightarrow \mathcal{D}$: This implication follows from Proposition 6.1.14(c).

Proof of the implication $\mathcal{B} \Longrightarrow \mathcal{A}$: Assume that Assertion $\mathcal{B}$ holds. If v is a nonempty proper suffix of w, then there exists an $u \in \mathfrak{A}^*$ satisfying $w = uv$. This u is nonempty because v is a proper suffix, and thus Assertion $\mathcal{B}$ yields $v > w$. Hence, every nonempty proper suffix v of w satisfies $v > w$. By the definition of a Lyndon word, this yields that w is Lyndon, so that Assertion $\mathcal{A}$ holds.

Proof of the implication $\mathcal{C} \Longrightarrow \mathcal{A}$: Assume that Assertion $\mathcal{C}$ holds. If w was not Lyndon, then Proposition 6.1.19(b) would yield nonempty words u and v such that $w = uv$ and $u \geq v$; this would contradict Assertion $\mathcal{C}$. Thus, w is Lyndon, and Assertion $\mathcal{A}$ holds.

Proof of the implication $\mathcal{D} \Longrightarrow \mathcal{A}$: Assume that Assertion $\mathcal{D}$ holds. If w was not Lyndon, then Proposition 6.1.19(b) would yield nonempty words u and v such that $w = uv$ and $uv \geq vu$; this would contradict Assertion $\mathcal{D}$. Thus, w is Lyndon, and Assertion $\mathcal{A}$ holds.

Now we have proven enough implications to conclude the equivalence of all four assertions. □

Theorem 6.1.20 connects our definition of Lyndon words with some of the definitions appearing in literature. For example, Lothaire [139, §5.1], Shirshov [202] and de Bruijn/Klarner [29, §4] define Lyndon words using Assertion $\mathcal{D}$ (note, however, that Shirshov takes $<$ instead of $>$ and calls Lyndon words "regular words"; also, de Bruijn/Klarner call Lyndon words "normal words"). Chen-Fox-Lyndon [38, §1], Reutenauer [182] and Radford [177] use our definition (but Chen-Fox-Lyndon call the Lyndon words "standard sequences", and Radford calls them "primes" and uses $<$ instead of $>$).

Theorem 6.1.20 appears (with different notations) in Zhou-Lu [229, Proposition 1.4]. The equivalence $\mathcal{D} \Longleftrightarrow \mathcal{A}$ of our Theorem 6.1.20 is equivalent to [139, Proposition 5.12] and to [38, $\mathfrak{A}'' = \mathfrak{A}'''$].

The following exercise provides a different (laborious) approach to Theorem 6.1.20:

Exercise 6.1.21. (a) Prove that if $u \in \mathfrak{A}^*$ and $v \in \mathfrak{A}^*$ are two words satisfying $uv < vu$, then there exists a nonempty suffix s of u satisfying $sv < v$.

(b) Give a new proof of Theorem 6.1.20 (avoiding the use of Proposition 6.1.19).

[**Hint:** For (a), perform strong induction on $\ell(u) + \ell(v)$, assume the contrary, and distinguish between the case when $u \leq v$ and the case when v is a prefix of u. For (b), use part (a) in proving the implication $\mathcal{D} \Longrightarrow \mathcal{B}$, and factor v as $v = u^m v'$ with m maximal in the proof of the implication $\mathcal{C} \Longrightarrow \mathcal{B}$.]

The following two exercises are taken from [91][281].

Exercise 6.1.22. Let w be a nonempty word. Prove that w is Lyndon if and only if every nonempty word t and every positive integer n satisfy (if $w \leq t^n$, then $w \leq t$).

Exercise 6.1.23. Let $w_1, w_2, \ldots, w_n$ be n Lyndon words, where n is a positive integer. Assume that $w_1 \leq w_2 \leq \cdots \leq w_n$ and $w_1 < w_n$. Show that $w_1 w_2 \cdots w_n$ is a Lyndon word.

The following exercise is a generalization (albeit not in an obvious way) of Exercise 6.1.23:

Exercise 6.1.24. Let $w_1, w_2, \ldots, w_n$ be n Lyndon words, where n is a positive integer. Assume that $w_i w_{i+1} \cdots w_n \geq w_1 w_2 \cdots w_n$ for every $i \in \{1, 2, \ldots, n\}$. Show that $w_1 w_2 \cdots w_n$ is a Lyndon word.

We are now ready to meet one of the most important features of Lyndon words: a bijection between all words and multisets of Lyndon words[282]; it is clear that such a bijection is vital for constructing polynomial generating sets of commutative algebras with bases indexed by words, such as QSym or shuffle algebras. This bijection is given by the *Chen-Fox-Lyndon factorization*:

Definition 6.1.25. Let w be a word. A *Chen-Fox-Lyndon factorization* (in short, *CFL factorization*) of w means a tuple $(a_1, a_2, \ldots, a_k)$ of Lyndon words satisfying $w = a_1 a_2 \cdots a_k$ and $a_1 \geq a_2 \geq \cdots \geq a_k$.

Example 6.1.26. The tuple $(23, 2, 14, 13323, 13, 12, 12, 1)$ is a CFL factorization of the word 23214133231312121 over the alphabet $\{1, 2, 3, \ldots\}$ (ordered by $1 < 2 < 3 < \cdots$), since 23, 2, 14, 13323, 13, 12, 12 and 1 are Lyndon words satisfying $23214133231312121 = 23 \cdot 2 \cdot 14 \cdot 13323 \cdot 13 \cdot 12 \cdot 12 \cdot 1$ and $23 \geq 2 \geq 14 \geq 13323 \geq 13 \geq 12 \geq 12 \geq 1$.

The bijection is given by the following *Chen-Fox-Lyndon theorem* ([93, Theorem 6.5.5], [139, Thm. 5.1.5], [177, part of Thm. 2.1.4]):

[281] Exercise 6.1.22 is more or less [91, Lemma 4.3] with a converse added; Exercise 6.1.23 is [91, Lemma 4.2].

[282] And it is not even the only such bijection: we will see another in Subsection 6.6.1.

Theorem 6.1.27. *Let w be a word. Then, there exists a unique CFL factorization of w.*

Before we prove this, we need to state and prove a lemma (which is [139, Proposition 5.1.6]):

Lemma 6.1.28. *Let $(a_1, a_2, \ldots, a_k)$ be a CFL factorization of a nonempty word w. Let p be a nonempty suffix of w. Then, $p \geq a_k$.*

Proof. We will prove Lemma 6.1.28 by induction over the (obviously) positive integer k.

Induction base: Assume that $k = 1$. Thus, $(a_1, a_2, \ldots, a_k) = (a_1)$ is a tuple of Lyndon words satisfying $w = a_1 a_2 \cdots a_k$. We have $w = a_1 a_2 \cdots a_k = a_1$ (since $k = 1$), so that w is a Lyndon word (since a_1 is a Lyndon word). Thus, Corollary 6.1.15 (applied to $v = p$) yields $p \geq w = a_1 = a_k$ (since $1 = k$). Thus, Lemma 6.1.28 is proven in the case $k = 1$. The induction base is complete.

Induction step: Let K be a positive integer. Assume (as the induction hypothesis) that Lemma 6.1.28 is proven for $k = K$. We now need to show that Lemma 6.1.28 holds for $k = K + 1$.

So let $(a_1, a_2, \ldots, a_{K+1})$ be a CFL factorization of a nonempty word w. Let p be a nonempty suffix of w. We need to prove that $p \geq a_{K+1}$.

By the definition of a CFL factorization, $(a_1, a_2, \ldots, a_{K+1})$ is a tuple of Lyndon words satisfying $w = a_1 a_2 \cdots a_{K+1}$ and $a_1 \geq a_2 \geq \cdots \geq a_{K+1}$. Let $w' = a_2 a_3 \cdots a_{K+1}$; then, $w = a_1 a_2 \cdots a_{K+1} = a_1 \underbrace{(a_2 a_3 \cdots a_{K+1})}_{=w'} = a_1 w'$. Hence, every nonempty suffix of w is either a nonempty suffix of w', or has the form qw' for a nonempty suffix q of a_1. Since p is a nonempty suffix of w, we thus must be in one of the following two cases:

Case 1: The word p is a nonempty suffix of w'.

Case 2: The word p has the form qw' for a nonempty suffix q of a_1.

Let us first consider Case 1. In this case, p is a nonempty suffix of w'. The K-tuple $(a_2, a_3, \ldots, a_{K+1})$ of Lyndon words satisfies $w' = a_2 a_3 \cdots a_{K+1}$ and $a_2 \geq a_3 \geq \cdots \geq a_{K+1}$; therefore, $(a_2, a_3, \ldots, a_{K+1})$ is a CFL factorization of w'. We can thus apply Lemma 6.1.28 to K, w' and $(a_2, a_3, \ldots, a_{K+1})$ instead of k, w and $(a_1, a_2, \ldots, a_k)$ (because we assumed that Lemma 6.1.28 is proven for $k = K$). As a result, we obtain that $p \geq a_{K+1}$. Thus, $p \geq a_{K+1}$ is proven in Case 1.

Let us now consider Case 2. In this case, p has the form qw' for a nonempty suffix q of a_1. Consider this q. Since a_1 is a Lyndon word, we have $q \geq a_1$ (by Corollary 6.1.15, applied to a_1 and q instead of w and v). Thus, $q \geq a_1 \geq a_2 \geq \cdots \geq a_{K+1}$, so that $p = qw' \geq q \geq a_{K+1}$. Thus, $p \geq a_{K+1}$ is proven in Case 2.

We have now proven $p \geq a_{K+1}$ in all cases. This proves that Lemma 6.1.28 holds for $k = K + 1$. The induction step is thus finished, and with it the proof of Lemma 6.1.28. □

Proof of Theorem 6.1.27. Let us first prove that there exists a CFL factorization of w.

Indeed, there clearly exists a tuple $(a_1, a_2, \ldots, a_k)$ of Lyndon words satisfying $w = a_1 a_2 \cdots a_k$ [283]. Fix such a tuple with **minimum** k. We claim that $a_1 \geq a_2 \geq \cdots \geq a_k$.

Indeed, if some $i \in \{1, 2, \ldots, k-1\}$ would satisfy $a_i < a_{i+1}$, then the word $a_i a_{i+1}$ would be Lyndon (by Proposition 6.1.16(a), applied to $u = a_i$ and $v = a_{i+1}$), whence $(a_1, a_2, \ldots, a_{i-1}, a_i a_{i+1}, a_{i+2}, a_{i+3}, \ldots, a_k)$ would also be a tuple of Lyndon words satisfying $w = a_1 a_2 \cdots a_{i-1} (a_i a_{i+1}) a_{i+2} a_{i+3} \cdots a_k$ but having length $k - 1 < k$, contradicting the fact that k is the minimum length of such a tuple. Hence, no $i \in \{1, 2, \ldots, k-1\}$ can satisfy $a_i < a_{i+1}$. In other words, every $i \in \{1, 2, \ldots, k-1\}$ satisfies $a_i \geq a_{i+1}$. In other words, $a_1 \geq a_2 \geq \cdots \geq a_k$. Thus, $(a_1, a_2, \ldots, a_k)$ is a CFL factorization of w, so we have shown that such a CFL factorization exists.

It remains to show that there exists at most one CFL factorization of w. We shall prove this by induction over $\ell(w)$. Thus, we fix a word w and assume that

$$\text{for every word } v \text{ with } \ell(v) < \ell(w), \text{ there exists at most one CFL factorization of } v. \tag{6.1.1}$$

We now have to prove that there exists at most one CFL factorization of w.

Indeed, let $(a_1, a_2, \ldots, a_k)$ and $(b_1, b_2, \ldots, b_m)$ be two CFL factorizations of w. We need to prove that $(a_1, a_2, \ldots, a_k) = (b_1, b_2, \ldots, b_m)$. If w is empty, then this is obvious, so we WLOG assume that it is not; thus, $k > 0$ and $m > 0$.

Since $(b_1, b_2, \ldots, b_m)$ is a CFL factorization of w, we have $w = b_1 b_2 \cdots b_m$, and thus b_m is a nonempty suffix of w. Thus, Lemma 6.1.28 (applied to $p = b_m$) yields $b_m \geq a_k$. The same argument (but with the

[283] For instance, the tuple $(w_1, w_2, \ldots, w_{\ell(w)})$ of one-letter words is a valid example (recall that one-letter words are always Lyndon).

roles of $(a_1, a_2, \ldots, a_k)$ and $(b_1, b_2, \ldots, b_m)$ switched) shows that $a_k \geq b_m$. Combined with $b_m \geq a_k$, this yields $a_k = b_m$. Now let $v = a_1 a_2 \cdots a_{k-1}$. Then, $(a_1, a_2, \ldots, a_{k-1})$ is a CFL factorization of v (since $a_1 \geq a_2 \geq \cdots \geq a_{k-1}$).

Since $(a_1, a_2, \ldots, a_k)$ is a CFL factorization of w, we have $w = a_1 a_2 \cdots a_k = \underbrace{a_1 a_2 \cdots a_{k-1}}_{=v} \underbrace{a_k}_{=b_m} = v b_m$, so that

$$v b_m = w = b_1 b_2 \cdots b_m = b_1 b_2 \cdots b_{m-1} b_m.$$

Cancelling b_m yields $v = b_1 b_2 \cdots b_{m-1}$. Thus, $(b_1, b_2, \ldots, b_{m-1})$ is a CFL factorization of v (since $b_1 \geq b_2 \geq \cdots \geq b_{m-1}$). Since $\ell(v) < \ell(w)$ (because $v = a_1 a_2 \cdots a_{k-1}$ is shorter than $w = a_1 a_2 \cdots a_k$), we can apply (6.1.1) to obtain that there exists at most one CFL factorization of v. But we already know two such CFL factorizations: $(a_1, a_2, \ldots, a_{k-1})$ and $(b_1, b_2, \ldots, b_{m-1})$. Thus, $(a_1, a_2, \ldots, a_{k-1}) = (b_1, b_2, \ldots, b_{m-1})$, which, combined with $a_k = b_m$, leads to $(a_1, a_2, \ldots, a_k) = (b_1, b_2, \ldots, b_m)$. This is exactly what we needed to prove. So we have shown (by induction) that there exists at most one CFL factorization of w. This completes the proof of Theorem 6.1.27. □

The CFL factorization allows us to count all Lyndon words of a given length if $\mathfrak{A}$ is finite:

Exercise 6.1.29. Assume that the alphabet $\mathfrak{A}$ is finite. Let $q = |\mathfrak{A}|$. Let μ be the number-theoretic Möbius function (defined as in Exercise 2.9.6). Show that the number of Lyndon words of length n equals $\dfrac{1}{n} \sum_{d|n} \mu(d) q^{n/d}$ for every positive integer n (where "$\sum_{d|n}$" means a sum over all positive divisors of n). [284]

Exercise 6.1.29 is a well-known result and appears, e.g., in [38, Theorem 1.5] or in [139, Section 5.1].

We will now study another kind of factorization: not of an arbitrary word into Lyndon words, but of a Lyndon word into two smaller Lyndon words. This factorization is called *standard factorization* ([139, §5.1]) or *canonical factorization* ([93, Lemma 6.5.33]); we only introduce it from the viewpoint we are interested in, namely its providing a way to do induction over Lyndon words[285]. Here is what we need to know:

Theorem 6.1.30. *Let w be a Lyndon word of length > 1. Let v be the (lexicographically) smallest nonempty* ***proper*** *suffix of w. Since v is a proper suffix of w, there exists a nonempty $u \in \mathfrak{A}^*$ such that $w = uv$. Consider this u. Then:*

(a) *The words u and v are Lyndon.*
(b) *We have $u < w < v$.*

Proof. Every nonempty proper suffix of v is $\geq v$ (since every nonempty proper suffix of v is a nonempty proper suffix of w, but v is the smallest such suffix) and therefore $> v$ (since a proper suffix of v cannot be $= v$). Combined with the fact that v is nonempty, this yields that v is Lyndon.

Since w is Lyndon, we know that every nonempty proper suffix of w is $> w$. Applied to the nonempty proper suffix v of w, this yields that $v > w$. Hence, $w < v$. Since v is nonempty, we have $u < uv = w < v$. This proves Theorem 6.1.30(b).

Let p be a nonempty proper suffix of u. Then, pv is a nonempty proper suffix of $uv = w$. Thus, $pv > w$ (since every nonempty proper suffix of w is $> w$). Thus, $pv > w = uv$, so that $uv < pv$. Thus, Proposition 6.1.2(e) (applied to $a = u$, $b = v$, $c = p$ and $d = v$) yields that either we have $u \leq p$ or the word p is a prefix of u.

Let us assume (for the sake of contradiction) that $p \leq u$. Then, $p < u$ (because p is a proper suffix of u, and therefore $p \neq u$). Hence, we cannot have $u \leq p$. Thus, the word p is a prefix of u (since either we have $u \leq p$ or the word p is a prefix of u). In other words, there exists a $q \in \mathfrak{A}^*$ such that $u = pq$. Consider this q. We have $w = \underbrace{u}_{=pq} v = pqv = p(qv)$, and thus qv is a proper suffix of w (proper because p is nonempty). Moreover, qv is nonempty (since v is nonempty). Hence, qv is a nonempty proper suffix of w. Since v is the smallest such suffix, this entails that $v \leq qv$. Proposition 6.1.2(b) (applied to $a = p$, $c = v$ and $d = qv$) thus yields $pv \leq pqv$. Hence, $pv \leq pqv = w$, which contradicts $pv > w$. This contradiction shows that our assumption (that $p \leq u$) was false. We thus have $p > u$.

[284] In particular, $\dfrac{1}{n} \sum_{d|n} \mu(d) q^{n/d}$ is an integer.

[285] e.g., allowing to solve Exercise 6.1.24 in a simpler way

We now have shown that $p > u$ whenever p is a nonempty proper suffix of u. Combined with the fact that u is nonempty, this shows that u is a Lyndon word. This completes the proof of Theorem 6.1.30(a). □

Another approach to the standard factorization is given in the following exercise:

Exercise 6.1.31. Let w be a Lyndon word of length > 1. Let v be the longest proper suffix of w such that v is Lyndon[286]. Since v is a proper suffix of w, there exists a nonempty $u \in \mathfrak{A}^*$ such that $w = uv$. Consider this u. Prove that:

(a) The words u and v are Lyndon.
(b) We have $u < w < v$.
(c) The words u and v are precisely the words u and v constructed in Theorem 6.1.30.

Notice that a well-known recursive characterization of Lyndon words [38, $\mathfrak{A}' = \mathfrak{A}''$] can be easily derived from Theorem 6.1.30 and Proposition 6.1.16(a). We will not dwell on it.

The following exercise surveys some variations on the characterizations of Lyndon words[287]:

Exercise 6.1.32. Let w be a nonempty word. Consider the following nine assertions:

- *Assertion $\mathcal{A}'$:* The word w is a power of a Lyndon word.
- *Assertion $\mathcal{B}'$:* If u and v are nonempty words satisfying $w = uv$, then either we have $v \geq w$ or the word v is a prefix of w.
- *Assertion $\mathcal{C}'$:* If u and v are nonempty words satisfying $w = uv$, then either we have $v \geq u$ or the word v is a prefix of u.
- *Assertion $\mathcal{D}'$:* If u and v are nonempty words satisfying $w = uv$, then we have $vu \geq uv$.
- *Assertion $\mathcal{E}'$:* If u and v are nonempty words satisfying $w = uv$, then either we have $v \geq u$ or the word v is a prefix of w.
- *Assertion $\mathcal{F}'$:* The word w is a prefix of a Lyndon word in $\mathfrak{A}^*$.
- *Assertion $\mathcal{F}''$:* Let m be an object not in the alphabet $\mathfrak{A}$. Let us equip the set $\mathfrak{A} \cup \{m\}$ with a total order which extends the total order on the alphabet $\mathfrak{A}$ and which satisfies ($a < m$ for every $a \in \mathfrak{A}$). Then, the word $wm \in (\mathfrak{A} \cup \{m\})^*$ (the concatenation of the word w with the one-letter word m) is a Lyndon word.
- *Assertion $\mathcal{G}'$:* There exists a Lyndon word $t \in \mathfrak{A}^*$, a positive integer ℓ and a prefix p of t (possibly empty) such that $w = t^\ell p$.
- *Assertion $\mathcal{H}'$:* There exists a Lyndon word $t \in \mathfrak{A}^*$, a nonnegative integer ℓ and a prefix p of t (possibly empty) such that $w = t^\ell p$.

(a) Prove the equivalence $\mathcal{A}' \Longleftrightarrow \mathcal{D}'$.
(b) Prove the equivalence $\mathcal{B}' \Longleftrightarrow \mathcal{C}' \Longleftrightarrow \mathcal{E}' \Longleftrightarrow \mathcal{F}'' \Longleftrightarrow \mathcal{G}' \Longleftrightarrow \mathcal{H}'$.
(c) Prove the implication $\mathcal{F}' \Longrightarrow \mathcal{B}'$.
(d) Prove the implication $\mathcal{D}' \Longrightarrow \mathcal{B}'$. (The implication $\mathcal{B}' \Longrightarrow \mathcal{D}'$ is false, as witnessed by the word 11211.)
(e) Prove that if there exists a letter $\mu \in \mathfrak{A}$ such that ($\mu > a$ for every letter a of w), then the equivalence $\mathcal{F}' \Longleftrightarrow \mathcal{F}''$ holds.
(f) Prove that if there exists a letter $\mu \in \mathfrak{A}$ such that ($\mu > a$ for some letter a of w), then the equivalence $\mathcal{F}' \Longleftrightarrow \mathcal{F}''$ holds.

The next exercise (based on work of Hazewinkel [92]) extends some of the above properties of Lyndon words (and words in general) to a more general setting, in which the alphabet $\mathfrak{A}$ is no longer required to be totally ordered, but only needs to be a poset:

Exercise 6.1.33. In this exercise, we shall loosen the requirement that the alphabet $\mathfrak{A}$ be a totally ordered set: Instead, we will only require $\mathfrak{A}$ to be a poset. The resulting more general setting will be called the *partial-order setting*, to distinguish it from the *total-order setting* in which $\mathfrak{A}$ is required to be a totally ordered set. All results in Chapter 6 so far address the total-order setting. In this exercise, we will generalize some of them to the partial-order setting.

[286] This is well-defined, because there exists at least one proper suffix v of w such that v is Lyndon. (Indeed, the last letter of w forms such a suffix, because it is a proper suffix of w (since w has length > 1) and is Lyndon (since it is a one-letter word, and since every one-letter word is Lyndon).)

[287] Compare this with [112, §7.2.11, Theorem Q].

All notions that we have defined in the total-order setting (the notion of a word, the relation $\leq$, the notion of a Lyndon word, etc.) are defined in precisely the same way in the partial-order setting. However, the poset $\mathfrak{A}^*$ is no longer totally ordered in the partial-order setting.

(a) Prove that Proposition 6.1.2 holds in the partial-order setting, as long as one replaces "a total order" by "a partial order" in part (a) of this Proposition.

(b) Prove (in the partial-order setting) that if $a, b, c, d \in \mathfrak{A}^*$ are four words such that the words ab and cd are comparable (with respect to the partial order $\leq$), then the words a and c are comparable.

(c) Prove that Proposition 6.1.4, Proposition 6.1.5, Corollary 6.1.6, Corollary 6.1.8, Exercise 6.1.9, Exercise 6.1.10, Exercise 6.1.11, Exercise 6.1.12, Proposition 6.1.14, Corollary 6.1.15, Proposition 6.1.16, Corollary 6.1.17, Proposition 6.1.18, Theorem 6.1.20, Exercise 6.1.21(a), Exercise 6.1.23, Exercise 6.1.24, Exercise 6.1.31(a) and Exercise 6.1.31(b) still hold in the partial-order setting.

(d) Find a counterexample to Exercise 6.1.22 in the partial-order setting.

(e) Salvage Exercise 6.1.22 in the partial-order setting (i.e., find a statement which is easily equivalent to this exercise in the total-order setting, yet true in the partial-order setting).

(f) In the partial-order setting, a *Hazewinkel-CFL factorization* of a word w will mean a tuple $(a_1, a_2, \ldots, a_k)$ of Lyndon words such that $w = a_1 a_2 \cdots a_k$ and such that no $i \in \{1, 2, \ldots, k-1\}$ satisfies $a_i < a_{i+1}$. Prove that every word w has a unique Hazewinkel-CFL factorization (in the partial-order setting).[288]

(g) Prove that Exercise 6.1.32 still holds in the partial-order setting.

The reader is invited to try extending other results to the partial-order setting (it seems that no research has been done on this except for Hazewinkel's [92]). We shall now, however, return to the total-order setting (which has the most known applications).

Another extension of the notion of Lyndon words has been introduced in 2018 by Dolce, Restivo and Reutenauer [53]; it is based on a generalized version of the lexicographic order, in which different letters are compared differently depending on their positions in the word (i.e., there is one total order for comparing first letters, another for comparing second letters, etc.).

Lyndon words are related to various other objects in mathematics, such as free Lie algebras (Subsection 6.1.1 below), shuffles and shuffle algebras (Sections 6.2 and 6.3 below), QSym (Sections 6.4 and 6.5), Markov chains on combinatorial Hopf algebras ([52]), de Bruijn sequences ([72], [159], [160], [112, §7.2.11, Algorithm F]), symmetric functions (specifically, the transition matrices between the bases $(h_\lambda)_{\lambda \in \mathrm{Par}}$, $(e_\lambda)_{\lambda \in \mathrm{Par}}$ and $(m_\lambda)_{\lambda \in \mathrm{Par}}$; see [117] for this), and the Burrows-Wheeler algorithm for data compression (see Remark 6.6.31 below for a quick idea, and [45], [81], [116] for more). They are also connected to *necklaces* (in the combinatorial sense) – a combinatorial object that also happens to be related to a lot of algebra ([185, Chapter 5], [48]). Let us survey the basics of this latter classical connection in an exercise:

Exercise 6.1.34. Let $\mathfrak{A}$ be any set (not necessarily totally ordered). Let C denote the infinite cyclic group, written multiplicatively. Fix a generator c of C. [289] Fix a positive integer n. The group C acts on $\mathfrak{A}^n$ from the left according to the rule

$$c \cdot (a_1, a_2, \ldots, a_n) = (a_2, a_3, \ldots, a_n, a_1) \qquad \text{for all } (a_1, a_2, \ldots, a_n) \in \mathfrak{A}^n.$$

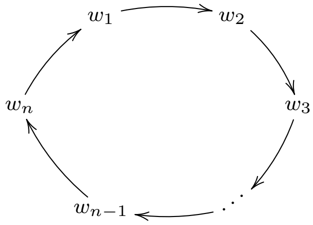
[290] The orbits of this C-action will be called *n-necklaces*[291]; they form a set partition of the set $\mathfrak{A}^n$.

[288] This result, as well as the validity of Proposition 6.1.16 in the partial-order setting, are due to Hazewinkel [92].

[289] So C is a group isomorphic to $(\mathbb{Z}, +)$, and the isomorphism $(\mathbb{Z}, +) \to C$ sends every $n \in \mathbb{Z}$ to c^n. (Recall that we write the binary operation of C as $\cdot$ instead of $+$.)

[290] In other words, c rotates any n-tuple of elements of $\mathfrak{A}$ cyclically to the left. Thus, $c^n \in C$ acts trivially on $\mathfrak{A}^n$, and so this action of C on $\mathfrak{A}^n$ factors through $C / \langle c^n \rangle$ (a cyclic group of order n).

[291] Classically, one visualizes them as necklaces of n beads of $|\mathfrak{A}|$ colors. (The colors are the elements of $\mathfrak{A}$.) For example, the necklace containing an n-tuple $(w_1, w_2, \ldots, w_n)$ is visualized as follows:



The n-necklace containing a given n-tuple $w \in \mathfrak{A}^n$ will be denoted by $[w]$.

(a) Prove that every n-necklace N is a finite nonempty set and satisfies $|N| \mid n$. (Recall that N is an orbit, thus a set; as usual, $|N|$ denotes the cardinality of this set.)

The *period* of an n-necklace N is defined as the positive integer $|N|$. (This $|N|$ is indeed a positive integer, since N is a finite nonempty set.)[292]

An n-necklace is said to be *aperiodic* if its period is n.

(b) Given any n-tuple $w = (w_1, w_2, \ldots, w_n) \in \mathfrak{A}^n$, prove that the n-necklace $[w]$ is aperiodic if and only if every $k \in \{1, 2, \ldots, n-1\}$ satisfies $(w_{k+1}, w_{k+2}, \ldots, w_n, w_1, w_2, \ldots, w_k) \neq w$.

From now on, we assume that the set $\mathfrak{A}$ is totally ordered. We use $\mathfrak{A}$ as our alphabet to define the notions of words, the lexicographic order, and Lyndon words. All notations that we introduced for words will thus be used for elements of $\mathfrak{A}^n$.

(c) Prove that every aperiodic n-necklace contains exactly one Lyndon word.
(d) If N is an n-necklace which is not aperiodic, then prove that N contains no Lyndon word.
(e) Show that the aperiodic n-necklaces are in bijection with Lyndon words of length n.

From now on, we assume that the set $\mathfrak{A}$ is finite. Define the number-theoretic Möbius function μ and the Euler totient function ϕ as in Exercise 2.9.6.

(f) Prove that the number of all aperiodic n-necklaces is

$$\frac{1}{n} \sum_{d \mid n} \mu(d) |\mathfrak{A}|^{n/d}.$$

(g) Prove that the number of all n-necklaces is

$$\frac{1}{n} \sum_{d \mid n} \phi(d) |\mathfrak{A}|^{n/d}.$$

(h) Solve Exercise 6.1.29 again.
(i) Forget that we fixed $\mathfrak{A}$. Show that every $q \in \mathbb{Z}$ satisfies $n \mid \sum_{d \mid n} \mu(d) q^{n/d}$ and $n \mid \sum_{d \mid n} \phi(d) q^{n/d}$.

[**Hint:** For (c), use Theorem 6.1.20. For (i), either use parts (f) and (g) and a trick to extend to q negative; or recall Exercise 2.9.8.]

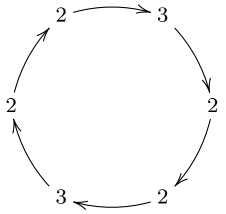
We will pick up the topic of necklaces again in Section 6.6, where we will connect it back to symmetric functions.

6.1.1. *Free Lie algebras.* In this brief subsection, we shall review the connection between Lyndon words and free Lie algebras (following [124, Kap. 4], but avoiding the generality of Hall sets in favor of just using Lyndon words). None of this material shall be used in the rest of these notes. We will only prove some basic results; for more thorough and comprehensive treatments of free Lie algebras, see [182], [27, Chapter 2] and [124, Kap. 4].

We begin with some properties of Lyndon words.

with $w_1, w_2, \ldots, w_n$ being the colors of the respective beads. The intuition behind this is that a necklace is an object that doesn't really change when we rotate it in its plane. However, to make this intuition match the definition, we need to think of a necklace as being stuck in its (fixed) plane, so that we cannot lift it up and turn it around, dropping it back to its plane in a reflected state.

[292] For example, the 6-necklace $[232232]$ – or, visually,



– has period 3, as it is a set of size 3 (with elements 232232, 322322 and 223223). The word "period" hints at the geometric meaning: If an n-necklace N is represented by coloring the vertices of a regular n-gon, then its period is the smallest positive integer d such that the colors are preserved when the n-gon is rotated by $2\pi d/n$.

Exercise 6.1.35. Let $w \in \mathfrak{A}^*$ be a nonempty word. Let v be the longest Lyndon suffix of w [293]. Let t be a Lyndon word. Then, t is the longest Lyndon suffix of wt if and only if we do not have $v < t$.

(We have written "we do not have $v < t$" instead of "$v \geq t$" in Exercise 6.1.35 for reasons of generalizability: This way, Exercise 6.1.35 generalizes to the partial-order setting introduced in Exercise 6.1.33, whereas the version with "$v \geq t$" does not.)

Exercise 6.1.36. Let $w \in \mathfrak{A}^*$ be a word of length > 1. Let v be the longest Lyndon proper suffix of w [294]. Let t be a Lyndon word. Then, t is the longest Lyndon proper suffix of wt if and only if we do not have $v < t$.

(Exercise 6.1.36, while being a trivial consequence of Exercise 6.1.35, is rather useful in the study of free Lie algebras. It generalizes both [38, Lemma (1.6)] (which is obtained by taking $w = c$, $v = b$ and $t = d$) and [139, Proposition 5.1.4] (which is obtained by taking $v = m$ and $t = n$).)

Definition 6.1.37. For the rest of Subsection 6.1.1, we let $\mathfrak{L}$ be the set of all Lyndon words (over the alphabet $\mathfrak{A}$).

Definition 6.1.38. Let w be a Lyndon word of length > 1. Let v be the longest proper suffix of w such that v is Lyndon. (This is well-defined, as we know from Exercise 6.1.31.) Since v is a proper suffix of w, there exists a nonempty $u \in \mathfrak{A}^*$ such that $w = uv$. Consider this u. (Clearly, this u is unique.) Theorem 6.1.30(a) shows that the words u and v are Lyndon. In other words, $u \in \mathfrak{L}$ and $v \in \mathfrak{L}$. Hence, $(u, v) \in \mathfrak{L} \times \mathfrak{L}$. The pair $(u, v) \in \mathfrak{L} \times \mathfrak{L}$ is called the *standard factorization* of w, and is denoted by $\operatorname{stf} w$.

For the sake of easier reference, we gather a few basic properties of the standard factorization:

Exercise 6.1.39. Let w be a Lyndon word of length > 1. Let $(g, h) = \operatorname{stf} w$. Prove the following:

(a) The word h is the longest Lyndon proper suffix of w.
(b) We have $w = gh$.
(c) We have $g < gh < h$.
(d) The word g is Lyndon.
(e) We have $g \in \mathfrak{L}$, $h \in \mathfrak{L}$, $\ell(g) < \ell(w)$ and $\ell(h) < \ell(w)$.
(f) Let t be a Lyndon word. Then, t is the longest Lyndon proper suffix of wt if and only if we do not have $h < t$.

Exercise 6.1.40. Let $\mathfrak{g}$ be a Lie algebra. For every Lyndon word w, let b_w be an element of $\mathfrak{g}$. Assume that for every Lyndon word w of length > 1, we have

$$b_w = [b_u, b_v], \qquad \text{where } (u, v) = \operatorname{stf} w. \tag{6.1.2}$$

Let B be the $\mathbf{k}$-submodule of $\mathfrak{g}$ spanned by the family $(b_w)_{w \in \mathfrak{L}}$.

(a) Prove that B is a Lie subalgebra of $\mathfrak{g}$.
(b) Let $\mathfrak{h}$ be a $\mathbf{k}$-Lie algebra. Let $f : B \to \mathfrak{h}$ be a $\mathbf{k}$-module homomorphism. Assume that whenever w is a Lyndon word of length > 1, we have

$$f([b_u, b_v]) = [f(b_u), f(b_v)], \qquad \text{where } (u, v) = \operatorname{stf} w. \tag{6.1.3}$$

Prove that f is a Lie algebra homomorphism.

[**Hint:** Given two words w and w', write $w \sim w'$ if and only if w' is a permutation of w. Part (a) follows from the fact that for any $(p, q) \in \mathfrak{L} \times \mathfrak{L}$ satisfying $p < q$, we have $[b_p, b_q] \in B_{pq,q}$, where $B_{h,s}$ denotes the $\mathbf{k}$-linear span of $\{b_w \mid w \in \mathfrak{L},\ w \sim h \text{ and } w < s\}$ for any two words h and s. Prove this fact by a double induction, first inducting over $\ell(pq)$, and then (for fixed $\ell(pq)$) inducting over the rank of q in lexicographic order (i.e., assume that the fact is already proven for every $q' < q$ instead of q). In the induction step, assume that $(p, q) \neq \operatorname{stf}(pq)$ (since otherwise the claim is rather obvious) and conclude that p has length > 1; thus, set $(u, v) = \operatorname{stf} p$, so that $\left[\underbrace{b_p}_{=[b_u, b_v]}, b_q\right] = [[b_u, b_v], b_q] = [[b_u, b_q], b_v] - [[b_v, b_q], b_u]$, and use Exercise 6.1.36 to obtain $v < q$.

[293] Of course, a Lyndon suffix of w just means a suffix p of w such that p is Lyndon.

[294] Of course, a Lyndon proper suffix of w just means a proper suffix p of w such that p is Lyndon.

The proof of (b) proceeds by a similar induction, piggybacking on the $[b_p, b_q] \in B_{pq,q}$ claim.]

Exercise 6.1.41. Let V be the free $\mathbf{k}$-module with basis $(x_a)_{a\in\mathfrak{A}}$. For every word $w \in \mathfrak{A}^*$, let x_w be the tensor $x_{w_1} \otimes x_{w_2} \otimes \cdots \otimes x_{w_{\ell(w)}}$. As we know from Example 1.1.2, the tensor algebra $T(V)$ is a free $\mathbf{k}$-module with basis $(x_w)_{w\in\mathfrak{A}^*}$. We regard V as a $\mathbf{k}$-submodule of $T(V)$.

The tensor algebra $T(V)$ becomes a Lie algebra via the commutator (i.e., its Lie bracket is defined by $[\alpha, \beta] = \alpha\beta - \beta\alpha$ for all $\alpha \in T(V)$ and $\beta \in T(V)$).

We define a sequence $(\mathfrak{g}_1, \mathfrak{g}_2, \mathfrak{g}_3, \ldots)$ of $\mathbf{k}$-submodules of $T(V)$ as follows: Recursively, we set $\mathfrak{g}_1 = V$, and for every $i \in \{2, 3, 4, \ldots\}$, we set $\mathfrak{g}_i = [V, \mathfrak{g}_{i-1}]$. Let $\mathfrak{g}$ be the $\mathbf{k}$-submodule $\mathfrak{g}_1 + \mathfrak{g}_2 + \mathfrak{g}_3 + \cdots$ of $T(V)$.

Prove the following:

(a) The $\mathbf{k}$-submodule $\mathfrak{g}$ is a Lie subalgebra of $T(V)$.
(b) If $\mathfrak{k}$ is any Lie subalgebra of $T(V)$ satisfying $V \subset \mathfrak{k}$, then $\mathfrak{g} \subset \mathfrak{k}$.

Now, for every $w \in \mathfrak{L}$, we define an element b_w of $T(V)$ as follows: We define b_w by recursion on the length of w. If the length of w is 1 [295], then we have $w = (a)$ for some letter $a \in \mathfrak{A}$, and we set $b_w = x_a$ for this letter a. If the length of w is > 1, then we set $b_w = [b_u, b_v]$, where $(u, v) = \operatorname{stf} w$ [296].

Prove the following:

(c) For every $w \in \mathfrak{L}$, we have
$$b_w \in x_w + \sum_{\substack{v\in\mathfrak{A}^{\ell(w)};\\ v>w}} \mathbf{k}x_v.$$
(d) The family $(b_w)_{w\in\mathfrak{L}}$ is a basis of the $\mathbf{k}$-module $\mathfrak{g}$.
(e) Let $\mathfrak{h}$ be any $\mathbf{k}$-Lie algebra. Let $\xi : \mathfrak{A} \to \mathfrak{h}$ be any map. Then, there exists a unique Lie algebra homomorphism $\Xi : \mathfrak{g} \to \mathfrak{h}$ such that every $a \in \mathfrak{A}$ satisfies $\Xi(x_a) = \xi(a)$.

Remark 6.1.42. Let V and $\mathfrak{g}$ be as in Exercise 6.1.41. In the language of universal algebra, the statement of Exercise 6.1.41(e) says that $\mathfrak{g}$ (or, to be more precise, the pair $(\mathfrak{g}, f)$, where $f : \mathfrak{A} \to \mathfrak{g}$ is the map sending each $a \in \mathfrak{A}$ to $x_a \in \mathfrak{g}$) satisfies the universal property of the free Lie algebra on the set $\mathfrak{A}$. Thus, this exercise allows us to call $\mathfrak{g}$ the *free Lie algebra* on $\mathfrak{A}$. Most authors define the free Lie algebra differently, but all reasonable definitions of a free Lie algebra[297] lead to isomorphic Lie algebras (because the universal property determines the free Lie algebra uniquely up to canonical isomorphism).

Notice that the Lie algebra $\mathfrak{g}$ does not depend on the total order on the alphabet $\mathfrak{A}$, but the basis $(b_w)_{w\in\mathfrak{L}}$ constructed in Exercise 6.1.41(d) does. There is no known basis of $\mathfrak{g}$ defined without ordering $\mathfrak{A}$.

It is worth noticing that our construction of $\mathfrak{g}$ proves not only that the free Lie algebra on $\mathfrak{A}$ exists, but also that this free Lie algebra can be realized as a Lie subalgebra of the (associative) algebra $T(V)$. Therefore, if we want to prove that a certain identity holds in every Lie algebra, we only need to check that this identity holds in every associative algebra (if all Lie brackets are replaced by commutators); the universal property of the free Lie algebra (i.e., Exercise 6.1.41(e)) will then ensure that this identity also holds in every Lie algebra $\mathfrak{h}$.

There is much more to say about free Lie algebras than what we have said here; in particular, there are connections to symmetric functions, necklaces, representations of symmetric groups and NSym. See [139, §5.3], [182], [27, Chapter 2], [124, §4] and [24] for further developments[298].

6.2. Shuffles and Lyndon words.

We will now connect the theory of Lyndon words with the notion of shuffle products. We have already introduced the latter notion in Definition 1.6.2, but we will now study it

[295] The length of any $w \in \mathfrak{L}$ must be at least 1. (Indeed, if $w \in \mathfrak{L}$, then the word w is Lyndon and thus nonempty, and hence its length must be at least 1.)

[296] This is well-defined, because b_u and b_v have already been defined. [*Proof.* Let $(u, v) = \operatorname{stf} w$. Then, Exercise 6.1.39(e) (applied to $(g, h) = (u, v)$) shows that $u \in \mathfrak{L}$, $v \in \mathfrak{L}$, $\ell(u) < \ell(w)$ and $\ell(v) < \ell(w)$. Recall that we are defining b_w by recursion on the length of w. Hence, b_p is already defined for every $p \in \mathfrak{L}$ satisfying $\ell(p) < \ell(w)$. Applying this to $p = u$, we see that b_u is already defined (since $u \in \mathfrak{L}$ and $\ell(u) < \ell(w)$). The same argument (but applied to v instead of u) shows that b_v is already defined. Hence, b_u and b_v have already been defined. Thus, b_w is well-defined by $b_w = [b_u, b_v]$, qed.]

[297] Here, we call a definition "reasonable" if the "free Lie algebra" it defines satisfies the universal property.

[298] The claim made in [24, page 2] that "$\{x_1, \ldots, x_n\}$ generates freely a Lie subalgebra of A_R" is essentially our Exercise 6.1.41(e).

more closely and introduce some more convenient notations (e.g., we will need a notation for single shuffles, not just the whole multiset).[299]

Definition 6.2.1. (a) Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$. Then, $\operatorname{Sh}_{n,m}$ denotes the subset

$$\left\{\sigma \in \mathfrak{S}_{n+m} \ : \ \sigma^{-1}(1) < \sigma^{-1}(2) < \cdots < \sigma^{-1}(n) ; \ \sigma^{-1}(n+1) < \sigma^{-1}(n+2) < \cdots < \sigma^{-1}(n+m)\right\}$$

of the symmetric group $\mathfrak{S}_{n+m}$.

(b) Let $u = (u_1, u_2, \ldots, u_n)$ and $v = (v_1, v_2, \ldots, v_m)$ be two words. If $\sigma \in \operatorname{Sh}_{n,m}$, then, $u \underset{\sigma}{\sqcup\!\sqcup} v$ will denote the word $\left(w_{\sigma(1)}, w_{\sigma(2)}, \ldots, w_{\sigma(n+m)}\right)$, where $(w_1, w_2, \ldots, w_{n+m})$ is the concatenation $u \cdot v = (u_1, u_2, \ldots, u_n, v_1, v_2, \ldots, v_m)$. We notice that the multiset of all letters of $u \underset{\sigma}{\sqcup\!\sqcup} v$ is the disjoint union of the multiset of all letters of u with the multiset of all letters of v. As a consequence, $\ell\left(u \underset{\sigma}{\sqcup\!\sqcup} v\right) = \ell(u) + \ell(v)$.

(c) Let $u = (u_1, u_2, \ldots, u_n)$ and $v = (v_1, v_2, \ldots, v_m)$ be two words. The *multiset of shuffles of u and v* is defined as the multiset $\left\{\left(w_{\sigma(1)}, w_{\sigma(2)}, \ldots, w_{\sigma(n+m)}\right) \ : \ \sigma \in \operatorname{Sh}_{n,m}\right\}_{\text{multiset}}$, where $(w_1, w_2, \ldots, w_{n+m})$ is the concatenation $u \cdot v = (u_1, u_2, \ldots, u_n, v_1, v_2, \ldots, v_m)$. In other words, the multiset of shuffles of u and v is the multiset

$$\left\{u \underset{\sigma}{\sqcup\!\sqcup} v \ : \ \sigma \in \operatorname{Sh}_{n,m}\right\}_{\text{multiset}}.$$

It is denoted by $u \sqcup\!\sqcup v$.

The next fact provides the main connection between Lyndon words and shuffles:

Theorem 6.2.2. *Let u and v be two words.*

Let $(a_1, a_2, \ldots, a_p)$ be the CFL factorization of u. Let $(b_1, b_2, \ldots, b_q)$ be the CFL factorization of v.

(a) *Let $(c_1, c_2, \ldots, c_{p+q})$ be the result of sorting the list $(a_1, a_2, \ldots, a_p, b_1, b_2, \ldots, b_q)$ in decreasing order*[300]*. Then, the lexicographically highest element of the multiset $u \sqcup\!\sqcup v$ is $c_1 c_2 \cdots c_{p+q}$ (and $(c_1, c_2, \ldots, c_{p+q})$ is the CFL factorization of this element).*

(b) *Let $\mathfrak{L}$ denote the set of all Lyndon words. If w is a Lyndon word and z is any word, let $\operatorname{mult}_w z$ denote the number of terms in the CFL factorization of z which are equal to w. The multiplicity with which the lexicographically highest element of the multiset $u \sqcup\!\sqcup v$ appears in the multiset $u \sqcup\!\sqcup v$ is $\prod_{w \in \mathfrak{L}} \dbinom{\operatorname{mult}_w u + \operatorname{mult}_w v}{\operatorname{mult}_w u}$. (This product is well-defined because almost all of its factors are 1.)*

(c) *If $a_i \geq b_j$ for every $i \in \{1, 2, \ldots, p\}$ and $j \in \{1, 2, \ldots, q\}$, then the lexicographically highest element of the multiset $u \sqcup\!\sqcup v$ is uv.*

(d) *If $a_i > b_j$ for every $i \in \{1, 2, \ldots, p\}$ and $j \in \{1, 2, \ldots, q\}$, then the multiplicity with which the word uv appears in the multiset $u \sqcup\!\sqcup v$ is 1.*

(e) *Assume that u is a Lyndon word. Also, assume that $u \geq b_j$ for every $j \in \{1, 2, \ldots, q\}$. Then, the lexicographically highest element of the multiset $u \sqcup\!\sqcup v$ is uv, and the multiplicity with which this word uv appears in the multiset $u \sqcup\!\sqcup v$ is $\operatorname{mult}_u v + 1$.*

Example 6.2.3. For this example, let u and v be the words $u = 23232$ and $v = 323221$ over the alphabet $\mathfrak{A} = \{1, 2, 3, \ldots\}$ with total order given by $1 < 2 < 3 < \cdots$. The CFL factorizations of u and v are $(23, 23, 2)$ and $(3, 23, 2, 2, 1)$, respectively. Thus, using the notations of Theorem 6.2.2, we have $p = 3$, $(a_1, a_2, \ldots, a_p) = (23, 23, 2)$, $q = 5$ and $(b_1, b_2, \ldots, b_q) = (3, 23, 2, 2, 1)$. Thus, Theorem 6.2.2(a) predicts that the lexicographically highest element of the multiset $u \sqcup\!\sqcup v$ is $c_1 c_2 c_3 c_4 c_5 c_6 c_7 c_8$, where $c_1, c_2, c_3, c_4, c_5, c_6, c_7, c_8$ are the words $23, 23, 2, 3, 23, 2, 2, 1$ listed in decreasing order (in other words, $(c_1, c_2, c_3, c_4, c_5, c_6, c_7, c_8) = (3, 23, 23, 23, 2, 2, 2, 1)$). In other words, Theorem 6.2.2(a) predicts that the lexicographically highest element of the multiset $u \sqcup\!\sqcup v$ is 32323232221. We could verify this by brute force, but this would be laborious since the multiset $u \sqcup\!\sqcup v$ has $\dbinom{5+6}{5} = 462$ elements (with multiplicities). Theorem 6.2.2(b) predicts that this lexicographically highest element 32323232221 appears in the multiset $u \sqcup\!\sqcup v$ with a multiplicity of

[299] Parts (a) and (c) of the below Definition 6.2.1 define notions which have already been introduced in Definition 1.6.2. Of course, the definitions of these notions are equivalent; however, the variables are differently labelled in the two definitions (for example, the variables u, v, w and σ of Definition 6.2.1(c) correspond to the variables a, b, c and w of Definition 1.6.2). The labels in Definition 6.2.1 have been chosen to match with the rest of Section 6.2.

[300] with respect to the total order on $\mathfrak{A}^*$ whose greater-or-equal relation is $\geq$

$\prod_{w\in\mathfrak{L}}\dbinom{\operatorname{mult}_w u+\operatorname{mult}_w v}{\operatorname{mult}_w u}$. This product $\prod_{w\in\mathfrak{L}}\dbinom{\operatorname{mult}_w u+\operatorname{mult}_w v}{\operatorname{mult}_w u}$ is infinite, but all but finitely many of its factors are 1 and therefore can be omitted; the only factors which are not 1 are those corresponding to Lyndon words w which appear both in the CFL factorization of u and in the CFL factorization of v (since for any other factor, at least one of the numbers $\operatorname{mult}_w u$ or $\operatorname{mult}_w v$ equals 0, and therefore the binomial coefficient $\dbinom{\operatorname{mult}_w u+\operatorname{mult}_w v}{\operatorname{mult}_w u}$ equals 1). Thus, in order to compute the product $\prod_{w\in\mathfrak{L}}\dbinom{\operatorname{mult}_w u+\operatorname{mult}_w v}{\operatorname{mult}_w u}$, we only need to multiply these factors. In our example, these are the factors for $w=23$ and for $w=2$ (these are the only Lyndon words which appear both in the CFL factorization $(23,23,2)$ of u and in the CFL factorization $(3,23,2,2,1)$ of v). So we have

$$\prod_{w\in\mathfrak{L}}\binom{\operatorname{mult}_w u+\operatorname{mult}_w v}{\operatorname{mult}_w u}=\underbrace{\binom{\operatorname{mult}_{23} u+\operatorname{mult}_{23} v}{\operatorname{mult}_{23} u}}_{=\binom{2+1}{2}=3}\underbrace{\binom{\operatorname{mult}_{2} u+\operatorname{mult}_{2} v}{\operatorname{mult}_{2} u}}_{=\binom{1+2}{1}=3}=3\cdot 3=9.$$

The word 32323232221 must thus appear in the multiset $u \sqcup\!\sqcup v$ with a multiplicity of 9. This, too, could be checked by brute force.

Theorem 6.2.2 (and Theorem 6.2.22 further below, which describes more precisely how the lexicographically highest element of $u \sqcup\!\sqcup v$ emerges by shuffling u and v) is fairly close to [177, Theorem 2.2.2] (and will be used for the same purposes), the main difference being that we are talking about the shuffle product of two (not necessarily Lyndon) words, while Radford (and most other authors) study the shuffle product of many Lyndon words.

In order to prove Theorem 6.2.2, we will need to make some stronger statements, for which we first have to introduce some more notation:

Definition 6.2.4. (a) If p and q are two integers, then $[p:q]^+$ denotes the interval $\{p+1,p+2,\ldots,q\}$ of $\mathbb{Z}$. Note that $\left|[p:q]^+\right|=q-p$ if $q\geq p$.

(b) If I and J are two nonempty intervals of $\mathbb{Z}$, then we say that $I<J$ if and only if every $i\in I$ and $j\in J$ satisfy $i<j$. This defines a partial order on the set of nonempty intervals of $\mathbb{Z}$. (Roughly speaking, $I<J$ if the interval I ends before J begins.)

(c) If w is a word with n letters (for some $n\in\mathbb{N}$), and I is an interval of $\mathbb{Z}$ such that $I\subset[0:n]^+$, then $w[I]$ will denote the word $(w_{p+1},w_{p+2},\ldots,w_q)$, where I is written in the form $I=[p:q]^+$ with $q\geq p$. Obviously, $\ell(w[I])=|I|=q-p$. A word of the form $w[I]$ for an interval $I\subset[0:n]^+$ (equivalently, a word which is a prefix of a suffix of w) is called a *factor* of w.

(d) Let α be a composition. Then, we define a tuple $\operatorname{intsys}\alpha$ of intervals of $\mathbb{Z}$ as follows: Write α in the form $(\alpha_1,\alpha_2,\ldots,\alpha_\ell)$ (so that $\ell=\ell(\alpha)$). Then, set $\operatorname{intsys}\alpha=(I_1,I_2,\ldots,I_\ell)$, where

$$I_i=\left[\sum_{k=1}^{i-1}\alpha_k:\sum_{k=1}^{i}\alpha_k\right]^+\qquad\text{for every } i\in\{1,2,\ldots,\ell\}.$$

This ℓ-tuple $\operatorname{intsys}\alpha$ is a tuple of nonempty intervals of $\mathbb{Z}$. This tuple $\operatorname{intsys}\alpha$ is called the *interval system corresponding to* α. (This is precisely the ℓ-tuple $(I_1,I_2,\ldots,I_\ell)$ constructed in Definition 4.3.4.) The length of the tuple $\operatorname{intsys}\alpha$ is $\ell(\alpha)$.

Example 6.2.5. (a) We have $[2:4]^+=\{3,4\}$ and $[3:3]^+=\varnothing$.

(b) We have $[2:4]^+<[4:5]^+<[6:8]^+$, but we have neither $[2:4]^+<[3:5]^+$ nor $[3:5]^+<[2:4]^+$.

(c) If w is the word 915352, then $w\left[[0:3]^+\right]=(w_1,w_2,w_3)=915$ and $w\left[[2:4]^+\right]=(w_3,w_4)=53$.

(d) If α is the composition $(4,1,4,2,3)$, then the interval system corresponding to α is

$$\begin{aligned}\operatorname{intsys}\alpha&=\left([0:4]^+,[4:5]^+,[5:9]^+,[9:11]^+,[11:14]^+\right)\\&=(\{1,2,3,4\},\{5\},\{6,7,8,9\},\{10,11\},\{12,13,14\}).\end{aligned}$$

The following properties of the notions introduced in the preceding definition are easy to check:

Remark 6.2.6. (a) If I and J are two nonempty intervals of $\mathbb{Z}$ satisfying $I<J$, then I and J are disjoint.

(b) If I and J are two disjoint nonempty intervals of $\mathbb{Z}$, then either $I < J$ or $J < I$.

(c) Let α be a composition. Write α in the form $(\alpha_1, \alpha_2, \ldots, \alpha_\ell)$ (so that $\ell = \ell(\alpha)$). The interval system $\operatorname{intsys} \alpha$ can be described as the unique ℓ-tuple $(I_1, I_2, \ldots, I_\ell)$ of nonempty intervals of $\mathbb{Z}$ satisfying the following three properties:
  - The intervals I_1, I_2, ..., I_ℓ form a set partition of the set $[0 : n]^+$, where $n = |\alpha|$.
  - We have $I_1 < I_2 < \cdots < I_\ell$.
  - We have $|I_i| = \alpha_i$ for every $i \in \{1, 2, \ldots, \ell\}$.

Exercise 6.2.7. Prove Remark 6.2.6.

The following two lemmas are collections of more or less trivial consequences of what it means to be an element of $\mathrm{Sh}_{n,m}$ and what it means to be a shuffle:

Lemma 6.2.8. *Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$. Let $\sigma \in \mathrm{Sh}_{n,m}$.*

(a) *If I is an interval of $\mathbb{Z}$ such that $I \subset [0 : n + m]^+$, then $\sigma(I) \cap [0 : n]^+$ and $\sigma(I) \cap [n : n + m]^+$ are intervals.*

(b) *Let K and L be nonempty intervals of $\mathbb{Z}$ such that $K \subset [0 : n]^+$ and $L \subset [0 : n]^+$ and $K < L$ and such that $K \cup L$ is an interval. Assume that $\sigma^{-1}(K)$ and $\sigma^{-1}(L)$ are intervals, but $\sigma^{-1}(K) \cup \sigma^{-1}(L)$ is not an interval. Then, there exists a nonempty interval $P \subset [n : n + m]^+$ such that $\sigma^{-1}(P)$, $\sigma^{-1}(K) \cup \sigma^{-1}(P)$ and $\sigma^{-1}(P) \cup \sigma^{-1}(L)$ are intervals and such that $\sigma^{-1}(K) < \sigma^{-1}(P) < \sigma^{-1}(L)$.*

(c) *Lemma 6.2.8(b) remains valid if "$K \subset [0 : n]^+$ and $L \subset [0 : n]^+$" and "$P \subset [n : n + m]^+$" are replaced by "$K \subset [n : n + m]^+$ and $L \subset [n : n + m]^+$" and "$P \subset [0 : n]^+$", respectively.*

Exercise 6.2.9. Prove Lemma 6.2.8.

Lemma 6.2.10. *Let u and v be two words. Let $n = \ell(u)$ and $m = \ell(v)$. Let $\sigma \in \mathrm{Sh}_{n,m}$.*

(a) *If I is an interval of $\mathbb{Z}$ satisfying either $I \subset [0 : n]^+$ or $I \subset [n : n + m]^+$, and if $\sigma^{-1}(I)$ is an interval, then*

$$\left(u \underset{\sigma}{\sqcup\!\sqcup} v\right)\left[\sigma^{-1}(I)\right] = (uv)[I]. \tag{6.2.1}$$

(b) *Assume that $u \underset{\sigma}{\sqcup\!\sqcup} v$ is the lexicographically highest element of the multiset $u \sqcup\!\sqcup v$. Let $I \subset [0 : n]^+$ and $J \subset [n : n + m]^+$ be two nonempty intervals. Assume that $\sigma^{-1}(I)$ and $\sigma^{-1}(J)$ are also intervals, that $\sigma^{-1}(I) < \sigma^{-1}(J)$, and that $\sigma^{-1}(I) \cup \sigma^{-1}(J)$ is an interval as well. Then, $(uv)[I] \cdot (uv)[J] \geq (uv)[J] \cdot (uv)[I]$.*

(c) *Lemma 6.2.10(b) remains valid if "$I \subset [0 : n]^+$ and $J \subset [n : n + m]^+$" is replaced by "$I \subset [n : n + m]^+$ and $J \subset [0 : n]^+$".*

Exercise 6.2.11. Prove Lemma 6.2.10.

[**Hint:** For (b), show that there exists a $\tau \in \mathrm{Sh}_{n,m}$ such that $u \underset{\tau}{\sqcup\!\sqcup} v$ differs from $u \underset{\sigma}{\sqcup\!\sqcup} v$ only in the order of the subwords $(uv)[I]$ and $(uv)[J]$.]

We are still a few steps away from stating our results in a way that allows comfortably proving Theorem 6.2.2. For the latter aim, we introduce the notion of *α-clumping permutations*, and characterize them in two ways:

Definition 6.2.12. Let $n \in \mathbb{N}$. Let α be a composition of n. Let $\ell = \ell(\alpha)$.

(a) For every set S of positive integers, let $\overrightarrow{S}$ denote the list of all elements of S in increasing order (with each element appearing exactly once). Notice that this list $\overrightarrow{S}$ is a word over the set of positive integers.

(b) For every $\tau \in \mathfrak{S}_\ell$, we define a permutation $\operatorname{iper}(\alpha, \tau) \in \mathfrak{S}_n$ as follows:

The interval system corresponding to α is an ℓ-tuple of intervals (since $\ell(\alpha) = \ell$); denote this ℓ-tuple by $(I_1, I_2, \ldots, I_\ell)$. Now, define $\operatorname{iper}(\alpha, \tau)$ to be the permutation in $\mathfrak{S}_n$ which (in one-line notation) is the word $\overrightarrow{I_{\tau(1)}}\,\overrightarrow{I_{\tau(2)}} \cdots \overrightarrow{I_{\tau(\ell)}}$ (a concatenation of ℓ words). This is well-defined[301]; hence, $\operatorname{iper}(\alpha, \tau) \in \mathfrak{S}_n$ is defined.

[301] In fact, from the properties of interval systems, we know that the intervals I_1, I_2, ..., I_ℓ form a set partition of the set $[0 : n]^+$. Hence, the intervals $I_{\tau(1)}$, $I_{\tau(2)}$, ..., $I_{\tau(\ell)}$ form a set partition of the set $[0 : n]^+$. As a consequence, the word

(c) The interval system corresponding to α is an ℓ-tuple of intervals (since $\ell(\alpha) = \ell$); denote this ℓ-tuple by $(I_1, I_2, \ldots, I_\ell)$.

A permutation $\sigma \in \mathfrak{S}_n$ is said to be *α-clumping* if every $i \in \{1, 2, \ldots, \ell\}$ has the two properties that:
- the set $\sigma^{-1}(I_i)$ is an interval;
- the restriction of the map σ^{-1} to the interval I_i is increasing.

Example 6.2.13. For this example, let $n = 7$ and $\alpha = (2, 1, 3, 1)$. Then, $\ell = \ell(\alpha) = 4$ and $(I_1, I_2, I_3, I_4) = (\{1, 2\}, \{3\}, \{4, 5, 6\}, \{7\})$ (where we are using the notations of Definition 6.2.12). Hence, $\overrightarrow{I_1} = 12$, $\overrightarrow{I_2} = 3$, $\overrightarrow{I_3} = 456$ and $\overrightarrow{I_4} = 7$.

(a) If $\tau \in \mathfrak{S}_\ell = \mathfrak{S}_4$ is the permutation $(2, 3, 1, 4)$, then $\operatorname{iper}(\alpha, \tau)$ is the permutation in $\mathfrak{S}_7$ which (in one-line notation) is the word $\overrightarrow{I_{\tau(1)}}\,\overrightarrow{I_{\tau(2)}}\,\overrightarrow{I_{\tau(3)}}\,\overrightarrow{I_{\tau(4)}} = \overrightarrow{I_2}\,\overrightarrow{I_3}\,\overrightarrow{I_1}\,\overrightarrow{I_4} = 3456127$.

If $\tau \in \mathfrak{S}_\ell = \mathfrak{S}_4$ is the permutation $(3, 1, 4, 2)$, then $\operatorname{iper}(\alpha, \tau)$ is the permutation in $\mathfrak{S}_7$ which (in one-line notation) is the word $\overrightarrow{I_{\tau(1)}}\,\overrightarrow{I_{\tau(2)}}\,\overrightarrow{I_{\tau(3)}}\,\overrightarrow{I_{\tau(4)}} = \overrightarrow{I_3}\,\overrightarrow{I_1}\,\overrightarrow{I_4}\,\overrightarrow{I_2} = 4561273$.

(b) The permutation $\sigma = (3, 7, 4, 5, 6, 1, 2) \in \mathfrak{S}_7$ (given here in one-line notation) is α-clumping, because:
- every $i \in \{1, 2, \ldots, \ell\} = \{1, 2, 3, 4\}$ has the property that $\sigma^{-1}(I_i)$ is an interval (namely, $\sigma^{-1}(I_1) = \sigma^{-1}(\{1, 2\}) = \{6, 7\}$, $\sigma^{-1}(I_2) = \sigma^{-1}(\{3\}) = \{1\}$, $\sigma^{-1}(I_3) = \sigma^{-1}(\{4, 5, 6\}) = \{3, 4, 5\}$ and $\sigma^{-1}(I_4) = \sigma^{-1}(\{7\}) = \{2\}$), and
- the restrictions of the map σ^{-1} to the intervals I_i are increasing (this means that $\sigma^{-1}(1) < \sigma^{-1}(2)$ and $\sigma^{-1}(4) < \sigma^{-1}(5) < \sigma^{-1}(6)$, since the one-element intervals I_2 and I_4 do not contribute anything to this condition).

Here is a more or less trivial observation:

Proposition 6.2.14. *Let $n \in \mathbb{N}$. Let α be a composition of n. Let $\ell = \ell(\alpha)$. Write α in the form $(\alpha_1, \alpha_2, \ldots, \alpha_\ell)$. The interval system corresponding to α is an ℓ-tuple of intervals (since $\ell(\alpha) = \ell$); denote this ℓ-tuple by $(I_1, I_2, \ldots, I_\ell)$. Let $\tau \in \mathfrak{S}_\ell$. Set $\sigma = \operatorname{iper}(\alpha, \tau)$.*

(a) *We have $\sigma^{-1}\left(I_{\tau(j)}\right) = \left[\sum_{k=1}^{j-1} \alpha_{\tau(k)} : \sum_{k=1}^{j} \alpha_{\tau(k)}\right]^+$ for every $j \in \{1, 2, \ldots, \ell\}$.*

(b) *For every $j \in \{1, 2, \ldots, \ell\}$, the restriction of the map σ^{-1} to the interval $I_{\tau(j)}$ is increasing.*

(c) *The permutation $\operatorname{iper}(\alpha, \tau)$ is α-clumping.*

(d) *Let $i \in \{1, 2, \ldots, \ell - 1\}$. Then, the sets $\sigma^{-1}\left(I_{\tau(i)}\right)$, $\sigma^{-1}\left(I_{\tau(i+1)}\right)$ and $\sigma^{-1}\left(I_{\tau(i)}\right) \cup \sigma^{-1}\left(I_{\tau(i+1)}\right)$ are nonempty intervals. Also, $\sigma^{-1}\left(I_{\tau(i)}\right) < \sigma^{-1}\left(I_{\tau(i+1)}\right)$.*

Exercise 6.2.15. Prove Proposition 6.2.14.

Proposition 6.2.16. *Let $n \in \mathbb{N}$. Let α be a composition of n. Let $\ell = \ell(\alpha)$.*

(a) *Define a map*

$$\begin{aligned} \operatorname{iper}_\alpha : \mathfrak{S}_\ell &\longrightarrow \{\omega \in \mathfrak{S}_n \mid \omega \text{ is } \alpha\text{-clumping}\}, \\ \tau &\longmapsto \operatorname{iper}(\alpha, \tau) \end{aligned}$$

[302]. *This map $\operatorname{iper}_\alpha$ is bijective.*

(b) *Let $\sigma \in \mathfrak{S}_n$ be an α-clumping permutation. Then, there exists a unique $\tau \in \mathfrak{S}_\ell$ satisfying $\sigma = \operatorname{iper}(\alpha, \tau)$.*

Exercise 6.2.17. Prove Proposition 6.2.16.

Next, we recall that the concatenation $\alpha \cdot \beta$ of two compositions α and β is defined in the same way as the concatenation of two words; if we regard compositions as words over the alphabet $\{1, 2, 3, \ldots\}$, then the concatenation $\alpha \cdot \beta$ of two compositions α and β **is** the concatenation $\alpha\beta$ of the words α and β. Thus, we are going to write $\alpha\beta$ for the concatenation $\alpha \cdot \beta$ of two compositions α and β from now on.

$\overrightarrow{I_{\tau(1)}}\,\overrightarrow{I_{\tau(2)}} \cdots \overrightarrow{I_{\tau(\ell)}}$ is a permutation of the word $12 \ldots n$, and so there exists a permutation in $\mathfrak{S}_n$ which (in one-line notation) is this word, qed.

[302] This map is well-defined because for every $\tau \in \mathfrak{S}_\ell$, the permutation $\operatorname{iper}(\alpha, \tau)$ is α-clumping (according to Proposition 6.2.14(c)).

Proposition 6.2.18. *Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$. Let α be a composition of n, and β be a composition of m. Let $p = \ell(\alpha)$ and $q = \ell(\beta)$. Let $\tau \in \mathfrak{S}_{p+q}$. Notice that* $\operatorname{iper}(\alpha\beta, \tau) \in \mathfrak{S}_{n+m}$ *(since $\alpha\beta$ is a composition of $n+m$ having length $\ell(\alpha\beta) = \ell(\alpha) + \ell(\beta) = p+q$). Then, $\tau \in \operatorname{Sh}_{p,q}$ if and only if* $\operatorname{iper}(\alpha\beta, \tau) \in \operatorname{Sh}_{n,m}$.

Exercise 6.2.19. Prove Proposition 6.2.18.

Here is one more simple fact:

Lemma 6.2.20. *Let u and v be two words. Let $n = \ell(u)$ and $m = \ell(v)$. Let α be a composition of n, and let β be a composition of m. Let $p = \ell(\alpha)$ and $q = \ell(\beta)$. The concatenation $\alpha\beta$ is a composition of $n+m$ having length $\ell(\alpha\beta) = \ell(\alpha) + \ell(\beta) = p+q$. Thus, the interval system corresponding to $\alpha\beta$ is a $(p+q)$-tuple of intervals which covers $[0 : n+m]^+$. Denote this $(p+q)$-tuple by $(I_1, I_2, \ldots, I_{p+q})$.*

Let $\tau \in \operatorname{Sh}_{p,q}$. Set $\sigma = \operatorname{iper}(\alpha\beta, \tau)$. Then,

$$u \underset{\sigma}{\sqcup\!\sqcup} v = (uv)\left[I_{\tau(1)}\right] \cdot (uv)\left[I_{\tau(2)}\right] \cdot \cdots \cdot (uv)\left[I_{\tau(p+q)}\right].$$

Exercise 6.2.21. Prove Lemma 6.2.20.

Having these notations and trivialities in place, we can say a bit more about the lexicographically highest element of a shuffle product than what was said in Theorem 6.2.2:

Theorem 6.2.22. *Let u and v be two words. Let $n = \ell(u)$ and $m = \ell(v)$.*

Let $(a_1, a_2, \ldots, a_p)$ be the CFL factorization of u. Let $(b_1, b_2, \ldots, b_q)$ be the CFL factorization of v.

Let α be the p-tuple $(\ell(a_1), \ell(a_2), \ldots, \ell(a_p))$. Then, α is a composition[303] *of length p and size $\sum_{k=1}^{p} \ell(a_k) = \ell\left(\underbrace{a_1 a_2 \cdots a_p}_{=u}\right) = \ell(u) = n$.*

Let β be the q-tuple $(\ell(b_1), \ell(b_2), \ldots, \ell(b_q))$. Then, β is a composition of length q and size $\sum_{k=1}^{q} \ell(b_k) = m$.[304]

Now, α is a composition of length p and size n, and β is a composition of length q and size m. Thus, the concatenation $\alpha\beta$ of these two tuples is a composition of length $p+q$ and size $n+m$. The interval system corresponding to this composition $\alpha\beta$ is a $(p+q)$-tuple (since said composition has length $p+q$); denote this $(p+q)$-tuple by $(I_1, I_2, \ldots, I_{p+q})$.

(a) *If $\tau \in \operatorname{Sh}_{p,q}$ satisfies $(uv)\left[I_{\tau(1)}\right] \geq (uv)\left[I_{\tau(2)}\right] \geq \cdots \geq (uv)\left[I_{\tau(p+q)}\right]$, and if we set $\sigma = \operatorname{iper}(\alpha\beta, \tau)$, then $\sigma \in \operatorname{Sh}_{n,m}$, and the word $u \underset{\sigma}{\sqcup\!\sqcup} v$ is the lexicographically highest element of the multiset $u \sqcup\!\sqcup v$.*

(b) *Let $\sigma \in \operatorname{Sh}_{n,m}$ be a permutation such that $u \underset{\sigma}{\sqcup\!\sqcup} v$ is the lexicographically highest element of the multiset $u \sqcup\!\sqcup v$. Then, there exists a unique permutation $\tau \in \operatorname{Sh}_{p,q}$ satisfying $(uv)\left[I_{\tau(1)}\right] \geq (uv)\left[I_{\tau(2)}\right] \geq \cdots \geq (uv)\left[I_{\tau(p+q)}\right]$ and $\sigma = \operatorname{iper}(\alpha\beta, \tau)$.*

Proof. Before we step to the actual proof, we need to make some preparation. First of all, $(I_1, I_2, \ldots, I_{p+q})$ is the interval system corresponding to the composition $\alpha\beta$. In other words,

$$(I_1, I_2, \ldots, I_{p+q}) = \operatorname{intsys}(\alpha\beta). \tag{6.2.2}$$

But since $\alpha = (\ell(a_1), \ell(a_2), \ldots, \ell(a_p))$ and $\beta = (\ell(b_1), \ell(b_2), \ldots, \ell(b_q))$, we have

$$\alpha\beta = (\ell(a_1), \ell(a_2), \ldots, \ell(a_p), \ell(b_1), \ell(b_2), \ldots, \ell(b_q)).$$

Thus, (6.2.2) rewrites as

$$(I_1, I_2, \ldots, I_{p+q}) = \operatorname{intsys}(\ell(a_1), \ell(a_2), \ldots, \ell(a_p), \ell(b_1), \ell(b_2), \ldots, \ell(b_q)).$$

By the definition of $\operatorname{intsys}(\ell(a_1), \ell(a_2), \ldots, \ell(a_p), \ell(b_1), \ell(b_2), \ldots, \ell(b_q))$, we thus have

$$I_i = \left[\sum_{k=1}^{i-1} \ell(a_k) : \sum_{k=1}^{i} \ell(a_k)\right]^+ \qquad \text{for every } i \in \{1, 2, \ldots, p\},$$

[303] since Lyndon words are nonempty, and thus $\ell(a_i) > 0$ for every i

[304] The proof of this is the same as the proof of the fact that α is a composition of length p and size $\sum_{k=1}^{p} \ell(\alpha_k) = n$.

and besides

$$I_{p+j}=\left[n+\sum_{k=1}^{j-1}\ell\left(b_k\right):n+\sum_{k=1}^{j}\ell\left(b_k\right)\right]^{+}\qquad\text{for every } j\in\{1,2,\ldots,q\}$$

(since $\sum_{k=1}^{p}\ell\left(a_k\right)=n$). Moreover, Remark 6.2.6(c) shows that $(I_1,I_2,\ldots,I_{p+q})$ is a $(p+q)$-tuple of nonempty intervals of $\mathbb{Z}$ and satisfies the following three properties:

- The intervals I_1, I_2, $\ldots$, I_{p+q} form a set partition of the set $[0:n+m]^{+}$.
- We have $I_1<I_2<\cdots<I_{p+q}$.
- We have $|I_i|=\ell\left(a_i\right)$ for every $i\in\{1,2,\ldots,p\}$ and $|I_{p+j}|=\ell\left(b_j\right)$ for every $j\in\{1,2,\ldots,q\}$.

Of course, every $i\in\{1,2,\ldots,p\}$ satisfies

$$I_i\subset[0:n]^{+}\qquad\text{and}\qquad(uv)\left[I_i\right]=u\left[I_i\right]=a_i. \tag{6.2.3}$$

Meanwhile, every $i\in\{p+1,p+2,\ldots,p+q\}$ satisfies

$$I_i\subset[n:n+m]^{+}\qquad\text{and}\qquad(uv)\left[I_i\right]=v\left[I_i-n\right]=b_{i-p} \tag{6.2.4}$$

(where I_i-n denotes the interval $\{k-n\ \mid\ k\in I_i\}$). We thus see that

$$(uv)\left[I_i\right]\ \text{is a Lyndon word}\qquad\text{for every } i\in\{1,2,\ldots,p+q\} \tag{6.2.5}$$

[305].

By the definition of a CFL factorization, we have $a_1\geq a_2\geq\cdots\geq a_p$ and $b_1\geq b_2\geq\cdots\geq b_q$.

We have $\sigma\in\operatorname{Sh}_{n,m}$, so that $\sigma^{-1}(1)<\sigma^{-1}(2)<\cdots<\sigma^{-1}(n)$ and $\sigma^{-1}(n+1)<\sigma^{-1}(n+2)<\cdots<\sigma^{-1}(n+m)$. In other words, the restriction of the map σ^{-1} to the interval $[0:n]^{+}$ is strictly increasing, and so is the restriction of the map σ^{-1} to the interval $[n:n+m]^{+}$.

(b) We will first show that

$$\text{if } J\subset[0:n]^{+}\text{ is an interval such that the word }(uv)[J]\text{ is Lyndon, then }\sigma^{-1}(J)\text{ is an interval.} \tag{6.2.6}$$

Proof of (6.2.6): We will prove (6.2.6) by strong induction over $|J|$.

So, fix some $N\in\mathbb{N}$. Assume (as the induction hypothesis) that (6.2.6) has been proven whenever $|J|<N$. We now need to prove (6.2.6) when $|J|=N$.

Let $J\subset[0:n]^{+}$ be an interval such that the word $(uv)[J]$ is Lyndon and such that $|J|=N$. We have to prove that $\sigma^{-1}(J)$ is an interval. This is obvious if $|J|=1$ (because in this case, $\sigma^{-1}(J)$ is a one-element set, thus trivially an interval). Hence, we WLOG assume that we don't have $|J|=1$. We also don't have $|J|=0$, because $(uv)[J]$ has to be Lyndon (and the empty word is not). So we have $|J|>1$. Now, $\ell((uv)[J])=|J|>1$, and thus $(uv)[J]$ is a Lyndon word of length >1. Let v' be the (lexicographically) smallest nonempty **proper** suffix of $(uv)[J]$. Since v' is a proper suffix of w, there exists a nonempty $u'\in\mathfrak{A}^{*}$ such that $(uv)[J]=u'v'$. Consider this u'.

Now, Theorem 6.1.30(a) (applied to $(uv)[J]$, u' and v' instead of w, u and v) yields that the words u' and v' are Lyndon. Also, Theorem 6.1.30(b) (applied to $(uv)[J]$, u' and v' instead of w, u and v) yields that $u'<(uv)[J]<v'$.

But from the fact that $(uv)[J]=u'v'$ with u' and v' both being nonempty, it becomes immediately clear that we can write J as a union of two disjoint nonempty intervals K and L such that $K<L$, $u'=(uv)[K]$ and $v'=(uv)[L]$. Consider these K and L. The intervals K and L are nonempty and have their sizes add up to $|J|$ (since they are disjoint and their union is J), and hence both must have size smaller than $|J|=N$. So $K\subset[0:n]^{+}$ is an interval of size $|K|<N$ having the property that $(uv)[K]$ is Lyndon (since $(uv)[K]=u'$ is Lyndon). Thus, we can apply (6.2.6) to K instead of J (because of the induction hypothesis). As a result, we conclude that $\sigma^{-1}(K)$ is an interval. Similarly, we can apply (6.2.6) to L instead of J (we know that $(uv)[L]$ is Lyndon since $(uv)[L]=v'$), and learn that $\sigma^{-1}(L)$ is an interval. The intervals $\sigma^{-1}(K)$ and $\sigma^{-1}(L)$ are both nonempty (since K and L are nonempty), and their union is $\sigma^{-1}(J)$ (because the union of K and L is J). The nonempty intervals K and L both are subsets of $[0:n]^{+}$ (since their union is $J\subset[0:n]^{+}$), and their union $K\cup L$ is an interval (since their union $K\cup L$ is J, and we know that J is an interval).

[305] Indeed, when $i\leq p$, this follows from (6.2.3) and the fact that a_i is Lyndon; whereas in the other case, this follows from (6.2.4) and the fact that b_{i-p} is Lyndon.

Now, assume (for the sake of contradiction) that $\sigma^{-1}(J)$ is not an interval. Since J is the union of K and L, we have $J = K \cup L$ and thus $\sigma^{-1}(J) = \sigma^{-1}(K \cup L) = \sigma^{-1}(K) \cup \sigma^{-1}(L)$ (since σ is a bijection). Therefore, $\sigma^{-1}(K) \cup \sigma^{-1}(L)$ is not an interval (since $\sigma^{-1}(J)$ is not an interval). Thus, Lemma 6.2.8(b) yields that there exists a nonempty interval $P \subset [n : n+m]^+$ such that $\sigma^{-1}(P)$, $\sigma^{-1}(K) \cup \sigma^{-1}(P)$ and $\sigma^{-1}(P) \cup \sigma^{-1}(L)$ are intervals and such that $\sigma^{-1}(K) < \sigma^{-1}(P) < \sigma^{-1}(L)$. Consider this P. Since P is nonempty, we have $|P| \neq 0$.

Lemma 6.2.10(b) (applied to K and P instead of I and J) yields

$$(uv)[K] \cdot (uv)[P] \geq (uv)[P] \cdot (uv)[K]. \tag{6.2.7}$$

Since $(uv)[K] = u'$, this rewrites as

$$u' \cdot (uv)[P] \geq (uv)[P] \cdot u'. \tag{6.2.8}$$

But Lemma 6.2.10(c) (applied to P and L instead of I and J) yields

$$(uv)[P] \cdot (uv)[L] \geq (uv)[L] \cdot (uv)[P]. \tag{6.2.9}$$

Since $(uv)[L] = v'$, this rewrites as

$$(uv)[P] \cdot v' \geq v' \cdot (uv)[P]. \tag{6.2.10}$$

Recall also that $u' < v'$, and that both words u' and v' are Lyndon. Now, Corollary 6.1.17 (applied to u', v' and $(uv)[P]$ instead of u, v and z) yields that $(uv)[P]$ is the empty word (because of (6.2.8) and (6.2.10)), so that $\ell((uv)[P]) = 0$. This contradicts $\ell((uv)[P]) = |P| \neq 0$. This contradiction shows that our assumption (that $\sigma^{-1}(J)$ is not an interval) was wrong. Hence, $\sigma^{-1}(J)$ is an interval. This completes the induction step, and thus (6.2.6) is proven.

Similarly to (6.2.6), we can show that

(6.2.11)

$$\text{if } J \subset [n : n+m]^+ \text{ is an interval such that the word } (uv)[J] \text{ is Lyndon, then } \sigma^{-1}(J) \text{ is an interval.}$$

Now, let $i \in \{1, 2, \ldots, p+q\}$ be arbitrary. We are going to prove that

$$\sigma^{-1}(I_i) \text{ is an interval.} \tag{6.2.12}$$

Proof of (6.2.12): We must be in one of the following two cases:

Case 1: We have $i \in \{1, 2, \ldots, p\}$.

Case 2: We have $i \in \{p+1, p+2, \ldots, p+q\}$.

Let us first consider Case 1. In this case, we have $i \in \{1, 2, \ldots, p\}$. Thus, $I_i \subset [0 : n]^+$ (by (6.2.3)). Also, (6.2.3) yields that $(uv)[I_i] = a_i$ is a Lyndon word. Hence, (6.2.6) (applied to $J = I_i$) yields that $\sigma^{-1}(I_i)$ is an interval. Thus, (6.2.12) is proven in Case 1.

Similarly, we can prove (6.2.12) in Case 2, using (6.2.4) and (6.2.11) instead of (6.2.3) and (6.2.6), respectively. Hence, (6.2.12) is proven.

So we know that $\sigma^{-1}(I_i)$ is an interval. But we also know that either $I_i \subset [0 : n]^+$ or $I_i \subset [n : n+m]^+$ (depending on whether $i \leq p$ or $i > p$). As a consequence, the restriction of the map σ^{-1} to the interval I_i is increasing (because the restriction of the map σ^{-1} to the interval $[0 : n]^+$ is strictly increasing, and so is the restriction of the map σ^{-1} to the interval $[n : n+m]^+$).

Now, let us forget that we fixed i. We thus have shown that every $i \in \{1, 2, \ldots, p+q\}$ has the two properties that:

- the set $\sigma^{-1}(I_i)$ is an interval;
- the restriction of the map σ^{-1} to the interval I_i is increasing.

In other words, the permutation σ is $(\alpha\beta)$-clumping (since $(I_1, I_2, \ldots, I_{p+q})$ is the interval system corresponding to the composition $\alpha\beta$). Hence, Proposition 6.2.16(b) (applied to $n+m$, $\alpha\beta$ and $p+q$ instead of n, α and ℓ) shows that there exists a unique $\tau \in \mathfrak{S}_{p+q}$ satisfying $\sigma = \operatorname{iper}(\alpha\beta, \tau)$. Thus, the uniqueness part of Theorem 6.2.22(b) (i.e., the claim that the τ in Theorem 6.2.22(b) is unique if it exists) is proven.

It now remains to prove the existence part of Theorem 6.2.22(b), i.e., to prove that there exists at least one permutation $\tau \in \operatorname{Sh}_{p,q}$ satisfying $(uv)\left[I_{\tau(1)}\right] \geq (uv)\left[I_{\tau(2)}\right] \geq \cdots \geq (uv)\left[I_{\tau(p+q)}\right]$ and $\sigma = \operatorname{iper}(\alpha\beta, \tau)$. We already know that there exists a unique $\tau \in \mathfrak{S}_{p+q}$ satisfying $\sigma = \operatorname{iper}(\alpha\beta, \tau)$. Consider this τ. We will now prove that $(uv)\left[I_{\tau(1)}\right] \geq (uv)\left[I_{\tau(2)}\right] \geq \cdots \geq (uv)\left[I_{\tau(p+q)}\right]$ and $\tau \in \operatorname{Sh}_{p,q}$. Once this is done, the existence part of Theorem 6.2.22(b) will be proven, and thus the proof of Theorem 6.2.22(b) will be complete.

Proposition 6.2.18 yields that $\tau \in \operatorname{Sh}_{p,q}$ if and only if $\operatorname{iper}(\alpha\beta, \tau) \in \operatorname{Sh}_{n,m}$. Since we know that $\operatorname{iper}(\alpha\beta, \tau) = \sigma \in \operatorname{Sh}_{n,m}$, we thus conclude that $\tau \in \operatorname{Sh}_{p,q}$. The only thing that remains to be proven now is that

$$(uv)\left[I_{\tau(1)}\right] \geq (uv)\left[I_{\tau(2)}\right] \geq \cdots \geq (uv)\left[I_{\tau(p+q)}\right]. \tag{6.2.13}$$

Proof of (6.2.13): We have $\tau \in \operatorname{Sh}_{p,q}$. In other words, $\tau^{-1}(1) < \tau^{-1}(2) < \cdots < \tau^{-1}(p)$ and $\tau^{-1}(p+1) < \tau^{-1}(p+2) < \cdots < \tau^{-1}(p+q)$. In other words, the restriction of the map τ^{-1} to the interval $[0:p]^+$ is strictly increasing, and so is the restriction of the map τ^{-1} to the interval $[p:p+q]^+$.

Let $i \in \{1, 2, \ldots, p+q-1\}$. We will show that

$$(uv)\left[I_{\tau(i)}\right] \geq (uv)\left[I_{\tau(i+1)}\right]. \tag{6.2.14}$$

Clearly, both $\tau(i)$ and $\tau(i+1)$ belong to $\{1, 2, \ldots, p+q\} = \{1, 2, \ldots, p\} \cup \{p+1, p+2, \ldots, p+q\}$. Thus, we must be in one of the following four cases:

Case 1: We have $\tau(i) \in \{1, 2, \ldots, p\}$ and $\tau(i+1) \in \{1, 2, \ldots, p\}$.

Case 2: We have $\tau(i) \in \{1, 2, \ldots, p\}$ and $\tau(i+1) \in \{p+1, p+2, \ldots, p+q\}$.

Case 3: We have $\tau(i) \in \{p+1, p+2, \ldots, p+q\}$ and $\tau(i+1) \in \{1, 2, \ldots, p\}$.

Case 4: We have $\tau(i) \in \{p+1, p+2, \ldots, p+q\}$ and $\tau(i+1) \in \{p+1, p+2, \ldots, p+q\}$.

Let us consider Case 1 first. In this case, we have $\tau(i) \in \{1, 2, \ldots, p\}$ and $\tau(i+1) \in \{1, 2, \ldots, p\}$. From the fact that the restriction of the map τ^{-1} to the interval $[0:p]^+$ is strictly increasing, we can easily deduce $\tau(i) < \tau(i+1)$ [306]. Therefore, $a_{\tau(i)} \geq a_{\tau(i+1)}$ (since $a_1 \geq a_2 \geq \cdots \geq a_p$).

But $(uv)\left[I_{\tau(i)}\right] = a_{\tau(i)}$ (by (6.2.3), applied to $\tau(i)$ instead of i) and $(uv)\left[I_{\tau(i+1)}\right] = a_{\tau(i+1)}$ (similarly). In view of these equalities, the inequality $a_{\tau(i)} \geq a_{\tau(i+1)}$ rewrites as $(uv)\left[I_{\tau(i)}\right] \geq (uv)\left[I_{\tau(i+1)}\right]$. Thus, (6.2.14) is proven in Case 1.

Similarly, we can show (6.2.14) in Case 4 (observing that $(uv)\left[I_{\tau(i)}\right] = b_{\tau(i)-p}$ and $(uv)\left[I_{\tau(i+1)}\right] = b_{\tau(i+1)-p}$ in this case).

Let us now consider Case 2. In this case, we have $\tau(i) \in \{1, 2, \ldots, p\}$ and $\tau(i+1) \in \{p+1, p+2, \ldots, p+q\}$. From $\tau(i) \in \{1, 2, \ldots, p\}$, we conclude that $I_{\tau(i)} \subset [0:n]^+$. From $\tau(i+1) \in \{p+1, p+2, \ldots, p+q\}$, we conclude that $I_{\tau(i+1)} \subset [n:n+m]^+$. The intervals $I_{\tau(i)}$ and $I_{\tau(i+1)}$ are clearly nonempty.

Proposition 6.2.14(d) (applied to $n+m$, $\alpha\beta$, $p+q$ and $(I_1, I_2, \ldots, I_{p+q})$ instead of n, α, ℓ and $(I_1, I_2, \ldots, I_\ell)$) yields that the sets $\sigma^{-1}\left(I_{\tau(i)}\right)$, $\sigma^{-1}\left(I_{\tau(i+1)}\right)$ and $\sigma^{-1}\left(I_{\tau(i)}\right) \cup \sigma^{-1}\left(I_{\tau(i+1)}\right)$ are nonempty intervals, and that we have $\sigma^{-1}\left(I_{\tau(i)}\right) < \sigma^{-1}\left(I_{\tau(i+1)}\right)$. Hence, Lemma 6.2.10(b) (applied to $I = I_{\tau(i)}$ and $J = I_{\tau(i+1)}$) yields

$$(uv)\left[I_{\tau(i)}\right] \cdot (uv)\left[I_{\tau(i+1)}\right] \geq (uv)\left[I_{\tau(i+1)}\right] \cdot (uv)\left[I_{\tau(i)}\right].$$

But $(uv)\left[I_{\tau(i)}\right]$ and $(uv)\left[I_{\tau(i+1)}\right]$ are Lyndon words (as a consequence of (6.2.5)). Thus, Proposition 6.1.18 (applied to $(uv)\left[I_{\tau(i)}\right]$ and $(uv)\left[I_{\tau(i+1)}\right]$ instead of u and v) shows that $(uv)\left[I_{\tau(i)}\right] \geq (uv)\left[I_{\tau(i+1)}\right]$ if and only if $(uv)\left[I_{\tau(i)}\right] \cdot (uv)\left[I_{\tau(i+1)}\right] \geq (uv)\left[I_{\tau(i+1)}\right] \cdot (uv)\left[I_{\tau(i)}\right]$. Since we know that $(uv)\left[I_{\tau(i)}\right] \cdot (uv)\left[I_{\tau(i+1)}\right] \geq (uv)\left[I_{\tau(i+1)}\right] \cdot (uv)\left[I_{\tau(i)}\right]$ holds, we thus conclude that $(uv)\left[I_{\tau(i)}\right] \geq (uv)\left[I_{\tau(i+1)}\right]$. Thus, (6.2.14) is proven in Case 2.

The proof of (6.2.14) in Case 3 is analogous to that in Case 2 (the main difference being that Lemma 6.2.10(c) is used in lieu of Lemma 6.2.10(b)).

Thus, (6.2.14) is proven in all possible cases. So we always have (6.2.14). In other words, $(uv)\left[I_{\tau(i)}\right] \geq (uv)\left[I_{\tau(i+1)}\right]$.

Now, forget that we fixed i. We hence have shown that $(uv)\left[I_{\tau(i)}\right] \geq (uv)\left[I_{\tau(i+1)}\right]$ for all $i \in \{1, 2, \ldots, p+q-1\}$. This proves (6.2.13), and thus completes our proof of Theorem 6.2.22(b).

(a) Let $\tau \in \operatorname{Sh}_{p,q}$ be such that

$$(uv)\left[I_{\tau(1)}\right] \geq (uv)\left[I_{\tau(2)}\right] \geq \cdots \geq (uv)\left[I_{\tau(p+q)}\right]. \tag{6.2.15}$$

Set $\sigma = \operatorname{iper}(\alpha\beta, \tau)$. Then, Proposition 6.2.18 yields that $\tau \in \operatorname{Sh}_{p,q}$ if and only if $\operatorname{iper}(\alpha\beta, \tau) \in \operatorname{Sh}_{n,m}$. Since we know that $\tau \in \operatorname{Sh}_{p,q}$, we can deduce from this that $\operatorname{iper}(\alpha\beta, \tau) \in \operatorname{Sh}_{n,m}$, so that $\sigma = \operatorname{iper}(\alpha\beta, \tau) \in \operatorname{Sh}_{n,m}$.

It remains to prove that the word $u \underset{\sigma}{\sqcup\!\sqcup} v$ is the lexicographically highest element of the multiset $u \sqcup\!\sqcup v$.

[306] *Proof.* Assume the contrary. Then, $\tau(i) \geq \tau(i+1)$. Since both $\tau(i)$ and $\tau(i+1)$ belong to $\{1, 2, \ldots, p\} = [0:p]^+$, this yields $\tau^{-1}(\tau(i)) \geq \tau^{-1}(\tau(i+1))$ (since the restriction of the map τ^{-1} to the interval $[0:p]^+$ is strictly increasing), which contradicts $\tau^{-1}(\tau(i)) = i < i+1 = \tau^{-1}(\tau(i+1))$. This contradiction proves the assumption wrong, qed.

It is clear that the multiset $u \sqcup\!\sqcup v$ has **some** lexicographically highest element. This element has the form $u \underset{\widetilde{\sigma}}{\sqcup\!\sqcup} v$ for some $\widetilde{\sigma} \in \operatorname{Sh}_{n,m}$ (because any element of this multiset has such a form). Consider this $\widetilde{\sigma}$. Theorem 6.2.22(b) (applied to $\widetilde{\sigma}$ instead of σ) yields that there exists a unique permutation $\widetilde{\tau} \in \operatorname{Sh}_{p,q}$ satisfying $(uv)\left[I_{\widetilde{\tau}(1)}\right] \geq (uv)\left[I_{\widetilde{\tau}(2)}\right] \geq \cdots \geq (uv)\left[I_{\widetilde{\tau}(p+q)}\right]$ and $\widetilde{\sigma} = \operatorname{iper}(\alpha\beta, \widetilde{\tau})$. (What we call $\widetilde{\tau}$ here is what has been called τ in Theorem 6.2.22(b).)

Now, the chain of inequalities $(uv)\left[I_{\widetilde{\tau}(1)}\right] \geq (uv)\left[I_{\widetilde{\tau}(2)}\right] \geq \cdots \geq (uv)\left[I_{\widetilde{\tau}(p+q)}\right]$ shows that the list $\left((uv)\left[I_{\widetilde{\tau}(1)}\right], (uv)\left[I_{\widetilde{\tau}(2)}\right], \ldots, (uv)\left[I_{\widetilde{\tau}(p+q)}\right]\right)$ is the result of sorting the list $((uv)[I_1], (uv)[I_2], \ldots, (uv)[I_{p+q}])$ in decreasing order. But the chain of inequalities (6.2.15) shows that the list $\left((uv)\left[I_{\tau(1)}\right], (uv)\left[I_{\tau(2)}\right], \ldots, (uv)\left[I_{\tau(p+q)}\right]\right)$ is the result of sorting the same list $((uv)[I_1], (uv)[I_2], \ldots, (uv)[I_{p+q}])$ in decreasing order. So each of the two lists $\left((uv)\left[I_{\widetilde{\tau}(1)}\right], (uv)\left[I_{\widetilde{\tau}(2)}\right], \ldots, (uv)\left[I_{\widetilde{\tau}(p+q)}\right]\right)$ and $\left((uv)\left[I_{\tau(1)}\right], (uv)\left[I_{\tau(2)}\right], \ldots, (uv)\left[I_{\tau(p+q)}\right]\right)$ is the result of sorting one and the same list $((uv)[I_1], (uv)[I_2], \ldots, (uv)[I_{p+q}])$ in decreasing order. Since the result of sorting a given list in decreasing order is unique, this yields

$$\left((uv)\left[I_{\widetilde{\tau}(1)}\right], (uv)\left[I_{\widetilde{\tau}(2)}\right], \ldots, (uv)\left[I_{\widetilde{\tau}(p+q)}\right]\right) = \left((uv)\left[I_{\tau(1)}\right], (uv)\left[I_{\tau(2)}\right], \ldots, (uv)\left[I_{\tau(p+q)}\right]\right).$$

Hence,

$$(uv)\left[I_{\widetilde{\tau}(1)}\right] \cdot (uv)\left[I_{\widetilde{\tau}(2)}\right] \cdot \cdots \cdot (uv)\left[I_{\widetilde{\tau}(p+q)}\right] = (uv)\left[I_{\tau(1)}\right] \cdot (uv)\left[I_{\tau(2)}\right] \cdot \cdots \cdot (uv)\left[I_{\tau(p+q)}\right]. \tag{6.2.16}$$

But Lemma 6.2.20 yields

$$u \underset{\sigma}{\sqcup\!\sqcup} v = (uv)\left[I_{\tau(1)}\right] \cdot (uv)\left[I_{\tau(2)}\right] \cdot \cdots \cdot (uv)\left[I_{\tau(p+q)}\right]. \tag{6.2.17}$$

Meanwhile, Lemma 6.2.20 (applied to $\widetilde{\tau}$ and $\widetilde{\sigma}$ instead of τ and σ) yields

$$\begin{aligned} u \underset{\widetilde{\sigma}}{\sqcup\!\sqcup} v &= (uv)\left[I_{\widetilde{\tau}(1)}\right] \cdot (uv)\left[I_{\widetilde{\tau}(2)}\right] \cdot \cdots \cdot (uv)\left[I_{\widetilde{\tau}(p+q)}\right] \\ &= (uv)\left[I_{\tau(1)}\right] \cdot (uv)\left[I_{\tau(2)}\right] \cdot \cdots \cdot (uv)\left[I_{\tau(p+q)}\right] \qquad (\text{by } (6.2.16)) \\ &= u \underset{\sigma}{\sqcup\!\sqcup} v \qquad (\text{by } (6.2.17)). \end{aligned}$$

Thus, $u \underset{\sigma}{\sqcup\!\sqcup} v$ is the lexicographically highest element of the multiset $u \sqcup\!\sqcup v$ (since we know that $u \underset{\widetilde{\sigma}}{\sqcup\!\sqcup} v$ is the lexicographically highest element of the multiset $u \sqcup\!\sqcup v$). This proves Theorem 6.2.22(a). □

Now, in order to prove Theorem 6.2.2, we record a very simple fact about counting shuffles:

Proposition 6.2.23. *Let $p \in \mathbb{N}$ and $q \in \mathbb{N}$. Let $\mathfrak{W}$ be a totally ordered set, and let $h : \{1, 2, \ldots, p+q\} \to \mathfrak{W}$ be a map. Assume that $h(1) \geq h(2) \geq \cdots \geq h(p)$ and $h(p+1) \geq h(p+2) \geq \cdots \geq h(p+q)$.*

For every $w \in \mathfrak{W}$, let $\mathfrak{a}(w)$ denote the number of all $i \in \{1, 2, \ldots, p\}$ satisfying $h(i) = w$, and let $\mathfrak{b}(w)$ denote the number of all $i \in \{p+1, p+2, \ldots, p+q\}$ satisfying $h(i) = w$.

Then, the number of $\tau \in \operatorname{Sh}_{p,q}$ satisfying $h(\tau(1)) \geq h(\tau(2)) \geq \cdots \geq h(\tau(p+q))$ is $\prod_{w \in \mathfrak{W}} \dbinom{\mathfrak{a}(w) + \mathfrak{b}(w)}{\mathfrak{a}(w)}$. (Of course, all but finitely many factors of this product are 1.)

Exercise 6.2.24. Prove Proposition 6.2.23.

Proof of Theorem 6.2.2. Let $n = \ell(u)$ and $m = \ell(v)$. Define α, β and $(I_1, I_2, \ldots, I_{p+q})$ as in Theorem 6.2.22.

Since $(a_1, a_2, \ldots, a_p)$ is the CFL factorization of u, we have $a_1 \geq a_2 \geq \cdots \geq a_p$ and $a_1 a_2 \cdots a_p = u$. Similarly, $b_1 \geq b_2 \geq \cdots \geq b_q$ and $b_1 b_2 \cdots b_q = v$.

From (6.2.3), we see that $(uv)[I_i] = a_i$ for every $i \in \{1, 2, \ldots, p\}$. From (6.2.4), we see that $(uv)[I_i] = b_{i-p}$ for every $i \in \{p+1, p+2, \ldots, p+q\}$. Combining these two equalities, we obtain

$$(uv)[I_i] = \begin{cases} a_i, & \text{if } i \leq p; \\ b_{i-p}, & \text{if } i > p \end{cases} \qquad \text{for every } i \in \{1, 2, \ldots, p+q\}. \tag{6.2.18}$$

In other words,

$$((uv)[I_1], (uv)[I_2], \ldots, (uv)[I_{p+q}]) = (a_1, a_2, \ldots, a_p, b_1, b_2, \ldots, b_q). \tag{6.2.19}$$

(a) Let z be the lexicographically highest element of the multiset $u \sqcup\!\sqcup v$. We must prove that $z = c_1 c_2 \cdots c_{p+q}$.

Since $z \in u \sqcup\!\sqcup v$, we can write z in the form $u \underset{\sigma}{\sqcup\!\sqcup} v$ for some $\sigma \in \operatorname{Sh}_{n,m}$ (since we can write any element of $u \sqcup\!\sqcup v$ in this form). Consider this σ. Then, $u \underset{\sigma}{\sqcup\!\sqcup} v = z$ is the lexicographically highest element of the multiset $u \sqcup\!\sqcup v$. Hence, Theorem 6.2.22(b) yields that there exists a unique permutation $\tau \in \operatorname{Sh}_{p,q}$ satisfying $(uv)\left[I_{\tau(1)}\right] \geq (uv)\left[I_{\tau(2)}\right] \geq \cdots \geq (uv)\left[I_{\tau(p+q)}\right]$ and $\sigma = \operatorname{iper}(\alpha\beta, \tau)$. Consider this τ.

Now, $\tau \in \operatorname{Sh}_{p,q} \subset \mathfrak{S}_{p+q}$ is a permutation, and thus the list $\left((uv)\left[I_{\tau(1)}\right], (uv)\left[I_{\tau(2)}\right], \ldots, (uv)\left[I_{\tau(p+q)}\right]\right)$ is a rearrangement of the list $((uv)[I_1], (uv)[I_2], \ldots, (uv)[I_{p+q}])$. Due to (6.2.19), this rewrites as follows: The list $\left((uv)\left[I_{\tau(1)}\right], (uv)\left[I_{\tau(2)}\right], \ldots, (uv)\left[I_{\tau(p+q)}\right]\right)$ is a rearrangement of the list $(a_1, a_2, \ldots, a_p, b_1, b_2, \ldots, b_q)$. Hence, $\left((uv)\left[I_{\tau(1)}\right], (uv)\left[I_{\tau(2)}\right], \ldots, (uv)\left[I_{\tau(p+q)}\right]\right)$ is the result of sorting the list $(a_1, a_2, \ldots, a_p, b_1, b_2, \ldots, b_q)$ in decreasing order (since $(uv)\left[I_{\tau(1)}\right] \geq (uv)\left[I_{\tau(2)}\right] \geq \cdots \geq (uv)\left[I_{\tau(p+q)}\right]$). But since the result of sorting the list $(a_1, a_2, \ldots, a_p, b_1, b_2, \ldots, b_q)$ in decreasing order is $(c_1, c_2, \ldots, c_{p+q})$, this becomes

$$\left((uv)\left[I_{\tau(1)}\right], (uv)\left[I_{\tau(2)}\right], \ldots, (uv)\left[I_{\tau(p+q)}\right]\right) = (c_1, c_2, \ldots, c_{p+q}).$$

Hence,

$$(uv)\left[I_{\tau(1)}\right] \cdot (uv)\left[I_{\tau(2)}\right] \cdot \cdots \cdot (uv)\left[I_{\tau(p+q)}\right] = c_1 \cdot c_2 \cdot \cdots \cdot c_{p+q}.$$

But Lemma 6.2.20 yields

$$u \underset{\sigma}{\sqcup\!\sqcup} v = (uv)\left[I_{\tau(1)}\right] \cdot (uv)\left[I_{\tau(2)}\right] \cdot \cdots \cdot (uv)\left[I_{\tau(p+q)}\right].$$

Altogether, we have

$$z = u \underset{\sigma}{\sqcup\!\sqcup} v = (uv)\left[I_{\tau(1)}\right] \cdot (uv)\left[I_{\tau(2)}\right] \cdot \cdots \cdot (uv)\left[I_{\tau(p+q)}\right] = c_1 \cdot c_2 \cdot \cdots \cdot c_{p+q} = c_1 c_2 \cdots c_{p+q}.$$

This proves Theorem 6.2.2(a).

(b) Recall that $u \sqcup\!\sqcup v = \left\{ u \underset{\sigma}{\sqcup\!\sqcup} v \ : \ \sigma \in \operatorname{Sh}_{n,m} \right\}_{\text{multiset}}$. Hence,

$$\begin{aligned}
&(\text{the multiplicity with which the lexicographically highest element of the multiset} \\
&\qquad u \sqcup\!\sqcup v \text{ appears in the multiset } u \sqcup\!\sqcup v) \\
&= \Big(\text{the number of all } \sigma \in \operatorname{Sh}_{n,m} \text{ such that } u \underset{\sigma}{\sqcup\!\sqcup} v \text{ is the} \\
&\qquad \text{lexicographically highest element of the multiset } u \sqcup\!\sqcup v \big).
\end{aligned}$$

However, for a given $\sigma \in \operatorname{Sh}_{n,m}$, we know that $u \underset{\sigma}{\sqcup\!\sqcup} v$ is the lexicographically highest element of the multiset $u \sqcup\!\sqcup v$ if and only if σ can be written in the form $\sigma = \operatorname{iper}(\alpha\beta, \tau)$ for some $\tau \in \operatorname{Sh}_{p,q}$ satisfying $(uv)\left[I_{\tau(1)}\right] \geq (uv)\left[I_{\tau(2)}\right] \geq \cdots \geq (uv)\left[I_{\tau(p+q)}\right]$. [307] Hence,

$$\begin{aligned}
&\Big(\text{the number of all } \sigma \in \operatorname{Sh}_{n,m} \text{ such that } u \underset{\sigma}{\sqcup\!\sqcup} v \text{ is the} \\
&\qquad \text{lexicographically highest element of the multiset } u \sqcup\!\sqcup v \big) \\
&= (\text{the number of all } \sigma \in \operatorname{Sh}_{n,m} \text{ which can be written in the form } \sigma = \operatorname{iper}(\alpha\beta, \tau) \\
&\qquad \text{for some } \tau \in \operatorname{Sh}_{p,q} \text{ satisfying } (uv)\left[I_{\tau(1)}\right] \geq (uv)\left[I_{\tau(2)}\right] \geq \cdots \geq (uv)\left[I_{\tau(p+q)}\right]\big) \\
&= \big(\text{the number of all } \tau \in \operatorname{Sh}_{p,q} \text{ satisfying } (uv)\left[I_{\tau(1)}\right] \geq (uv)\left[I_{\tau(2)}\right] \geq \cdots \geq (uv)\left[I_{\tau(p+q)}\right]\big)
\end{aligned}$$

[307] In fact, the "if" part of this assertion follows from Theorem 6.2.22(a), whereas its "only if" part follows from Theorem 6.2.22(b).

(because if a $\sigma \in \operatorname{Sh}_{n,m}$ can be written in the form $\sigma = \operatorname{iper}(\alpha\beta, \tau)$ for some $\tau \in \operatorname{Sh}_{p,q}$ satisfying $(uv)\left[I_{\tau(1)}\right] \geq (uv)\left[I_{\tau(2)}\right] \geq \cdots \geq (uv)\left[I_{\tau(p+q)}\right]$, then σ can be written **uniquely** in this form[308]). Thus,

$$\begin{aligned}
&(\text{the multiplicity with which the lexicographically highest element of the multiset}\\
&\qquad u \sqcup\!\sqcup v \text{ appears in the multiset } u \sqcup\!\sqcup v)\\
&= \Big(\text{the number of all } \sigma \in \operatorname{Sh}_{n,m} \text{ such that } u \underset{\sigma}{\sqcup\!\sqcup} v \text{ is the}\\
&\qquad \text{lexicographically highest element of the multiset } u \sqcup\!\sqcup v)\\
&= \left(\text{the number of all } \tau \in \operatorname{Sh}_{p,q} \text{ satisfying } (uv)\left[I_{\tau(1)}\right] \geq (uv)\left[I_{\tau(2)}\right] \geq \cdots \geq (uv)\left[I_{\tau(p+q)}\right]\right). \qquad (6.2.20)
\end{aligned}$$

Now, define a map $h : \{1, 2, \ldots, p+q\} \to \mathfrak{L}$ by

$$h(i) = \begin{cases} a_i, & \text{if } i \leq p; \\ b_{i-p}, & \text{if } i > p \end{cases} \qquad \text{for every } i \in \{1, 2, \ldots, p+q\}.$$

Then, $h(1) \geq h(2) \geq \cdots \geq h(p)$ (because this is just a rewriting of $a_1 \geq a_2 \geq \cdots \geq a_p$) and $h(p+1) \geq h(p+2) \geq \cdots \geq h(p+q)$ (since this is just a rewriting of $b_1 \geq b_2 \geq \cdots \geq b_q$). For every $w \in \mathfrak{L}$, the number of all $i \in \{1, 2, \ldots, p\}$ satisfying $h(i) = w$ is

$$\begin{aligned}
&\left|\left\{ i \in \{1, 2, \ldots, p\} \mid \underbrace{h(i)}_{=a_i} = w \right\}\right|\\
&= |\{i \in \{1, 2, \ldots, p\} \mid a_i = w\}|\\
&= (\text{the number of terms in the list } (a_1, a_2, \ldots, a_p) \text{ which are equal to } w)\\
&= (\text{the number of terms in the CFL factorization of } u \text{ which are equal to } w)\\
&\qquad (\text{since the list } (a_1, a_2, \ldots, a_p) \text{ is the CFL factorization of } u)\\
&= \operatorname{mult}_w u
\end{aligned}$$

(because $\operatorname{mult}_w u$ is defined as the number of terms in the CFL factorization of u which are equal to w). Similarly, for every $w \in \mathfrak{L}$, the number of all $i \in \{p+1, p+2, \ldots, p+q\}$ satisfying $h(i) = w$ equals $\operatorname{mult}_w v$. Thus, we can apply Proposition 6.2.23 to $\mathfrak{W} = \mathfrak{L}$, $\mathfrak{a}(w) = \operatorname{mult}_w u$ and $\mathfrak{b}(w) = \operatorname{mult}_w v$. As a result, we see that the number of $\tau \in \operatorname{Sh}_{p,q}$ satisfying $h(\tau(1)) \geq h(\tau(2)) \geq \cdots \geq h(\tau(p+q))$ is $\prod_{w \in \mathfrak{L}} \binom{\operatorname{mult}_w u + \operatorname{mult}_w v}{\operatorname{mult}_w u}$. In other words,

$$\begin{aligned}
&(\text{the number of all } \tau \in \operatorname{Sh}_{p,q} \text{ satisfying } h(\tau(1)) \geq h(\tau(2)) \geq \cdots \geq h(\tau(p+q)))\\
&= \prod_{w \in \mathfrak{L}} \binom{\operatorname{mult}_w u + \operatorname{mult}_w v}{\operatorname{mult}_w u}. \qquad (6.2.21)
\end{aligned}$$

However, for every $i \in \{1, 2, \ldots, p+q\}$, we have

$$h(i) = \begin{cases} a_i, & \text{if } i \leq p; \\ b_{i-p}, & \text{if } i > p \end{cases} = (uv)[I_i] \qquad (\text{by } (6.2.18)).$$

[308] *Proof.* Let $\sigma \in \operatorname{Sh}_{n,m}$ be such that σ can be written in the form $\sigma = \operatorname{iper}(\alpha\beta, \tau)$ for some $\tau \in \operatorname{Sh}_{p,q}$ satisfying $(uv)\left[I_{\tau(1)}\right] \geq (uv)\left[I_{\tau(2)}\right] \geq \cdots \geq (uv)\left[I_{\tau(p+q)}\right]$. Then, the word $u \underset{\sigma}{\sqcup\!\sqcup} v$ is the lexicographically highest element of the multiset $u \sqcup\!\sqcup v$ (according to Theorem 6.2.22(a)). Hence, there exists a unique permutation $\tau \in \operatorname{Sh}_{p,q}$ satisfying $(uv)\left[I_{\tau(1)}\right] \geq (uv)\left[I_{\tau(2)}\right] \geq \cdots \geq (uv)\left[I_{\tau(p+q)}\right]$ and $\sigma = \operatorname{iper}(\alpha\beta, \tau)$ (according to Theorem 6.2.22(b)). In other words, σ can be written **uniquely** in the form $\sigma = \operatorname{iper}(\alpha\beta, \tau)$ for some $\tau \in \operatorname{Sh}_{p,q}$ satisfying $(uv)\left[I_{\tau(1)}\right] \geq (uv)\left[I_{\tau(2)}\right] \geq \cdots \geq (uv)\left[I_{\tau(p+q)}\right]$, qed.

Hence, for any $\tau \in \operatorname{Sh}_{p,q}$, the condition $h(\tau(1)) \geq h(\tau(2)) \geq \cdots \geq h(\tau(p+q))$ is equivalent to $(uv)\left[I_{\tau(1)}\right] \geq (uv)\left[I_{\tau(2)}\right] \geq \cdots \geq (uv)\left[I_{\tau(p+q)}\right]$. Thus,

$$\begin{aligned}
&\left(\text{the number of all } \tau \in \operatorname{Sh}_{p,q} \text{ satisfying } \underbrace{h(\tau(1)) \geq h(\tau(2)) \geq \cdots \geq h(\tau(p+q))}_{\substack{\text{this is equivalent to} \\ (uv)\left[I_{\tau(1)}\right] \geq (uv)\left[I_{\tau(2)}\right] \geq \cdots \geq (uv)\left[I_{\tau(p+q)}\right]}}\right) \\
&= \left(\text{the number of all } \tau \in \operatorname{Sh}_{p,q} \text{ satisfying } (uv)\left[I_{\tau(1)}\right] \geq (uv)\left[I_{\tau(2)}\right] \geq \cdots \geq (uv)\left[I_{\tau(p+q)}\right]\right) \\
&= (\text{the multiplicity with which the lexicographically highest element of the multiset} \\
&\qquad u \sqcup\!\sqcup v \text{ appears in the multiset } u \sqcup\!\sqcup v)
\end{aligned}$$

(by (6.2.20)). Compared with (6.2.21), this yields

$$\begin{aligned}
&(\text{the multiplicity with which the lexicographically highest element of the multiset} \\
&\qquad u \sqcup\!\sqcup v \text{ appears in the multiset } u \sqcup\!\sqcup v) \\
&= \prod_{w \in \mathfrak{L}} \binom{\operatorname{mult}_w u + \operatorname{mult}_w v}{\operatorname{mult}_w u}.
\end{aligned}$$

This proves Theorem 6.2.2(b).

(c) We shall use the notations of Theorem 6.2.2(a) and Theorem 6.2.2(b).

Assume that $a_i \geq b_j$ for every $i \in \{1, 2, \ldots, p\}$ and $j \in \{1, 2, \ldots, q\}$. This, combined with $a_1 \geq a_2 \geq \cdots \geq a_p$ and $b_1 \geq b_2 \geq \cdots \geq b_q$, yields that $a_1 \geq a_2 \geq \cdots \geq a_p \geq b_1 \geq b_2 \geq \cdots \geq b_q$. Thus, the list $(a_1, a_2, \ldots, a_p, b_1, b_2, \ldots, b_q)$ is weakly decreasing. Thus, the result of sorting the list $(a_1, a_2, \ldots, a_p, b_1, b_2, \ldots, b_q)$ in decreasing order is the list $(a_1, a_2, \ldots, a_p, b_1, b_2, \ldots, b_q)$ itself. But since this result is $(c_1, c_2, \ldots, c_{p+q})$, this shows that $(c_1, c_2, \ldots, c_{p+q}) = (a_1, a_2, \ldots, a_p, b_1, b_2, \ldots, b_q)$. Hence, $c_1 c_2 \cdots c_{p+q} = \underbrace{a_1 a_2 \cdots a_p}_{=u} \underbrace{b_1 b_2 \cdots b_q}_{=v} = uv$. Now, Theorem 6.2.2(a) yields that the lexicographically highest element of the multiset $u \sqcup\!\sqcup v$ is $c_1 c_2 \cdots c_{p+q} = uv$. This proves Theorem 6.2.2(c).

(d) We shall use the notations of Theorem 6.2.2(a) and Theorem 6.2.2(b).

Assume that $a_i > b_j$ for every $i \in \{1, 2, \ldots, p\}$ and $j \in \{1, 2, \ldots, q\}$. Thus, $a_i \geq b_j$ for every $i \in \{1, 2, \ldots, p\}$ and $j \in \{1, 2, \ldots, q\}$. Hence, Theorem 6.2.2(c) yields that the lexicographically highest element of the multiset $u \sqcup\!\sqcup v$ is uv. Therefore, Theorem 6.2.2(b) shows that the multiplicity with which this word uv appears in the multiset $u \sqcup\!\sqcup v$ is $\prod_{w \in \mathfrak{L}} \binom{\operatorname{mult}_w u + \operatorname{mult}_w v}{\operatorname{mult}_w u}$.

Now, every $w \in \mathfrak{L}$ satisfies $\binom{\operatorname{mult}_w u + \operatorname{mult}_w v}{\operatorname{mult}_w u} = 1$ [309]. Thus, as we know, the multiplicity with which this word uv appears in the multiset $u \sqcup\!\sqcup v$ is $\prod_{w \in \mathfrak{L}} \underbrace{\binom{\operatorname{mult}_w u + \operatorname{mult}_w v}{\operatorname{mult}_w u}}_{=1} = \prod_{w \in \mathfrak{L}} 1 = 1$. This proves Theorem 6.2.2(d).

(e) We shall use the notations of Theorem 6.2.2(a) and Theorem 6.2.2(b).

Since u is a Lyndon word, the 1-tuple (u) is the CFL factorization of u. Hence, we can apply Theorem 6.2.2(c) to 1 and (u) instead of p and $(a_1, a_2, \ldots, a_p)$. As a result, we conclude that the lexicographically highest element of the multiset $u \sqcup\!\sqcup v$ is uv. It remains to prove that the multiplicity with which this word uv appears in the multiset $u \sqcup\!\sqcup v$ is $\operatorname{mult}_u v + 1$.

[309] *Proof.* Assume the contrary. Then, there exists at least one $w \in \mathfrak{L}$ such that $\binom{\operatorname{mult}_w u + \operatorname{mult}_w v}{\operatorname{mult}_w u} \neq 1$. Consider this w. Both $\operatorname{mult}_w u$ and $\operatorname{mult}_w v$ must be positive (since $\binom{\operatorname{mult}_w u + \operatorname{mult}_w v}{\operatorname{mult}_w u} \neq 1$). Since $\operatorname{mult}_w u$ is positive, there must be at least one term in the CFL factorization of u which is equal to w. In other words, there is at least one $i \in \{1, 2, \ldots, p\}$ satisfying $a_i = w$ (since $(a_1, a_2, \ldots, a_p)$ is the CFL factorization of u). Similarly, there is at least one $j \in \{1, 2, \ldots, q\}$ satisfying $b_j = w$. These i and j satisfy $a_i = w = b_j$, which contradicts $a_i > b_j$. This contradiction shows that our assumption was false, qed.

For every $w \in \mathfrak{L}$ satisfying $w \neq u$, we have

$$\operatorname{mult}_w u = 0 \tag{6.2.22}$$

[310]. Also, $\operatorname{mult}_u u = 1$ (for a similar reason). But uv is the lexicographically highest element of the multiset $u \sqcup\!\sqcup v$. Hence, the multiplicity with which the word uv appears in the multiset $u \sqcup\!\sqcup v$ is the multiplicity with which the lexicographically highest element of the multiset $u \sqcup\!\sqcup v$ appears in the multiset $u \sqcup\!\sqcup v$. According to Theorem 6.2.2(b), the latter multiplicity is

$$\begin{aligned}
&\prod_{w\in\mathfrak{L}} \binom{\operatorname{mult}_w u + \operatorname{mult}_w v}{\operatorname{mult}_w u} \\
&= \underbrace{\binom{\operatorname{mult}_u u + \operatorname{mult}_u v}{\operatorname{mult}_u u}}_{\substack{=\binom{1+\operatorname{mult}_u v}{1} \\ (\text{since } \operatorname{mult}_u u = 1)}} \cdot \prod_{\substack{w\in\mathfrak{L};\\ w\neq u}} \underbrace{\binom{\operatorname{mult}_w u + \operatorname{mult}_w v}{\operatorname{mult}_w u}}_{\substack{=\binom{0+\operatorname{mult}_w v}{0} \\ (\text{since } \operatorname{mult}_w u = 0 \text{ (by (6.2.22))})}} \qquad (\text{since } u \in \mathfrak{L}) \\
&= \underbrace{\binom{1+\operatorname{mult}_u v}{1}}_{=1+\operatorname{mult}_u v = \operatorname{mult}_u v + 1} \cdot \prod_{\substack{w\in\mathfrak{L};\\ w\neq u}} \underbrace{\binom{0+\operatorname{mult}_w v}{0}}_{=1} = \left(\operatorname{mult}_u v + 1\right) \cdot \underbrace{\prod_{\substack{w\in\mathfrak{L};\\ w\neq u}} 1}_{=1} = \operatorname{mult}_u v + 1.
\end{aligned}$$

This proves Theorem 6.2.2(e). □

As an application of our preceding results, we can prove a further necessary and sufficient criterion for a word to be Lyndon; this criterion is due to Chen/Fox/Lyndon [38, $\mathfrak{A}'' = \mathfrak{A}''''$]:

Exercise 6.2.25. Let $w \in \mathfrak{A}^*$ be a nonempty word. Prove that w is Lyndon if and only if for any two nonempty words $u \in \mathfrak{A}^*$ and $v \in \mathfrak{A}^*$ satisfying $w = uv$, there exists at least one $s \in u \sqcup\!\sqcup v$ satisfying $s > w$.

6.3. **Radford's theorem on the shuffle algebra.** We recall that our goal in Chapter 6 is to exhibit an algebraically independent generating set of the $\mathbf{k}$-algebra QSym. Having the notion of Lyndon words – which will, to some extent, but not literally, parametrize this generating set – in place, we could start the construction of this generating set immediately. However, it might come off as rather unmotivated this way, and so we begin with some warmups. First, we shall prove Radford's theorem on the shuffle algebra.

Definition 6.3.1. A *polynomial algebra* will mean a $\mathbf{k}$-algebra which is isomorphic to the polynomial ring $\mathbf{k}[x_i \mid i \in I]$ as a $\mathbf{k}$-algebra (for some indexing set I). Note that I need not be finite.

Equivalently, a polynomial algebra can be defined as a $\mathbf{k}$-algebra which has an algebraically independent (over $\mathbf{k}$) generating set. Yet equivalently, a polynomial algebra can be defined as a $\mathbf{k}$-algebra which is isomorphic to the symmetric algebra of a free $\mathbf{k}$-module.

Keep in mind that when we say that a certain bialgebra A is a polynomial algebra, we are making no statement about the coalgebra structure on A. The isomorphism from A to the symmetric algebra of a free $\mathbf{k}$-module need not be a coalgebra isomorphism, and the algebraically independent generating set of A need not consist of primitives. Thus, showing that a bialgebra A is a polynomial algebra does not trivialize the study of its bialgebraic structure.

Remark 6.3.2. Let V be a $\mathbf{k}$-module, and let $\mathfrak{A}$ be a totally ordered set. Let b_a be an element of V for every $a \in \mathfrak{A}$. Consider the shuffle algebra $\operatorname{Sh}(V)$ (defined in Definition 1.6.7).

For every word $w \in \mathfrak{A}^*$ over the alphabet $\mathfrak{A}$, let us define an element b_w of $\operatorname{Sh}(V)$ by $b_w = b_{w_1} b_{w_2} \cdots b_{w_\ell}$, where ℓ is the length of w. (The multiplication used here is that of $T(V)$, not that of $\operatorname{Sh}(V)$; the latter is denoted by $\underline{\sqcup\!\sqcup}$.)

Let $u \in \mathfrak{A}^*$ and $v \in \mathfrak{A}^*$ be two words over the alphabet $\mathfrak{A}$. Let $n = \ell(u)$ and $m = \ell(v)$. Then,

$$b_u \,\underline{\sqcup\!\sqcup}\, b_v = \sum_{\sigma \in \operatorname{Sh}_{n,m}} b_{u \underset{\sigma}{\sqcup\!\sqcup} v}.$$

[310] *Proof of (6.2.22):* Let $w \in \mathfrak{L}$ be such that $w \neq u$. Then, the number of terms in the list (u) which are equal to w is 0. Since (u) is the CFL factorization of u, this rewrites as follows: The number of terms in the CFL factorization of u which are equal to w is 0. In other words, $\operatorname{mult}_w u = 0$. This proves (6.2.22).

Exercise 6.3.3. Prove Remark 6.3.2.

[**Hint:** This follows from the definition of $\shuffle$.]

We can now state Radford's theorem [177, Theorem 3.1.1(e)]:

Theorem 6.3.4. *Assume that $\mathbb{Q}$ is a subring of $\mathbf{k}$. Let V be a free $\mathbf{k}$-module with a basis $(b_a)_{a\in\mathfrak{A}}$, where $\mathfrak{A}$ is a totally ordered set. Then, the shuffle algebra $\operatorname{Sh}(V)$ (defined in Definition 1.6.7) is a polynomial $\mathbf{k}$-algebra. An algebraically independent generating set of $\operatorname{Sh}(V)$ can be constructed as follows:*

For every word $w \in \mathfrak{A}^*$ over the alphabet $\mathfrak{A}$, let us define an element b_w of $\operatorname{Sh}(V)$ by $b_w = b_{w_1} b_{w_2} \cdots b_{w_\ell}$, where ℓ is the length of w. (The multiplication used here is that of $T(V)$, not that of $\operatorname{Sh}(V)$; the latter is denoted by $\shuffle$.) Let $\mathfrak{L}$ denote the set of all Lyndon words over the alphabet $\mathfrak{A}$. Then, $(b_w)_{w\in\mathfrak{L}}$ is an algebraically independent generating set of the $\mathbf{k}$-algebra $\operatorname{Sh}(V)$.

Example 6.3.5. For this example, let $\mathfrak{A}$ be the alphabet $\{1, 2, 3, \ldots\}$ with total order given by $1 < 2 < 3 < \cdots$, and assume that $\mathbb{Q}$ is a subring of $\mathbf{k}$. Let V be the free $\mathbf{k}$-module with basis $(b_a)_{a\in\mathfrak{A}}$. We use the notations of Theorem 6.3.4. Then, Theorem 6.3.4 yields that $(b_w)_{w\in\mathfrak{L}}$ is an algebraically independent generating set of the $\mathbf{k}$-algebra $\operatorname{Sh}(V)$. Here are some examples of elements of $\operatorname{Sh}(V)$ written as polynomials in this generating set:

$$\begin{aligned}
b_{12} &= b_{12} \qquad \text{(the word 12 itself is Lyndon)};\\
b_{21} &= b_1 \shuffle b_2 - b_{12};\\
b_{11} &= \frac{1}{2} b_1 \shuffle b_1;\\
b_{123} &= b_{123} \qquad \text{(the word 123 itself is Lyndon)};\\
b_{132} &= b_{132} \qquad \text{(the word 132 itself is Lyndon)};\\
b_{213} &= b_2 \shuffle b_{13} - b_{123} - b_{132};\\
b_{231} &= b_{23} \shuffle b_1 - b_2 \shuffle b_{13} + b_{132};\\
b_{312} &= b_3 \shuffle b_{12} - b_{123} - b_{132};\\
b_{321} &= b_1 \shuffle b_2 \shuffle b_3 - b_{23} \shuffle b_1 - b_3 \shuffle b_{12} + b_{123};\\
b_{112} &= b_{112} \qquad \text{(the word 112 itself is Lyndon)};\\
b_{121} &= b_{12} \shuffle b_1 - 2 b_{112};\\
b_{1212} &= \frac{1}{2} b_{12} \shuffle b_{12} - 2 b_{1122};\\
b_{4321} &= b_1 \shuffle b_2 \shuffle b_3 \shuffle b_4 - b_1 \shuffle b_2 \shuffle b_{34} - b_1 \shuffle b_{23} \shuffle b_4 - b_{12} \shuffle b_3 \shuffle b_4\\
&\qquad + b_1 \shuffle b_{234} + b_{12} \shuffle b_{34} + b_{123} \shuffle b_4 - b_{1234}.
\end{aligned}$$

[311]

Note that Theorem 6.3.4 cannot survive without the condition that $\mathbb{Q}$ be a subring of $\mathbf{k}$. For instance, for any $v \in V$, we have $v \shuffle v = 2vv$ in $\operatorname{Sh}(V)$, which vanishes if $2 = 0$ in $\mathbf{k}$; this stands in contrast to the fact that polynomial $\mathbf{k}$-algebras are integral domains when $\mathbf{k}$ itself is one. We will see that QSym is less sensitive towards the base ring in this regard (although proving that QSym is a polynomial algebra is much easier when $\mathbb{Q}$ is a subring of $\mathbf{k}$).

[311] A pattern emerges in the formulas for b_{21}, b_{321} and b_{4321}: for every $n \in \mathbb{N}$, we have

$$b_{(n,n-1,\ldots,1)} = \sum_{\alpha\in\operatorname{Comp}_n} (-1)^{n-\ell(\alpha)}\, b_{\mathbf{d}_1(\alpha)} \shuffle b_{\mathbf{d}_2(\alpha)} \shuffle \cdots \shuffle b_{\mathbf{d}_{\ell(\alpha)}(\alpha)},$$

where $(\mathbf{d}_1(\alpha)) \cdot (\mathbf{d}_2(\alpha)) \cdot \cdots \cdot (\mathbf{d}_{\ell(\alpha)}(\alpha))$ is the factorization of the word $(1, 2, \ldots, n)$ into factors of length $\alpha_1, \alpha_2, \ldots, \alpha_\ell$ (where $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_\ell)$). This can be proved by an application of Lemma 5.2.7(a) (as it is easy to see that for any composition α of n, we have

$$\begin{aligned}
b_{\mathbf{d}_1(\alpha)} \shuffle b_{\mathbf{d}_2(\alpha)} \shuffle \cdots \shuffle b_{\mathbf{d}_{\ell(\alpha)}(\alpha)} &= \left(\text{the sum of } b_\pi \text{ for all words } \pi \in \mathfrak{S}_n \text{ satisfying } \operatorname{Des}\left(\pi^{-1}\right) \subset D(\alpha)\right)\\
&= \sum_{\substack{\beta\in\operatorname{Comp}_n;\\ \beta \text{ coarsens } \alpha}} \sum_{\substack{\pi\in\mathfrak{S}_n;\\ \gamma(\pi^{-1})=\beta}} b_\pi,
\end{aligned}$$

where $\gamma\left(\pi^{-1}\right)$ denotes the composition τ of n satisfying $D(\tau) = \operatorname{Des}\left(\pi^{-1}\right)$).

Remark 6.3.6. Theorem 6.3.4 can be contrasted with the following fact: If $\mathbb{Q}$ is a subring of $\mathbf{k}$, then the shuffle algebra $\operatorname{Sh}(V)$ of **any** $\mathbf{k}$-module V (not necessarily free!) is isomorphic (as a $\mathbf{k}$-algebra) to the symmetric algebra $\operatorname{Sym}\left((\ker\epsilon)/(\ker\epsilon)^2\right)$ (by Theorem 1.7.29(e), applied to $A = \operatorname{Sh}(V)$). This fact is closely related to Theorem 6.3.4, but neither follows from it (since Theorem 6.3.4 only considers the case of free $\mathbf{k}$-modules V) nor yields it (since this fact does not provide explicit generators for the $\mathbf{k}$-module $(\ker\epsilon)/(\ker\epsilon)^2$ and thus for the $\mathbf{k}$-algebra $\operatorname{Sh}(V)$).

In our proof of Theorem 6.3.4 (but not only there), we will use part (a) of the following lemma[312], which makes proving that certain families indexed by Lyndon words generate certain $\mathbf{k}$-algebras more comfortable:

Lemma 6.3.7. *Let A be a commutative $\mathbf{k}$-algebra. Let $\mathfrak{A}$ be a totally ordered set. Let $\mathfrak{L}$ be the set of all Lyndon words over the alphabet $\mathfrak{A}$. Let b_w be an element of A for every $w \in \mathfrak{L}$. For every word $u \in \mathfrak{A}^*$, define an element $\mathbf{b}_u$ of A by $\mathbf{b}_u = b_{a_1} b_{a_2} \cdots b_{a_p}$, where $(a_1, a_2, \ldots, a_p)$ is the CFL factorization of u.*

(a) *The family $(b_w)_{w\in\mathfrak{L}}$ is an algebraically independent generating set of the $\mathbf{k}$-algebra A if and only if the family $(\mathbf{b}_u)_{u\in\mathfrak{A}^*}$ is a basis of the $\mathbf{k}$-module A.*

(b) *The family $(b_w)_{w\in\mathfrak{L}}$ generates the $\mathbf{k}$-algebra A if and only if the family $(\mathbf{b}_u)_{u\in\mathfrak{A}^*}$ spans the $\mathbf{k}$-module A.*

(c) *Assume that the $\mathbf{k}$-algebra A is graded. Let $\operatorname{wt} : \mathfrak{A} \to \{1, 2, 3, \ldots\}$ be any map such that for every $N \in \{1, 2, 3, \ldots\}$, the set $\operatorname{wt}^{-1}(N)$ is finite.*

For every word $w \in \mathfrak{A}^*$, define an element $\operatorname{Wt}(w) \in \mathbb{N}$ by $\operatorname{Wt}(w) = \operatorname{wt}(w_1) + \operatorname{wt}(w_2) + \cdots + \operatorname{wt}(w_k)$, where k is the length of w.

Assume that for every $w \in \mathfrak{L}$, the element b_w of A is homogeneous of degree $\operatorname{Wt}(w)$.

Assume further that the $\mathbf{k}$-module A has a basis $(g_u)_{u\in\mathfrak{A}^*}$ having the property that for every $u \in \mathfrak{A}^*$, the element g_u of A is homogeneous of degree $\operatorname{Wt}(u)$.

Assume also that the family $(b_w)_{w\in\mathfrak{L}}$ generates the $\mathbf{k}$-algebra A.

Then, this family $(b_w)_{w\in\mathfrak{L}}$ is an algebraically independent generating set of the $\mathbf{k}$-algebra A.

Exercise 6.3.8. Prove Lemma 6.3.7.

[**Hint:** For (a) and (b), notice that the $\mathbf{b}_u$ are the "monomials" in the b_w. For (c), use Exercise 2.5.18(b) in every homogeneous component of A.]

The main workhorse of our proof of Theorem 6.3.4 will be the following consequence of Theorem 6.2.2(c):

Proposition 6.3.9. *Let V be a free $\mathbf{k}$-module with a basis $(b_a)_{a\in\mathfrak{A}}$, where $\mathfrak{A}$ is a totally ordered set.*

For every word $w \in \mathfrak{A}^*$ over the alphabet $\mathfrak{A}$, let us define an element b_w of $\operatorname{Sh}(V)$ by $b_w = b_{w_1} b_{w_2} \cdots b_{w_\ell}$, where ℓ is the length of w. (The multiplication used here is that of $T(V)$, not that of $\operatorname{Sh}(V)$; the latter is denoted by $\sqcup\!\sqcup$.)

For every word $u \in \mathfrak{A}^*$, define an element $\mathbf{b}_u$ by $\mathbf{b}_u = b_{a_1} \sqcup\!\sqcup\, b_{a_2} \sqcup\!\sqcup \cdots \sqcup\!\sqcup\, b_{a_p}$, where $(a_1, a_2, \ldots, a_p)$ is the CFL factorization of u.

If $\ell \in \mathbb{N}$ and if $x \in \mathfrak{A}^\ell$ is a word, then there is a family $(\eta_{x,y})_{y\in\mathfrak{A}^\ell} \in \mathbb{N}^{\mathfrak{A}^\ell}$ of elements of $\mathbb{N}$ satisfying

$$\mathbf{b}_x = \sum_{\substack{y\in\mathfrak{A}^\ell;\\ y\le x}} \eta_{x,y} b_y$$

and $\eta_{x,x} \neq 0$ (in $\mathbb{N}$).

Before we prove this, let us show a very simple lemma:

Lemma 6.3.10. *Let $\mathfrak{A}$ be a totally ordered set. Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$. Let $\sigma \in \operatorname{Sh}_{n,m}$.*

(a) If u, v and v' are three words satisfying $\ell(u) = n$, $\ell(v) = m$, $\ell(v') = m$ and $v' < v$, then $u \underset{\sigma}{\sqcup\!\sqcup} v' < u \underset{\sigma}{\sqcup\!\sqcup} v$.

(b) If u, u' and v are three words satisfying $\ell(u) = n$, $\ell(u') = n$, $\ell(v) = m$ and $u' < u$, then $u' \underset{\sigma}{\sqcup\!\sqcup} v < u \underset{\sigma}{\sqcup\!\sqcup} v$.

(c) If u, v and v' are three words satisfying $\ell(u) = n$, $\ell(v) = m$, $\ell(v') = m$ and $v' \le v$, then $u \underset{\sigma}{\sqcup\!\sqcup} v' \le u \underset{\sigma}{\sqcup\!\sqcup} v$.

Exercise 6.3.11. Prove Lemma 6.3.10.

[312] And in a later proof, we will also use its part (c) (which is tailored for application to QSym).

Exercise 6.3.12. Prove Proposition 6.3.9.

[**Hint:** Proceed by induction over ℓ. In the induction step, apply Theorem 6.2.2(c)[313] to $u = a_1$ and $v = a_2 a_3 \cdots a_p$, where $(a_1, a_2, \ldots, a_p)$ is the CFL factorization of x. Use Lemma 6.3.10 to get rid of smaller terms.]

Exercise 6.3.13. Prove Theorem 6.3.4.

[**Hint:** According to Lemma 6.3.7(a), it suffices to show that the family $(\mathbf{b}_u)_{u \in \mathfrak{A}^*}$ defined in Proposition 6.3.9 is a basis of the $\mathbf{k}$-module $\operatorname{Sh}(V)$. When $\mathfrak{A}$ is finite, the latter can be proven by triangularity using Proposition 6.3.9. Reduce the general case to that of finite $\mathfrak{A}$.]

6.4. **Polynomial freeness of** QSym**: statement and easy parts.**

Definition 6.4.1. For the rest of Section 6.4 and for Section 6.5, we introduce the following notations: We let $\mathfrak{A}$ be the totally ordered set $\{1, 2, 3, \ldots\}$ with its natural order (that is, $1 < 2 < 3 < \cdots$.) Thus, the words over $\mathfrak{A}$ are precisely the compositions. That is, $\mathfrak{A}^* = \operatorname{Comp}$. We let $\mathfrak{L}$ denote the set of all Lyndon words over $\mathfrak{A}$. These Lyndon words are also called *Lyndon compositions*.

A natural question is how many Lyndon compositions of a given size exist. While we will not use the answer, we nevertheless record it:

Exercise 6.4.2. Show that the number of Lyndon compositions of size n equals

$$\frac{1}{n} \sum_{d \mid n} \mu(d) \left(2^{n/d} - 1\right) = \frac{1}{n} \sum_{d \mid n} \mu(d) \, 2^{n/d} - \delta_{n,1}$$

for every positive integer n (where "$\sum_{d \mid n}$" means a sum over all positive divisors of n, and where μ is the number-theoretic Möbius function).

[**Hint:** One solution is similar to the solution of Exercise 6.1.29 using CFL factorization. Another proceeds by defining a bijection between Lyndon compositions and Lyndon words over a two-letter alphabet $\{\mathbf{0}, \mathbf{1}\}$ (with $\mathbf{0} < \mathbf{1}$) which are $\neq \mathbf{1}$. [314]]

Let us now state Hazewinkel's result ([89, Theorem 8.1], [93, §6.7]) which is the main goal of Chapter 6:

Theorem 6.4.3. *The* $\mathbf{k}$*-algebra* QSym *is a polynomial algebra. It is isomorphic, as a graded* $\mathbf{k}$*-algebra, to the* $\mathbf{k}$*-algebra* $\mathbf{k}[x_w \mid w \in \mathfrak{L}]$*. Here, the grading on* $\mathbf{k}[x_w \mid w \in \mathfrak{L}]$ *is defined by setting* $\deg(x_w) = \sum_{i=1}^{\ell(w)} w_i$ *for every* $w \in \mathfrak{L}$*.*

We shall prove Theorem 6.4.3 in the next section (Section 6.5). But the particular case of Theorem 6.4.3 when $\mathbb{Q}$ is a subring of $\mathbf{k}$ can be proven more easily; we state it as a proposition:

Proposition 6.4.4. *Assume that* $\mathbb{Q}$ *is a subring of* $\mathbf{k}$*. Then, Theorem 6.4.3 holds.*

We will give two proofs of Proposition 6.4.4 in this Section 6.4; a third proof of Proposition 6.4.4 will immediately result from the proof of Theorem 6.4.3 in Section 6.5. (There **is** virtue in giving three different proofs, as they all construct different isomorphisms $\mathbf{k}[x_w \mid w \in \mathfrak{L}] \to \operatorname{QSym}$.)

Our first proof – originating in Malvenuto's [145, Corollaire 4.20] – can be given right away; it relies on Exercise 5.4.12:

First proof of Proposition 6.4.4. Let V be the free $\mathbf{k}$-module with basis $(\mathfrak{b}_n)_{n \in \{1,2,3,\ldots\}}$. Endow the $\mathbf{k}$-module V with a grading by assigning to each basis vector $\mathfrak{b}_n$ the degree n. Exercise 5.4.12(k) shows that QSym is isomorphic to the shuffle algebra $\operatorname{Sh}(V)$ (defined as in Proposition 1.6.7) as Hopf algebras. By being a

[313] Or Theorem 6.2.2(e), if you prefer.

[314] This bijection is obtained by restricting the bijection

$$\operatorname{Comp} \to \{w \in \{\mathbf{0}, \mathbf{1}\}^* \mid w \text{ does not start with } \mathbf{1}\},$$
$$(\alpha_1, \alpha_2, \ldots, \alpha_\ell) \mapsto \mathbf{0}\mathbf{1}^{\alpha_1 - 1} \mathbf{0}\mathbf{1}^{\alpha_2 - 1} \cdots \mathbf{0}\mathbf{1}^{\alpha_\ell - 1}$$

(where $\mathbf{01}^k$ is to be read as $\mathbf{0}\left(\mathbf{1}^k\right)$, not as $(\mathbf{01})^k$) to the set of Lyndon compositions. The idea behind this bijection is well-known in the Grothendieck-Teichmüller community: see, e.g., [94, §3.1] (and see [77, Note 5.16] for a different appearance of this idea).

bit more careful, we can obtain the slightly stronger result that QSym is isomorphic to the shuffle algebra $\operatorname{Sh}(V)$ as **graded** Hopf algebras[315]. In particular, $\operatorname{QSym} \cong \operatorname{Sh}(V)$ as graded $\mathbf{k}$-algebras.

Theorem 6.3.4 (applied to $b_a = \mathfrak{b}_a$) yields that the shuffle algebra $\operatorname{Sh}(V)$ is a polynomial $\mathbf{k}$-algebra, and that an algebraically independent generating set of $\operatorname{Sh}(V)$ can be constructed as follows:

For every word $w \in \mathfrak{A}^*$ over the alphabet $\mathfrak{A}$, let us define an element $\mathfrak{b}_w$ of $\operatorname{Sh}(V)$ by $\mathfrak{b}_w = \mathfrak{b}_{w_1}\mathfrak{b}_{w_2}\cdots\mathfrak{b}_{w_\ell}$, where ℓ is the length of w. (The multiplication used here is that of $T(V)$, not that of $\operatorname{Sh}(V)$; the latter is denoted by $\sqcup\!\sqcup$.) Then, $(\mathfrak{b}_w)_{w\in\mathfrak{L}}$ is an algebraically independent generating set of the $\mathbf{k}$-algebra $\operatorname{Sh}(V)$.

For every $w \in \mathfrak{A}^*$, we have $\mathfrak{b}_w = \mathfrak{b}_{w_1}\mathfrak{b}_{w_2}\cdots\mathfrak{b}_{w_{\ell(w)}}$ (by the definition of $\mathfrak{b}_w$). For every $w \in \mathfrak{A}^*$, the element $\mathfrak{b}_w = \mathfrak{b}_{w_1}\mathfrak{b}_{w_2}\cdots\mathfrak{b}_{w_{\ell(w)}}$ of $\operatorname{Sh}(V)$ is homogeneous of degree $\sum_{i=1}^{\ell(w)} \underbrace{\deg(\mathfrak{b}_{w_i})}_{=w_i} = \sum_{i=1}^{\ell(w)} w_i$.

Now, define a grading on the $\mathbf{k}$-algebra $\mathbf{k}[x_w \mid w \in \mathfrak{L}]$ by setting $\deg(x_w) = \sum_{i=1}^{\ell(w)} w_i$ for every $w \in \mathfrak{L}$. By the universal property of the polynomial algebra $\mathbf{k}[x_w \mid w \in \mathfrak{L}]$, we can define a $\mathbf{k}$-algebra homomorphism $\Phi : \mathbf{k}[x_w \mid w \in \mathfrak{L}] \to \operatorname{Sh}(V)$ by setting

$$\Phi(x_w) = \mathfrak{b}_w \qquad \text{for every } w \in \mathfrak{L}.$$

This homomorphism Φ is a $\mathbf{k}$-algebra isomorphism (since $(\mathfrak{b}_w)_{w\in\mathfrak{L}}$ is an algebraically independent generating set of the $\mathbf{k}$-algebra $\operatorname{Sh}(V)$) and is graded (because for every $w \in \mathfrak{L}$, the element $\mathfrak{b}_w$ of $\operatorname{Sh}(V)$ is homogeneous of degree $\sum_{i=1}^{\ell(w)} w_i = \deg(x_w)$). Thus, Φ is an isomorphism of graded $\mathbf{k}$-algebras. Hence, $\operatorname{Sh}(V) \cong \mathbf{k}[x_w \mid w \in \mathfrak{L}]$ as graded $\mathbf{k}$-algebras. Altogether, $\operatorname{QSym} \cong \operatorname{Sh}(V) \cong \mathbf{k}[x_w \mid w \in \mathfrak{L}]$ as graded $\mathbf{k}$-algebras. Thus, QSym is a polynomial algebra. This proves Theorem 6.4.3 under the assumption that $\mathbb{Q}$ be a subring of $\mathbf{k}$. In other words, this proves Proposition 6.4.4. □

Our second proof of Proposition 6.4.4 comes from Hazewinkel/Gubareni/Kirichenko [93] (where Proposition 6.4.4 appears as [93, Theorem 6.5.13]). This proof will construct an explicit algebraically independent family generating the $\mathbf{k}$-algebra QSym. [316] The generating set will be very unsophisticated: it will be $(M_\alpha)_{\alpha\in\mathfrak{L}}$, where $\mathfrak{A}$ and $\mathfrak{L}$ are as in Theorem 6.4.3. Here, we are using the fact that words over the alphabet $\{1, 2, 3, \ldots\}$ are the same thing as compositions, so, in particular, a monomial quasisymmetric function M_α is defined for every such word α.

It takes a bit of work to show that this family indeed fits the bill. We begin with a corollary of Proposition 5.1.3 that is essentially obtained by throwing away all non-bijective maps f:

Proposition 6.4.5. *Let $\alpha \in \mathfrak{A}^*$ and $\beta \in \mathfrak{A}^*$. Then,*

$$M_\alpha M_\beta = \sum_{\gamma \in \alpha \sqcup\!\sqcup \beta} M_\gamma + (\text{a sum of terms of the form } M_\delta \text{ with } \delta \in \mathfrak{A}^* \text{ satisfying } \ell(\delta) < \ell(\alpha) + \ell(\beta)).$$

[317]

Exercise 6.4.6. Prove Proposition 6.4.5.

[**Hint:** Recall what was said about the $p = \ell + m$ case in Example 5.1.4.]

Corollary 6.4.7. *Let $\alpha \in \mathfrak{A}^*$ and $\beta \in \mathfrak{A}^*$. Then, $M_\alpha M_\beta$ is a sum of terms of the form M_δ with $\delta \in \mathfrak{A}^*$ satisfying $\ell(\delta) \leq \ell(\alpha) + \ell(\beta)$.*

Exercise 6.4.8. Prove Corollary 6.4.7.

We now define a partial order on the compositions of a given nonnegative integer:

[315] *Proof.* In the solution of Exercise 5.4.12(k), we have shown that $\operatorname{QSym} \cong T(V)^o$ as graded Hopf algebras. But Remark 1.6.9(b) shows that the Hopf algebra $T(V)^o$ is naturally isomorphic to the shuffle algebra $\operatorname{Sh}(V^o)$ as Hopf algebras; it is easy to see that the natural isomorphism $T(V)^o \to \operatorname{Sh}(V^o)$ is graded (because it is the direct sum of the isomorphisms $(V^{\otimes n})^o \to (V^o)^{\otimes n}$ over all $n \in \mathbb{N}$, and each of these isomorphisms is graded). Hence, $T(V)^o \cong \operatorname{Sh}(V^o)$ as graded Hopf algebras. But $V^o \cong V$ as graded $\mathbf{k}$-modules (since V is of finite type), and thus $\operatorname{Sh}(V^o) \cong \operatorname{Sh}(V)$ as graded Hopf algebras. Altogether, we obtain $\operatorname{QSym} \cong T(V)^o \cong \operatorname{Sh}(V^o) \cong \operatorname{Sh}(V)$ as graded Hopf algebras, qed.

[316] We could, of course, obtain such a family from our above proof as well (this is done by Malvenuto in [145, Corollaire 4.20]), but it won't be a very simple one.

[317] The sum $\sum_{\gamma\in\alpha\sqcup\!\sqcup\beta} M_\gamma$ ranges over the **multiset** $\alpha \sqcup\!\sqcup \beta$; if an element appears several times in $\alpha \sqcup\!\sqcup \beta$, then it has accordingly many addends corresponding to it.

Definition 6.4.9. Let $n \in \mathbb{N}$. We define a binary relation $\underset{\text{wll}}{\leq}$ on the set Comp_n as follows: For two compositions α and β in Comp_n, we set $\alpha \underset{\text{wll}}{\leq} \beta$ if and only if

$$\text{either } \ell(\alpha) < \ell(\beta) \text{ or } (\ell(\alpha) = \ell(\beta) \text{ and } \alpha \leq \beta \text{ in lexicographic order}).$$

This binary relation $\underset{\text{wll}}{\leq}$ is the smaller-or-equal relation of a total order on Comp_n; we refer to said total order as the *wll-order* on Comp_n, and we denote by $\underset{\text{wll}}{<}$ the smaller relation of this total order.

Notice that if α and β are two compositions satisfying $\ell(\alpha) = \ell(\beta)$, then $\alpha \leq \beta$ in lexicographic order if and only if $\alpha \leq \beta$ with respect to the relation $\leq$ defined in Definition 6.1.1.

A remark about the name "wll-order" is in order. We have taken this notation from [89, Definition 6.7.14], where it is used for an extension of this order to the whole set Comp. We will never use this extension, as we will only ever compare two compositions of the same integer.[318]

We now state a fact which is similar (and plays a similar role) to Proposition 6.3.9:

Proposition 6.4.10. *For every composition $u \in \operatorname{Comp} = \mathfrak{A}^*$, define an element $\mathbf{M}_u \in \operatorname{QSym}$ by $\mathbf{M}_u = M_{a_1} M_{a_2} \cdots M_{a_p}$, where $(a_1, a_2, \ldots, a_p)$ is the CFL factorization of the word u.*

If $n \in \mathbb{N}$ and if $x \in \operatorname{Comp}_n$, then there is a family $(\eta_{x,y})_{y \in \operatorname{Comp}_n} \in \mathbb{N}^{\operatorname{Comp}_n}$ of elements of $\mathbb{N}$ satisfying

$$\mathbf{M}_x = \sum_{\substack{y \in \operatorname{Comp}_n; \\ y \underset{\text{wll}}{\leq} x}} \eta_{x,y} M_y$$

and $\eta_{x,x} \neq 0$ (in $\mathbb{N}$).

Before we prove it, let us show the following lemma:

Lemma 6.4.11. *Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$. Let $u \in \operatorname{Comp}_n$ and $v \in \operatorname{Comp}_m$. Let z be the lexicographically highest element of the multiset $u \sqcup\!\sqcup\, v$.*

(a) We have $z \in \operatorname{Comp}_{n+m}$.

(b) There exists a positive integer h such that

$$M_u M_v = h M_z + \left(\text{a sum of terms of the form } M_w \text{ with } w \in \operatorname{Comp}_{n+m} \text{ satisfying } w \underset{\text{wll}}{<} z\right).$$

(c) Let $v' \in \operatorname{Comp}_m$ be such that $v' \underset{\text{wll}}{<} v$. Then,

$$M_u M_{v'} = \left(\text{a sum of terms of the form } M_w \text{ with } w \in \operatorname{Comp}_{n+m} \text{ satisfying } w \underset{\text{wll}}{<} z\right).$$

Exercise 6.4.12. Prove Lemma 6.4.11.

[**Hint:** For (b), set h to be the multiplicity with which the word z appears in the multiset $u \sqcup\!\sqcup\, v$, then use Proposition 6.4.5 and notice that $M_u M_v$ is homogeneous of degree $n+m$. For (c), use (b) for v' instead of v and notice that Lemma 6.3.10(a) shows that the lexicographically highest element of the multiset $u \sqcup\!\sqcup\, v'$ is $\underset{\text{wll}}{<} z$.]

Exercise 6.4.13. Prove Proposition 6.4.10.

[**Hint:** Proceed by strong induction over n. In the induction step, let $(a_1, a_2, \ldots, a_p)$ be the CFL factorization of x, and set $u = a_1$ and $v = a_2 a_3 \cdots a_p$; then apply Proposition 6.4.10 to v instead of x, and multiply the resulting equality $\mathbf{M}_v = \sum_{\substack{y \in \operatorname{Comp}_{|v|}; \\ y \underset{\text{wll}}{\leq} v}} \eta_{v,y} M_y$ with M_u to obtain an expression for $M_u \mathbf{M}_v = \mathbf{M}_x$. Use Lemma 6.4.11 to show that this expression has the form $\sum_{\substack{y \in \operatorname{Comp}_n; \\ y \underset{\text{wll}}{\leq} x}} \eta_{x,y} M_y$ with $\eta_{x,x} \neq 0$; here it helps to remember that the lexicographically highest element of the multiset $u \sqcup\!\sqcup\, v$ is $uv = x$ (by Theorem 6.2.2(c)).]

[318] In [89, Definition 6.7.14], the name "wll-order" is introduced as an abbreviation for "**w**eight first, then **l**ength, then **l**exicographic" (in the sense that two compositions are first compared by their weights, then, if the weights are equal, by their lengths, and finally, if the lengths are also equal, by the lexicographic order). For us, the alternative explanation "**w**ord **l**ength, then **l**exicographic" serves just as well.

We are almost ready to give our second proof of Proposition 6.4.4; our last step is the following proposition:

Proposition 6.4.14. *Assume that $\mathbb{Q}$ is a subring of* $\mathbf{k}$. *Then,* $(M_w)_{w \in \mathfrak{L}}$ *is an algebraically independent generating set of the* $\mathbf{k}$*-algebra* QSym.

Exercise 6.4.15. Prove Proposition 6.4.14.

[**Hint:** Define $\mathbf{M}_u$ for every $u \in \operatorname{Comp}$ as in Proposition 6.4.10. Conclude from Proposition 6.4.10 that, for every $n \in \mathbb{N}$, the family $(\mathbf{M}_u)_{u \in \operatorname{Comp}_n}$ expands invertibly triangularly[319] (with respect to the total order $\underset{\mathrm{wll}}{\leq}$ on Comp_n) with respect to the basis $(M_u)_{u \in \operatorname{Comp}_n}$ of QSym_n. Conclude that this family $(\mathbf{M}_u)_{u \in \operatorname{Comp}_n}$ is a basis of QSym_n itself, and so the whole family $(\mathbf{M}_u)_{u \in \operatorname{Comp}}$ is a basis of QSym. Conclude using Lemma 6.3.7(a).]

Second proof of Proposition 6.4.4. Proposition 6.4.14 yields that $(M_w)_{w \in \mathfrak{L}}$ is an algebraically independent generating set of the $\mathbf{k}$-algebra QSym.

Define a grading on the $\mathbf{k}$-algebra $\mathbf{k}[x_w \mid w \in \mathfrak{L}]$ by setting $\deg(x_w) = \sum_{i=1}^{\ell(w)} w_i$ for every $w \in \mathfrak{L}$. By the universal property of the polynomial algebra $\mathbf{k}[x_w \mid w \in \mathfrak{L}]$, we can define a $\mathbf{k}$-algebra homomorphism $\Phi : \mathbf{k}[x_w \mid w \in \mathfrak{L}] \to \operatorname{QSym}$ by setting

$$\Phi(x_w) = M_w \qquad \text{for every } w \in \mathfrak{L}.$$

This homomorphism Φ is a $\mathbf{k}$-algebra isomorphism (since $(M_w)_{w \in \mathfrak{L}}$ is an algebraically independent generating set of the $\mathbf{k}$-algebra QSym) and is graded (because for every $w \in \mathfrak{L}$, the element M_w of QSym is homogeneous of degree $|w| = \sum_{i=1}^{\ell(w)} w_i = \deg(x_w)$). Thus, Φ is an isomorphism of graded $\mathbf{k}$-algebras. Hence, $\operatorname{QSym} \cong \mathbf{k}[x_w \mid w \in \mathfrak{L}]$ as graded $\mathbf{k}$-algebras. In particular, this shows that QSym is a polynomial algebra. This proves Theorem 6.4.3 under the assumption that $\mathbb{Q}$ be a subring of $\mathbf{k}$. Proposition 6.4.4 is thus proven again. □

6.5. **Polynomial freeness of** QSym**: the general case.**

We now will prepare for proving Theorem 6.4.3 without any assumptions on $\mathbf{k}$. In our proof, we follow [89] and [93, §6.7], but without using the language of plethysm and Frobenius maps. We start with the following definition:

Definition 6.5.1. Let α be a composition. Write α in the form $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_\ell)$ with $\ell = \ell(\alpha)$.

(a) Let $\operatorname{SIS}(\ell)$ denote the set of all strictly increasing ℓ-tuples $(i_1, i_2, \ldots, i_\ell)$ of positive integers.[320] For every ℓ-tuple $\mathbf{i} = (i_1, i_2, \ldots, i_\ell) \in \operatorname{SIS}(\ell)$, we denote the monomial $x_{i_1}^{\alpha_1} x_{i_2}^{\alpha_2} \cdots x_{i_\ell}^{\alpha_\ell}$ by $\mathbf{x}_{\mathbf{i}}^{\alpha}$. This $\mathbf{x}_{\mathbf{i}}^{\alpha}$ is a monomial of degree $\alpha_1 + \alpha_2 + \cdots + \alpha_\ell = |\alpha|$. Then,

$$M_\alpha = \sum_{\mathbf{i} \in \operatorname{SIS}(\ell)} \mathbf{x}_{\mathbf{i}}^{\alpha}. \tag{6.5.1}$$

[321]

[319] See Definition 11.1.16(b) for the meaning of this.

[320] "Strictly increasing" means that $i_1 < i_2 < \cdots < i_\ell$ here. Of course, the elements of $\operatorname{SIS}(\ell)$ are in 1-to-1 correspondence with ℓ-element subsets of $\{1, 2, 3, \ldots\}$.

[321] *Proof of (6.5.1):* By the definition of M_α, we have

$$M_\alpha = \underbrace{\sum_{i_1 < i_2 < \cdots < i_\ell \text{ in } \{1,2,3,\ldots\}}}_{= \sum_{(i_1, i_2, \ldots, i_\ell) \in \operatorname{SIS}(\ell)}} x_{i_1}^{\alpha_1} x_{i_2}^{\alpha_2} \cdots x_{i_\ell}^{\alpha_\ell} = \sum_{(i_1, i_2, \ldots, i_\ell) \in \operatorname{SIS}(\ell)} x_{i_1}^{\alpha_1} x_{i_2}^{\alpha_2} \cdots x_{i_\ell}^{\alpha_\ell} = \sum_{\mathbf{i} = (i_1, i_2, \ldots, i_\ell) \in \operatorname{SIS}(\ell)} \underbrace{x_{i_1}^{\alpha_1} x_{i_2}^{\alpha_2} \cdots x_{i_\ell}^{\alpha_\ell}}_{\substack{= \mathbf{x}_{\mathbf{i}}^{\alpha} \\ (\text{by the definition of } \mathbf{x}_{\mathbf{i}}^{\alpha})}}$$
$$= \sum_{\mathbf{i} = (i_1, i_2, \ldots, i_\ell) \in \operatorname{SIS}(\ell)} \mathbf{x}_{\mathbf{i}}^{\alpha} = \sum_{\mathbf{i} \in \operatorname{SIS}(\ell)} \mathbf{x}_{\mathbf{i}}^{\alpha},$$

qed.

(b) Consider the ring $\mathbf{k}[[\mathbf{x}]]$ endowed with the coefficientwise topology[322]. The family $(\mathbf{x}_{\mathbf{i}}^{\alpha})_{\mathbf{i}\in\mathrm{SIS}(\ell)}$ of elements of $\mathbf{k}[[\mathbf{x}]]$ is power-summable[323]. Hence, for every $f \in \Lambda$, there is a well-defined power series $f\left((\mathbf{x}_{\mathbf{i}}^{\alpha})_{\mathbf{i}\in\mathrm{SIS}(\ell)}\right) \in \mathbf{k}[[\mathbf{x}]]$ obtained by "evaluating" f at $(\mathbf{x}_{\mathbf{i}}^{\alpha})_{\mathbf{i}\in\mathrm{SIS}(\ell)}$ [324]. In particular, for every $s \in \mathbb{Z}$, we can evaluate the symmetric function $e_s \in \Lambda$ [325] at $(\mathbf{x}_{\mathbf{i}}^{\alpha})_{\mathbf{i}\in\mathrm{SIS}(\ell)}$. The resulting power series $e_s\left((\mathbf{x}_{\mathbf{i}}^{\alpha})_{\mathbf{i}\in\mathrm{SIS}(\ell)}\right) \in \mathbf{k}[[\mathbf{x}]]$ will be denoted $M_{\alpha}^{\langle s\rangle}$. Thus,

$$M_{\alpha}^{\langle s\rangle} = e_s\left((\mathbf{x}_{\mathbf{i}}^{\alpha})_{\mathbf{i}\in\mathrm{SIS}(\ell)}\right).$$

The power series $M_{\alpha}^{\langle s\rangle}$ are the power series $e_s(\alpha)$ in [93]. We will shortly (in Corollary 6.5.8(a)) see that $M_{\alpha}^{\langle s\rangle} \in \mathrm{QSym}$ (although this is also easy to prove by inspection). Here are some examples of $M_{\alpha}^{\langle s\rangle}$:

[322] This topology is defined as follows:

We endow the ring $\mathbf{k}$ with the discrete topology. Then, we can regard the $\mathbf{k}$-module $\mathbf{k}[[\mathbf{x}]]$ as a direct product of infinitely many copies of $\mathbf{k}$ (by identifying every power series in $\mathbf{k}[[\mathbf{x}]]$ with the family of its coefficients). Hence, the product topology is a well-defined topology on $\mathbf{k}[[\mathbf{x}]]$; this topology is denoted as the *coefficientwise topology*. A sequence $(a_n)_{n\in\mathbb{N}}$ of power series converges to a power series a with respect to this topology if and only if for every monomial $\mathfrak{m}$, all sufficiently high $n \in \mathbb{N}$ satisfy

$$(\text{the coefficient of } \mathfrak{m} \text{ in } a_n) = (\text{the coefficient of } \mathfrak{m} \text{ in } a).$$

Note that this is **not** the topology obtained by taking the completion of $\mathbf{k}[x_1, x_2, x_3, \ldots]$ with respect to the standard grading (in which all x_i have degree 1). (The latter completion is actually a smaller ring than $\mathbf{k}[[\mathbf{x}]]$.)

[323] Let us define what "power-summable" means for us:

A family $(n_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}} \in \mathbb{N}^{\mathbf{I}}$ (where $\mathbf{I}$ is some set) is said to be *finitely supported* if all but finitely many $\mathbf{i} \in \mathbf{I}$ satisfy $n_{\mathbf{i}} = 0$.

If $(n_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}} \in \mathbb{N}^{\mathbf{I}}$ is a finitely supported family, then $\sum_{\mathbf{i}\in\mathbf{I}} n_{\mathbf{i}}$ is a well-defined element of $\mathbb{N}$. If $N \in \mathbb{N}$, then a family $(n_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}} \in \mathbb{N}^{\mathbf{I}}$ will be called *$(\leq N)$-supported* if it is finitely supported and satisfies $\sum_{\mathbf{i}\in\mathbf{I}} n_{\mathbf{i}} \leq N$.

We say that a family $(s_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}} \in R^{\mathbf{I}}$ of elements of a topological commutative $\mathbf{k}$-algebra R is *power-summable* if it satisfies the following property: For every $N \in \mathbb{N}$, the sum

$$\sum_{\substack{(n_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}}\in\mathbb{N}^{\mathbf{I}};\\ (n_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}} \text{ is } (\leq N)\text{-supported}}} \alpha_{(n_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}}} \prod_{\mathbf{i}\in\mathbf{I}} s_{\mathbf{i}}^{n_{\mathbf{i}}}$$

converges in the topology on R for every choice of scalars $\alpha_{(n_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}}} \in \mathbf{k}$ corresponding to all $(\leq N)$-supported $(n_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}} \in \mathbb{N}^{\mathbf{I}}$. In our specific case, we consider $\mathbf{k}[[\mathbf{x}]]$ as a topological commutative $\mathbf{k}$-algebra, where the topology is the coefficientwise topology. The fact that the family $(\mathbf{x}_{\mathbf{i}}^{\alpha})_{\mathbf{i}\in\mathrm{SIS}(\ell)}$ is power-summable then can be proven as follows:

- If $\alpha \neq \varnothing$, then this fact follows from the (easily-verified) observation that every given monomial in the variables $x_1, x_2, x_3, \ldots$ can be written as a product of monomials of the form $\mathbf{x}_{\mathbf{i}}^{\alpha}$ (with $\mathbf{i} \in \mathrm{SIS}(\ell)$) in only finitely many ways.
- If $\alpha = \varnothing$, then this fact follows by noticing that $(\mathbf{x}_{\mathbf{i}}^{\alpha})_{\mathbf{i}\in\mathrm{SIS}(\ell)}$ is a finite family (indeed, $\mathrm{SIS}(\ell) = \mathrm{SIS}(0) = \{()\}$), and every finite family is power-summable.

[324] Here is how this power series $f\left((\mathbf{x}_{\mathbf{i}}^{\alpha})_{\mathbf{i}\in\mathrm{SIS}(\ell)}\right)$ is formally defined:

Let R be any topological commutative $\mathbf{k}$-algebra, and let $(s_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}} \in R^{\mathbf{I}}$ be any power-summable family of elements of R. Assume that the indexing set $\mathbf{I}$ is countably infinite, and fix a bijection $\mathrm{j} : \{1, 2, 3, \ldots\} \to \mathbf{I}$. Let $g \in R(\mathbf{x})$ be arbitrary. Then, we can substitute $s_{\mathrm{j}(1)}$, $s_{\mathrm{j}(2)}$, $s_{\mathrm{j}(3)}$, $\ldots$ for the variables x_1, x_2, x_3, $\ldots$ in g, thus obtaining an infinite sum which converges in R (in fact, its convergence follows from the fact that the family $(s_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}} \in R^{\mathbf{I}}$ is power-summable). The value of this sum will be denoted by $g\left((s_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}}\right)$. In general, this value depends on the choice of the bijection j, so the notation $g\left((s_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}}\right)$ is unambiguous only if this bijection j is chosen once and for all. However, when $g \in \Lambda$, one can easily see that the choice of j has no effect on $g\left((s_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}}\right)$.

We can still define $g\left((s_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}}\right)$ when the set $\mathbf{I}$ is finite instead of being countably infinite. In this case, we only need to modify our above definition as follows: Instead of fixing a bijection $\mathrm{j} : \{1, 2, 3, \ldots\} \to \mathbf{I}$, we now fix a bijection $\mathrm{j} : \{1, 2, \ldots, |\mathbf{I}|\} \to \mathbf{I}$, and instead of substituting $s_{\mathrm{j}(1)}$, $s_{\mathrm{j}(2)}$, $s_{\mathrm{j}(3)}$, $\ldots$ for the variables x_1, x_2, x_3, $\ldots$ in g, we now substitute $s_{\mathrm{j}(1)}$, $s_{\mathrm{j}(2)}$, $\ldots$, $s_{\mathrm{j}(|\mathbf{I}|)}$, 0, 0, 0, $\ldots$ for the variables x_1, x_2, x_3, $\ldots$ in g. Again, the same observations hold as before: $g\left((s_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}}\right)$ is independent on j if $g \in \Lambda$.

Hence, $g\left((s_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}}\right)$ is well-defined for every $g \in R(\mathbf{x})$, every countable (i.e., finite or countably infinite) set $\mathbf{I}$, every topological commutative $\mathbf{k}$-algebra R and every power-summable family $(s_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}} \in R^{\mathbf{I}}$ of elements of R, as long as a bijection j is chosen. In particular, we can apply this to $g = f$, $\mathbf{I} = \mathrm{SIS}(\ell)$, $R = \mathbf{k}[[\mathbf{x}]]$ and $(s_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}} = (\mathbf{x}_{\mathbf{i}}^{\alpha})_{\mathbf{i}\in\mathrm{SIS}(\ell)}$, choosing j to be the bijection which sends every positive integer k to the k-th smallest element of $\mathrm{SIS}(\ell)$ in the lexicographic order. (Of course, since $f \in \Lambda$, the choice of j is irrelevant.)

[325] Recall that $e_0 = 1$, and that $e_s = 0$ for $s < 0$.

Example 6.5.2. If α is a composition and ℓ denotes its length $\ell(\alpha)$, then

$$M_\alpha^{\langle 0\rangle} = \underbrace{e_0}_{=1}\left(\left(\mathbf{x}_\mathbf{i}^\alpha\right)_{\mathbf{i}\in\operatorname{SIS}(\ell)}\right) = 1\left(\left(\mathbf{x}_\mathbf{i}^\alpha\right)_{\mathbf{i}\in\operatorname{SIS}(\ell)}\right) = 1$$

and

$$M_\alpha^{\langle 1\rangle} = e_1\left(\left(\mathbf{x}_\mathbf{i}^\alpha\right)_{\mathbf{i}\in\operatorname{SIS}(\ell)}\right) = \sum_{\mathbf{i}\in\operatorname{SIS}(\ell)} \mathbf{x}_\mathbf{i}^\alpha = M_\alpha \qquad \text{(by (6.5.1))}$$

and[326]

$$M_\alpha^{\langle 2\rangle} = e_2\left(\left(\mathbf{x}_\mathbf{i}^\alpha\right)_{\mathbf{i}\in\operatorname{SIS}(\ell)}\right) = \sum_{\substack{\mathbf{i}\in\operatorname{SIS}(\ell),\ \mathbf{j}\in\operatorname{SIS}(\ell);\\ \mathbf{i}<\mathbf{j}}} \mathbf{x}_\mathbf{i}^\alpha \mathbf{x}_\mathbf{j}^\alpha$$

(where the notation "$\mathbf{i} < \mathbf{j}$" should be interpreted with respect to an arbitrary but fixed total order on the set $\operatorname{SIS}(\ell)$ – for example, the lexicographic order). Applying the last of these three equalities to $\alpha = (2, 1)$, we obtain

$$\begin{aligned}
M_{(2,1)}^{\langle 2\rangle} &= \sum_{\substack{\mathbf{i}\in\operatorname{SIS}(2),\ \mathbf{j}\in\operatorname{SIS}(2),\\ \mathbf{i}<\mathbf{j}}} \mathbf{x}_\mathbf{i}^{(2,1)}\mathbf{x}_\mathbf{j}^{(2,1)} = \sum_{\substack{(i_1,i_2)\in\operatorname{SIS}(2),\ (j_1,j_2)\in\operatorname{SIS}(2);\\ (i_1,i_2)<(j_1,j_2)}} \underbrace{\mathbf{x}_{(i_1,i_2)}^{(2,1)}}_{=x_{i_1}^2 x_{i_2}^1} \underbrace{\mathbf{x}_{(j_1,j_2)}^{(2,1)}}_{=x_{j_1}^2 x_{j_2}^1} \\
&= \sum_{\substack{(i_1,i_2)\in\operatorname{SIS}(2),\ (j_1,j_2)\in\operatorname{SIS}(2);\\ (i_1,i_2)<(j_1,j_2)}} x_{i_1}^2 x_{i_2}^1 x_{j_1}^2 x_{j_2}^1 \\
&= \underbrace{\sum_{\substack{i_1<i_2;\ j_1<j_2;\\ i_1<j_1}} x_{i_1}^2 x_{i_2}^1 x_{j_1}^2 x_{j_2}^1}_{=M_{(2,1,2,1)}+M_{(2,3,1)}+2M_{(2,2,1,1)}+M_{(2,2,2)}} + \underbrace{\sum_{\substack{i_1<i_2;\ j_1<j_2;\\ i_1=j_1;\ i_2<j_2}} x_{i_1}^2 x_{i_2}^1 x_{j_1}^2 x_{j_2}^1}_{=M_{(4,1,1)}} \\
&\qquad \left(\text{here, we have WLOG assumed that the order on } \operatorname{SIS}(2) \text{ is lexicographic}\right) \\
&= M_{(2,1,2,1)} + M_{(2,3,1)} + 2M_{(2,2,1,1)} + M_{(2,2,2)} + M_{(4,1,1)}.
\end{aligned}$$

Of course, every negative integer s satisfies $M_\alpha^{\langle s\rangle} = \underbrace{e_s}_{=0}\left(\left(\mathbf{x}_\mathbf{i}^\alpha\right)_{\mathbf{i}\in\operatorname{SIS}(\ell)}\right) = 0$.

There is a determinantal formula for the $s!M_\alpha^{\langle s\rangle}$ (and thus also for $M_\alpha^{\langle s\rangle}$ when $s!$ is invertible in $\mathbf{k}$), but in order to state it, we need to introduce one more notation:

Definition 6.5.3. Let $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_\ell)$ be a composition, and let k be a positive integer. Then, $\alpha\{k\}$ will denote the composition $(k\alpha_1, k\alpha_2, \ldots, k\alpha_\ell)$. Clearly, $\ell(\alpha\{k\}) = \ell(\alpha)$ and $|\alpha\{k\}| = k|\alpha|$.

Exercise 6.5.4. Let α be a composition. Write the composition α in the form $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_\ell)$ with $\ell = \ell(\alpha)$.

(a) Show that the s-th power-sum symmetric function $p_s \in \Lambda$ satisfies

$$p_s\left(\left(\mathbf{x}_\mathbf{i}^\alpha\right)_{\mathbf{i}\in\operatorname{SIS}(\ell)}\right) = M_{\alpha\{s\}}$$

for every positive integer s.

(b) Let us fix a total order on the set $\operatorname{SIS}(\ell)$ (for example, the lexicographic order). Show that the s-th elementary symmetric function $e_s \in \Lambda$ satisfies

$$M_\alpha^{\langle s\rangle} = e_s\left(\left(\mathbf{x}_\mathbf{i}^\alpha\right)_{\mathbf{i}\in\operatorname{SIS}(\ell)}\right) = \sum_{\substack{(\mathbf{i}_1,\mathbf{i}_2,\ldots,\mathbf{i}_s)\in(\operatorname{SIS}(\ell))^s;\\ \mathbf{i}_1<\mathbf{i}_2<\cdots<\mathbf{i}_s}} \mathbf{x}_{\mathbf{i}_1}^\alpha \mathbf{x}_{\mathbf{i}_2}^\alpha \cdots \mathbf{x}_{\mathbf{i}_s}^\alpha$$

for every $s \in \mathbb{N}$.

(c) Let $s \in \mathbb{N}$, and let n be a positive integer. Let $e_s^{\langle n\rangle}$ be the symmetric function $\sum_{i_1<i_2<\cdots<i_s} x_{i_1}^n x_{i_2}^n \cdots x_{i_s}^n \in \Lambda$. Then, show that

$$M_{\alpha\{n\}}^{\langle s\rangle} = e_s^{\langle n\rangle}\left(\left(\mathbf{x}_\mathbf{i}^\alpha\right)_{\mathbf{i}\in\operatorname{SIS}(\ell)}\right).$$

[326] This is not completely obvious, but easy to check (see Exercise 6.5.4(b)).

(d) Let $s \in \mathbb{N}$, and let n be a positive integer. Prove that there exists a polynomial $P \in \mathbf{k}[z_1, z_2, z_3, \ldots]$ such that $M_{\alpha\{n\}}^{\langle s\rangle} = P\left(M_\alpha^{\langle 1\rangle}, M_\alpha^{\langle 2\rangle}, M_\alpha^{\langle 3\rangle}, \ldots\right)$.

[**Hint:** For (a), (b) and (c), apply the definition of $f\left((\mathbf{x}_\mathbf{i}^\alpha)_{\mathbf{i}\in\operatorname{SIS}(\ell)}\right)$ with f a symmetric function[327]. For (d), recall that Λ is generated by $e_1, e_2, e_3, \ldots$.]

Exercise 6.5.5. Let $s \in \mathbb{N}$. Show that the composition (1) satisfies $M_{(1)}^{\langle s\rangle} = e_s$.

Proposition 6.5.6. *Let $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_\ell)$ be a composition.*

(a) Let $n \in \mathbb{N}$. Define a matrix $A_n^{\langle\alpha\rangle} = \left(a_{i,j}^{\langle\alpha\rangle}\right)_{i,j=1,2,\ldots,n}$ by

$$a_{i,j}^{\langle\alpha\rangle} = \begin{cases} M_{\alpha\{i-j+1\}}, & \text{if } i \geq j; \\ i, & \text{if } i = j-1; \\ 0, & \text{if } i < j-1 \end{cases} \qquad \text{for all } (i,j) \in \{1,2,\ldots,n\}^2 .$$

This matrix $A_n^{\langle\alpha\rangle}$ looks as follows:

$$A_n^{\langle\alpha\rangle} = \begin{pmatrix} M_{\alpha\{1\}} & 1 & 0 & \cdots & 0 & 0 \\ M_{\alpha\{2\}} & M_{\alpha\{1\}} & 2 & \cdots & 0 & 0 \\ M_{\alpha\{3\}} & M_{\alpha\{2\}} & M_{\alpha\{1\}} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ M_{\alpha\{n-1\}} & M_{\alpha\{n-2\}} & M_{\alpha\{n-3\}} & \cdots & M_{\alpha\{1\}} & n-1 \\ M_{\alpha\{n\}} & M_{\alpha\{n-1\}} & M_{\alpha\{n-2\}} & \cdots & M_{\alpha\{2\}} & M_{\alpha\{1\}} \end{pmatrix} .$$

Then, $\det\left(A_n^{\langle\alpha\rangle}\right) = n! M_\alpha^{\langle n\rangle}$.

(b) Let n be a positive integer. Define a matrix $B_n^{\langle\alpha\rangle} = \left(b_{i,j}^{\langle\alpha\rangle}\right)_{i,j=1,2,\ldots,n}$ by

$$b_{i,j}^{\langle\alpha\rangle} = \begin{cases} i M_\alpha^{\langle i\rangle}, & \text{if } j = 1; \\ M_\alpha^{\langle i-j+1\rangle}, & \text{if } j > 1 \end{cases} \qquad \text{for all } (i,j) \in \{1,2,\ldots,n\}^2 .$$

[327] There are two subtleties that need to be addressed:

- the fact that the definition of $f\left((\mathbf{x}_\mathbf{i}^\alpha)_{\mathbf{i}\in\operatorname{SIS}(\ell)}\right)$ distinguishes between two cases depending on whether or not $\operatorname{SIS}(\ell)$ is finite;
- the fact that the total order on the set $\{1,2,3,\ldots\}$ (which appears in the summation subscript in the equality $e_s = \sum_{\substack{(i_1,i_2,\ldots,i_s)\in\{1,2,3,\ldots\}^s; \\ i_1<i_2<\cdots<i_s}} x_{i_1}x_{i_2}\cdots x_{i_s}$) has nothing to do with the total order on the set $\operatorname{SIS}(\ell)$ (which appears in the summation subscript in $\sum_{\substack{(\mathbf{i}_1,\mathbf{i}_2,\ldots,\mathbf{i}_s)\in(\operatorname{SIS}(\ell))^s; \\ \mathbf{i}_1<\mathbf{i}_2<\cdots<\mathbf{i}_s}} \mathbf{x}_{\mathbf{i}_1}^\alpha \mathbf{x}_{\mathbf{i}_2}^\alpha \cdots \mathbf{x}_{\mathbf{i}_s}^\alpha$). For instance, the former total order is well-founded, whereas the latter may and may not be. So there is (generally) no bijection between $\{1,2,3,\ldots\}$ and $\operatorname{SIS}(\ell)$ preserving these orders (even if $\operatorname{SIS}(\ell)$ is infinite). Fortunately, this does not matter much, because the total order is only being used to ensure that every product of s distinct elements appears exactly once in the sum.

The matrix $B_n^{\langle \alpha \rangle}$ looks as follows:

$$B_n^{\langle \alpha \rangle} = \begin{pmatrix} M_\alpha^{\langle 1 \rangle} & M_\alpha^{\langle 0 \rangle} & M_\alpha^{\langle -1 \rangle} & \cdots & M_\alpha^{\langle -n+3 \rangle} & M_\alpha^{\langle -n+2 \rangle} \\ 2M_\alpha^{\langle 2 \rangle} & M_\alpha^{\langle 1 \rangle} & M_\alpha^{\langle 0 \rangle} & \cdots & M_\alpha^{\langle -n+4 \rangle} & M_\alpha^{\langle -n+3 \rangle} \\ 3M_\alpha^{\langle 3 \rangle} & M_\alpha^{\langle 2 \rangle} & M_\alpha^{\langle 1 \rangle} & \cdots & M_\alpha^{\langle -n+5 \rangle} & M_\alpha^{\langle -n+4 \rangle} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ (n-1)\, M_\alpha^{\langle n-1 \rangle} & M_\alpha^{\langle n-2 \rangle} & M_\alpha^{\langle n-3 \rangle} & \cdots & M_\alpha^{\langle 1 \rangle} & M_\alpha^{\langle 0 \rangle} \\ nM_\alpha^{\langle n \rangle} & M_\alpha^{\langle n-1 \rangle} & M_\alpha^{\langle n-2 \rangle} & \cdots & M_\alpha^{\langle 2 \rangle} & M_\alpha^{\langle 1 \rangle} \end{pmatrix}$$

$$= \begin{pmatrix} M_\alpha^{\langle 1 \rangle} & 1 & 0 & \cdots & 0 & 0 \\ 2M_\alpha^{\langle 2 \rangle} & M_\alpha^{\langle 1 \rangle} & 1 & \cdots & 0 & 0 \\ 3M_\alpha^{\langle 3 \rangle} & M_\alpha^{\langle 2 \rangle} & M_\alpha^{\langle 1 \rangle} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ (n-1)\, M_\alpha^{\langle n-1 \rangle} & M_\alpha^{\langle n-2 \rangle} & M_\alpha^{\langle n-3 \rangle} & \cdots & M_\alpha^{\langle 1 \rangle} & 1 \\ nM_\alpha^{\langle n \rangle} & M_\alpha^{\langle n-1 \rangle} & M_\alpha^{\langle n-2 \rangle} & \cdots & M_\alpha^{\langle 2 \rangle} & M_\alpha^{\langle 1 \rangle} \end{pmatrix}.$$

Then, $\det\left(B_n^{\langle \alpha \rangle}\right) = M_{\alpha\{n\}}$.

Exercise 6.5.7. Prove Proposition 6.5.6.

[**Hint:** Substitute $\left(\mathbf{x}_{\mathbf{i}}^{\alpha}\right)_{\mathbf{i}\in\operatorname{SIS}(\ell)}$ for the variable set in Exercise 2.9.13, and recall Exercise 6.5.4(a).]

Corollary 6.5.8. *Let α be a composition. Let $s \in \mathbb{Z}$.*

(a) *We have $M_\alpha^{\langle s \rangle} \in \operatorname{QSym}$.*

(b) *We have $M_\alpha^{\langle s \rangle} \in \operatorname{QSym}_{s|\alpha|}$.*

Exercise 6.5.9. Prove Corollary 6.5.8.

We make one further definition:

Definition 6.5.10. Let α be a nonempty composition. Then, we denote by $\gcd \alpha$ the greatest common divisor of the parts of α. (For instance, $\gcd(8, 6, 4) = 2$.) We also define $\operatorname{red} \alpha$ to be the composition $\left(\dfrac{\alpha_1}{\gcd \alpha}, \dfrac{\alpha_2}{\gcd \alpha}, \ldots, \dfrac{\alpha_\ell}{\gcd \alpha}\right)$, where α is written in the form $(\alpha_1, \alpha_2, \ldots, \alpha_\ell)$.

We say that a nonempty composition α is *reduced* if $\gcd \alpha = 1$.

We define $\mathfrak{RL}$ to be the set of all reduced Lyndon compositions. In other words, $\mathfrak{RL} = \{w \in \mathfrak{L} \mid w \text{ is reduced}\}$ (since $\mathfrak{L}$ is the set of all Lyndon compositions).

Hazewinkel, in [93, proof of Thm. 6.7.5], denotes $\mathfrak{RL}$ by $eLYN$, calling reduced Lyndon compositions "elementary Lyndon words".

Remark 6.5.11. Let α be a nonempty composition.

(a) We have $\alpha = (\operatorname{red} \alpha)\{\gcd \alpha\}$.

(b) The composition α is Lyndon if and only if the composition $\operatorname{red} \alpha$ is Lyndon.

(c) The composition $\operatorname{red} \alpha$ is reduced.

(d) If α is reduced, then $\operatorname{red} \alpha = \alpha$.

(e) If $s \in \{1, 2, 3, \ldots\}$, then the composition $\alpha\{s\}$ is nonempty and satisfies $\operatorname{red}(\alpha\{s\}) = \operatorname{red} \alpha$ and $\gcd(\alpha\{s\}) = s \gcd \alpha$.

(f) We have $(\gcd \alpha)\, |\operatorname{red} \alpha| = |\alpha|$.

Exercise 6.5.12. Prove Remark 6.5.11.

Our goal in this section is now to prove the following result of Hazewinkel:

Theorem 6.5.13. *The family $\left(M_w^{\langle s \rangle}\right)_{(w,s)\in\mathfrak{RL}\times\{1,2,3,\ldots\}}$ is an algebraically independent generating set of the $\mathbf{k}$-algebra* QSym.

This will (almost) immediately yield Theorem 6.4.3.

Our first step towards proving Theorem 6.5.13 is the following observation:

Lemma 6.5.14. *The family* $\left(M_w^{\langle s\rangle}\right)_{(w,s)\in\mathfrak{RL}\times\{1,2,3,\ldots\}}$ *is a reindexing of the family* $\left(M_{\operatorname{red}\alpha}^{\langle\gcd\alpha\rangle}\right)_{\alpha\in\mathfrak{L}}$.

Exercise 6.5.15. Prove Lemma 6.5.14.

Next, we show a lemma:

Lemma 6.5.16. *Let α be a nonempty composition. Let $s\in\mathbb{N}$. Then,*

$$s!M_\alpha^{\langle s\rangle}-M_\alpha^s\in\sum_{\substack{\beta\in\operatorname{Comp}_{s|\alpha|};\\ \ell(\beta)\leq(s-1)\ell(\alpha)}}\mathbf{k}M_\beta. \tag{6.5.2}$$

(That is, $s!M_\alpha^{\langle s\rangle}-M_\alpha^s$ is a $\mathbf{k}$-linear combination of terms of the form M_β with β ranging over the compositions of $s\,|\alpha|$ satisfying $\ell\left(\beta\right)\leq\left(s-1\right)\ell\left(\alpha\right)$.)

Exercise 6.5.17. Prove Lemma 6.5.16.

[**Hint:** There are two approaches: One is to apply Proposition 6.5.6(a) and expand the determinant; the other is to argue which monomials can appear in $s!M_\alpha^{\langle s\rangle}-M_\alpha^s$.]

We now return to studying products of monomial quasisymmetric functions:

Lemma 6.5.18. *Let $n\in\mathbb{N}$ and $m\in\mathbb{N}$. Let $u\in\operatorname{Comp}_n$ and $v\in\operatorname{Comp}_m$. Let z be the lexicographically highest element of the multiset $u\sqcup\!\sqcup\, v$. Let h be the multiplicity with which the word z appears in the multiset $u\sqcup\!\sqcup\, v$. Then,*[328]

$$M_uM_v=hM_z+\left(\text{a sum of terms of the form } M_w \text{ with } w\in\operatorname{Comp}_{n+m} \text{ satisfying } w\underset{\text{wll}}{<}z\right).$$

Proof of Lemma 6.5.18. Lemma 6.5.18 was shown during the proof of Lemma 6.4.11(b). □

Corollary 6.5.19. *Let $n\in\mathbb{N}$ and $m\in\mathbb{N}$. Let $u\in\operatorname{Comp}_n$ and $v\in\operatorname{Comp}_m$. Regard u and v as words in $\mathfrak{A}^*$. Assume that u is a Lyndon word. Let $(b_1,b_2,\ldots,b_q)$ be the CFL factorization of the word v.*

Assume that $u\geq b_j$ for every $j\in\{1,2,\ldots,q\}$. Let

$$h=1+\left|\{j\in\{1,2,\ldots,q\}\ \mid\ b_j=u\}\right|.$$

Then,

$$M_uM_v=hM_{uv}+\left(\text{a sum of terms of the form } M_w \text{ with } w\in\operatorname{Comp}_{n+m} \text{ satisfying } w\underset{\text{wll}}{<}uv\right).$$

Exercise 6.5.20. Prove Corollary 6.5.19.

[**Hint:** Apply Lemma 6.5.18, and notice that uv is the lexicographically highest element of the multiset $u\sqcup\!\sqcup\, v$ (by Theorem 6.2.2(e)), and that h is the multiplicity with which this word uv appears in the multiset $u\sqcup\!\sqcup\, v$ (this is a rewriting of Theorem 6.2.2(e)).]

Corollary 6.5.21. *Let $k\in\mathbb{N}$ and $s\in\mathbb{N}$. Let $x\in\operatorname{Comp}_k$ be such that x is a Lyndon word. Then:*

(a) *The lexicographically highest element of the multiset $x\sqcup\!\sqcup\, x^s$ is x^{s+1}.*

(b) *We have*

$$M_xM_{x^s}=\left(s+1\right)M_{x^{s+1}}+\left(\text{a sum of terms of the form } M_w \text{ with } w\in\operatorname{Comp}_{(s+1)k} \text{ satisfying } w\underset{\text{wll}}{<}x^{s+1}\right).$$

(c) *Let $t\in\operatorname{Comp}_{sk}$ be such that $t\underset{\text{wll}}{<}x^s$. Then,*

$$M_xM_t=\left(\text{a sum of terms of the form } M_w \text{ with } w\in\operatorname{Comp}_{(s+1)k} \text{ satisfying } w\underset{\text{wll}}{<}x^{s+1}\right).$$

[328] The following equality makes sense because we have $z\in\operatorname{Comp}_{n+m}$ (by Lemma 6.4.11(a)).

Exercise 6.5.22. Prove Corollary 6.5.21.

[**Hint:** Notice that $\left(\underbrace{x, x, \ldots, x}_{s \text{ times}}\right)$ is the CFL factorization of the word x^s. Now, part (a) of Corollary 6.5.21 follows from Theorem 6.2.2(c), part (b) follows from Corollary 6.5.19, and part (c) from Lemma 6.4.11(c) (using part (a)).]

Corollary 6.5.23. *Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$. Let $u \in \operatorname{Comp}_n$ and $v \in \operatorname{Comp}_m$. Regard u and v as words in $\mathfrak{A}^*$. Let $(a_1, a_2, \ldots, a_p)$ be the CFL factorization of u. Let $(b_1, b_2, \ldots, b_q)$ be the CFL factorization of the word v. Assume that $a_i > b_j$ for every $i \in \{1, 2, \ldots, p\}$ and $j \in \{1, 2, \ldots, q\}$. Then,*

$$M_u M_v = M_{uv} + \left(\text{a sum of terms of the form } M_w \text{ with } w \in \operatorname{Comp}_{n+m} \text{ satisfying } w \underset{\text{wll}}{<} uv\right).$$

Exercise 6.5.24. Prove Corollary 6.5.23.

[**Hint:** Combine Lemma 6.5.18 with the parts (c) and (d) of Theorem 6.2.2.]

Corollary 6.5.25. *Let $n \in \mathbb{N}$. Let $u \in \operatorname{Comp}_n$ be a nonempty composition. Regard u as a word in $\mathfrak{A}^*$. Let $(a_1, a_2, \ldots, a_p)$ be the CFL factorization of u. Let $k \in \{1, 2, \ldots, p-1\}$ be such that $a_k > a_{k+1}$. Let x be the word $a_1 a_2 \cdots a_k$, and let y be the word $a_{k+1} a_{k+2} \cdots a_p$. Then,*

$$M_u = M_x M_y - \left(\text{a sum of terms of the form } M_w \text{ with } w \in \operatorname{Comp}_n \text{ satisfying } w \underset{\text{wll}}{<} u\right).$$

Exercise 6.5.26. Prove Corollary 6.5.25.

[**Hint:** Apply Corollary 6.5.23 to x, y, $|x|$, $|y|$, k, $p-k$, $(a_1, a_2, \ldots, a_k)$ and $(a_{k+1}, a_{k+2}, \ldots, a_p)$ instead of u, v, n, m, p, q, $(a_1, a_2, \ldots, a_p)$ and $(b_1, b_2, \ldots, b_q)$; then, notice that $xy = u$ and $|x| + |y| = n$.]

Corollary 6.5.27. *Let $k \in \mathbb{N}$. Let $x \in \operatorname{Comp}_k$ be a composition. Assume that x is a Lyndon word. Let $s \in \mathbb{N}$. Then,*

$$M_x^s - s! M_{x^s} \in \sum_{\substack{w \in \operatorname{Comp}_{sk};\\ w \underset{\text{wll}}{<} x^s}} \mathbf{k} M_w.$$

(Recall that x^s is defined to be the word $\underbrace{xx \cdots x}_{s \text{ times}}$.)

Exercise 6.5.28. Prove Corollary 6.5.27.

[**Hint:** Rewrite the claim of Corollary 6.5.27 in the form $M_x^s \in s! M_{x^s} + \sum_{\substack{w \in \operatorname{Comp}_{sk};\\ w \underset{\text{wll}}{<} x^s}} \mathbf{k} M_w$. This can be proven by induction over s, where in the induction step we need the following two observations:

(1) We have $M_x M_{x^s} \in (s+1) M_{x^{s+1}} + \sum_{\substack{w \in \operatorname{Comp}_{(s+1)k};\\ w \underset{\text{wll}}{<} x^{s+1}}} \mathbf{k} M_w$.

(2) For every $t \in \operatorname{Comp}_{sk}$ satisfying $t \underset{\text{wll}}{<} x^s$, we have $M_x M_t \in \sum_{\substack{w \in \operatorname{Comp}_{(s+1)k};\\ w \underset{\text{wll}}{<} x^{s+1}}} \mathbf{k} M_w$.

These two observations follow from parts (b) and (c) of Corollary 6.5.21.]

Corollary 6.5.29. *Let $k \in \mathbb{N}$. Let $x \in \operatorname{Comp}_k$ be a composition. Assume that x is a Lyndon word. Let $s \in \mathbb{N}$. Then,*

$$M_x^{\langle s \rangle} - M_{x^s} \in \sum_{\substack{w \in \operatorname{Comp}_{sk};\\ w \underset{\text{wll}}{<} x^s}} \mathbf{k} M_w.$$

(Recall that x^s is defined to be the word $\underbrace{xx \cdots x}_{s \text{ times}}$.)

Exercise 6.5.30. Prove Corollary 6.5.29.

[**Hint:** Lemma 6.5.16 (applied to $\alpha = x$) yields

$$s! M_x^{\langle s \rangle} - M_x^s \in \sum_{\substack{\beta \in \operatorname{Comp}_{sk};\\ \ell(\beta) \leq (s-1)\ell(x)}} M_\beta = \sum_{\substack{w \in \operatorname{Comp}_{sk};\\ \ell(w) \leq (s-1)\ell(x)}} \mathbf{k} M_w \subset \sum_{\substack{w \in \operatorname{Comp}_{sk};\\ w \underset{\text{wll}}{<} x^s}} \mathbf{k} M_w$$

[329]. Adding this to the claim of Corollary 6.5.27, obtain $s! M_x^{\langle s \rangle} - s! M_{x^s} \in \sum_{\substack{w \in \operatorname{Comp}_{sk};\\ w \underset{\text{wll}}{<} x^s}} \mathbf{k} M_w$, that is, $s! \left(M_x^{\langle s \rangle} - M_{x^s} \right) \in \sum_{\substack{w \in \operatorname{Comp}_{sk};\\ w \underset{\text{wll}}{<} x^s}} \mathbf{k} M_w$. It remains to get rid of the $s!$ on the left hand side. Assume WLOG that $\mathbf{k} = \mathbb{Z}$, and argue that every $f \in \operatorname{QSym}$ satisfying $s! \cdot f \in \sum_{\substack{w \in \operatorname{Comp}_{sk};\\ w \underset{\text{wll}}{<} x^s}} \mathbf{k} M_w$ must itself lie in $\sum_{\substack{w \in \operatorname{Comp}_{sk};\\ w \underset{\text{wll}}{<} x^s}} \mathbf{k} M_w$.]

We are now ready to prove Theorem 6.5.13:

Exercise 6.5.31. Prove Theorem 6.5.13.

[**Hint:** Lemma 6.5.14 yields that the family $\left(M_w^{\langle s \rangle} \right)_{(w,s) \in \mathfrak{RL} \times \{1,2,3,\ldots\}}$ is a reindexing of the family $\left(M_{\operatorname{red} w}^{\langle \gcd w \rangle} \right)_{w \in \mathfrak{L}}$. Hence, it is enough to prove that the family $\left(M_{\operatorname{red} w}^{\langle \gcd w \rangle} \right)_{w \in \mathfrak{L}}$ is an algebraically independent generating set of the $\mathbf{k}$-algebra QSym. The latter claim, in turn, will follow from Lemma 6.3.7(c)[330] once it is proven that the family $\left(M_{\operatorname{red} w}^{\langle \gcd w \rangle} \right)_{w \in \mathfrak{L}}$ generates the $\mathbf{k}$-algebra QSym. So it remains to show that the family $\left(M_{\operatorname{red} w}^{\langle \gcd w \rangle} \right)_{w \in \mathfrak{L}}$ generates the $\mathbf{k}$-algebra QSym.

Let U denote the $\mathbf{k}$-subalgebra of QSym generated by $\left(M_{\operatorname{red} w}^{\langle \gcd w \rangle} \right)_{w \in \mathfrak{L}}$. It then suffices to prove that $U = \operatorname{QSym}$. To this purpose, it is enough to prove that

$$M_\beta \in U \qquad \text{for every composition } \beta. \tag{6.5.3}$$

For every reduced Lyndon composition α and every $j \in \{1,2,3,\ldots\}$, the quasisymmetric function $M_\alpha^{\langle j \rangle}$ is an element of the family $\left(M_{\operatorname{red} w}^{\langle \gcd w \rangle} \right)_{w \in \mathfrak{L}}$ and thus belongs to U. Combine this with Exercise 6.5.4(d) to see that

$$M_\beta^{\langle s \rangle} \in U \qquad \text{for every Lyndon composition } \beta \text{ and every } s \in \{1,2,3,\ldots\} \tag{6.5.4}$$

(because every Lyndon composition β can be written as $\alpha\{n\}$ for a reduced Lyndon composition α and an $n \in \{1,2,3,\ldots\}$). Now, prove (6.5.3) by strong induction: first, induct on $|\beta|$, and then, for fixed $|\beta|$, induct on β in the wll-order. The induction step looks as follows: Fix some composition α, and assume (as induction hypothesis) that:

- (6.5.3) holds for every composition β satisfying $|\beta| < |\alpha|$;
- (6.5.3) holds for every composition β satisfying $|\beta| = |\alpha|$ and $\beta \underset{\text{wll}}{<} \alpha$.

It remains to prove that (6.5.3) holds for $\beta = \alpha$. In other words, it remains to prove that $M_\alpha \in U$.

Let $(a_1, a_2, \ldots, a_p)$ be the CFL factorization of the word α. Assume WLOG that $p \neq 0$ (else, all is trivial). We are in one of the following two cases:

Case 1: All of the words a_1, a_2, $\ldots$, a_p are equal.

Case 2: Not all of the words a_1, a_2, $\ldots$, a_p are equal.

[329] since every $w \in \operatorname{Comp}_{sk}$ with the property that $\ell(w) \leq (s-1)\ell(x)$ must satisfy $w \underset{\text{wll}}{<} x^s$

[330] applied to $A = \operatorname{QSym}$, $b_w = M_{\operatorname{red} w}^{\langle \gcd w \rangle}$, $\operatorname{wt}(N) = N$ and $g_u = M_u$

In Case 2, there exists a $k \in \{1, 2, \ldots, p-1\}$ satisfying $a_k > a_{k+1}$ (since $a_1 \geq a_2 \geq \cdots \geq a_p$), and thus Corollary 6.5.25 (applied to $u = \alpha$, $n = |\alpha|$, $x = a_1 a_2 \cdots a_k$ and $y = a_{k+1} a_{k+2} \cdots a_p$) shows that

$$\begin{aligned} M_\alpha &= \underbrace{M_{a_1 a_2 \cdots a_k}}_{\substack{\in U \\ \text{(by the induction} \\ \text{hypothesis)}}} \underbrace{M_{a_{k+1} a_{k+2} \cdots a_p}}_{\substack{\in U \\ \text{(by the induction} \\ \text{hypothesis)}}} \\ &\qquad - \left(\text{a sum of terms of the form } \underbrace{M_w}_{\substack{\in U \\ \text{(by the induction} \\ \text{hypothesis)}}} \text{ with } w \in \operatorname{Comp}_{|\alpha|} \text{ satisfying } w \underset{\text{wll}}{<} \alpha \right) \\ &\in UU - (\text{a sum of terms in } U) \subset U. \end{aligned}$$

Hence, it only remains to deal with Case 1. In this case, set $x = a_1 = a_2 = \cdots = a_p$. Thus, $\alpha = a_1 a_2 \cdots a_p = x^p$, whence $|\alpha| = p\,|x|$. But Corollary 6.5.29 (applied to $s = p$ and $k = |x|$) yields

$$\begin{aligned} M_x^{\langle p \rangle} - M_{x^p} &\in \sum_{\substack{w \in \operatorname{Comp}_{p|x|}; \\ w \underset{\text{wll}}{<} x^p}} \mathbf{k} M_w = \sum_{\substack{w \in \operatorname{Comp}_{|\alpha|}; \\ w \underset{\text{wll}}{<} \alpha}} \mathbf{k} \underbrace{M_w}_{\substack{\in U \\ \text{(by the induction} \\ \text{hypothesis)}}} \qquad (\text{since } p\,|x| = |\alpha| \text{ and } x^p = \alpha) \\ &\subset \sum_{\substack{w \in \operatorname{Comp}_N; \\ w \underset{\text{wll}}{<} \alpha}} \mathbf{k} U \subset U, \end{aligned}$$

so that $M_{x^p} \in \underbrace{M_x^{\langle p \rangle}}_{\substack{\in U \\ \text{(by (6.5.4))}}} - U \subset U - U \subset U$. This rewrites as $M_\alpha \in U$ (since $\alpha = x^p$). So $M_\alpha \in U$ is proven in both Cases 1 and 2, and thus the induction proof of (6.5.3) is finished.]

Exercise 6.5.32. Prove Theorem 6.4.3.

Of course, this proof of Theorem 6.4.3 yields a new (third) proof for Proposition 6.4.4.

We notice the following corollary of our approach to Theorem 6.4.3:

Corollary 6.5.33. *The* Λ*-algebra* QSym *is a polynomial algebra (over* Λ*).*

Exercise 6.5.34. Prove Corollary 6.5.33.

[**Hint:** The algebraically independent generating set $\left(M_w^{\langle s \rangle} \right)_{(w,s) \in \mathfrak{RL} \times \{1,2,3,\ldots\}}$ of QSym contains the elements $M_{(1)}^{\langle s \rangle} = e_s \in \Lambda$ for all $s \in \{1, 2, 3, \ldots\}$.]

6.6. The Gessel-Reutenauer bijection and symmetric functions.

In this section, we shall discuss the Gessel-Reutenauer bijection between words and multisets of aperiodic necklaces, and use it to study another family of symmetric functions.

The Gessel-Reutenauer bijection was studied in [82], where it was applied to various enumeration problems (e.g., counting permutations in $\mathfrak{S}_n$ with given descent set and given cycle type); it is also closely related to the Burrows-Wheeler bijection used in data compression ([45]), and to the structure of free Lie algebras ([81], [182]). We shall first introduce the Gessel-Reutenauer bijection and study it combinatorially in Subsection 6.6.1; then, in the following Subsection 6.6.2, we shall apply it to symmetric functions.

6.6.1. *Necklaces and the Gessel-Reutenauer bijection.*

We begin with definitions, some of which have already been made in Exercise 6.1.34:

Definition 6.6.1. Throughout Section 6.6, we shall freely use Definition 6.1.1 and Definition 6.1.13. We fix a totally ordered alphabet $\mathfrak{A}$. (This alphabet can be arbitrary, although most examples will use $\mathfrak{A} = \{1 < 2 < 3 < \cdots\}$.)

Let C denote the infinite cyclic group, written multiplicatively. Fix a generator c of C. [331]

[331] So C is a group isomorphic to $(\mathbb{Z}, +)$, and the isomorphism $(\mathbb{Z}, +) \to C$ sends every $n \in \mathbb{Z}$ to c^n. (Recall that we write the binary operation of C as $\cdot$ instead of $+$.)

For any positive integer n, the group C acts on $\mathfrak{A}^n$ from the left according to the rule

$$c \cdot (a_1, a_2, \ldots, a_n) = (a_2, a_3, \ldots, a_n, a_1) \qquad \text{for all } (a_1, a_2, \ldots, a_n) \in \mathfrak{A}^n .$$

[332] The orbits of this C-action will be called *n-necklaces*[333]; they form a set partition of the set $\mathfrak{A}^n$.

The n-necklace containing a given n-tuple $w \in \mathfrak{A}^n$ will be denoted by $[w]$.

A *necklace* shall mean an n-necklace for some positive integer n. Thus, for each nonempty word w, there is a well-defined necklace $[w]$ (namely, $[w]$ is an n-necklace, where $n = \ell(w)$).

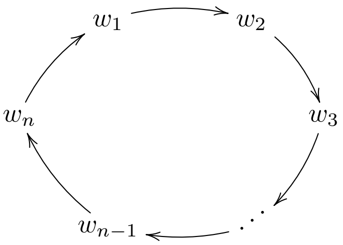
The *period* of a necklace N is defined as the positive integer $|N|$. (This $|N|$ is indeed a positive integer, since N is a finite nonempty set[334].)

An n-necklace is said to be *aperiodic* if its period is n.

Example 6.6.2. Let $\mathfrak{A}$ be the alphabet $\{1 < 2 < 3 < \cdots\}$. The orbit of the word 223 under the C-action is the 3-necklace $\{223, 232, 322\}$; it is an aperiodic 3-necklace. The orbit of the word 223223 under the C-action is the 6-necklace $\{223223, 232232, 322322\}$; it is not aperiodic (since it has period 3). The orbit of any nonempty word $w = (w_1, w_2, \ldots, w_n) \in \mathfrak{A}^n$ is the n-necklace

$$\{(w_i, w_{i+1}, \ldots, w_n, w_1, w_2, \ldots, w_{i-1}) \mid i \in \{1, 2, \ldots, n\}\} .$$

We can draw this n-necklace on the plane as follows:



It is easy to see that the notion of an "aperiodic necklace" we just defined is equivalent to the notion of a "primitive necklace" used in Exercise 4.6.4(b).

Exercise 6.1.34(a) shows that any n-necklace for any positive integer n is a finite nonempty set. In other words, any necklace is a finite nonempty set.

Let us next introduce some notations regarding words and permutations. We recall that a cycle of a permutation $\tau \in \mathfrak{S}_n$ is an orbit under the action of τ on $\{1, 2, \ldots, n\}$. (This orbit can be a 1-element set, when τ has fixed points.) We begin with a basic definition:

Definition 6.6.3. Let $\tau \in \mathfrak{S}_n$ be a permutation. Let $h \in \{1, 2, \ldots, n\}$.

(a) We let $\operatorname{ord}_\tau(h)$ denote the smallest positive integer i such that $\tau^i(h) = h$. (Basic properties of permutations show that this i exists.)

(b) Let $w = (w_1, w_2, \ldots, w_n) \in \mathfrak{A}^n$ be a word. Then, $w_{\tau,h}$ shall denote the word $w_{\tau^1(h)} w_{\tau^2(h)} \cdots w_{\tau^k(h)}$, where $k = \operatorname{ord}_\tau(h)$.

Example 6.6.4. Let τ be the permutation $3142765 \in \mathfrak{S}_7$ (in one-line notation). Then, $\operatorname{ord}_\tau(1) = 4$ (since $\tau^4(1) = 1$, but $\tau^i(1) \neq 1$ for every positive integer $i < 4$). Likewise, $\operatorname{ord}_\tau(2) = 4$ and $\operatorname{ord}_\tau(3) = 4$ and $\operatorname{ord}_\tau(4) = 4$ and $\operatorname{ord}_\tau(5) = 2$ and $\operatorname{ord}_\tau(6) = 1$ and $\operatorname{ord}_\tau(7) = 2$.

Now, let w be the word $4112524 \in \mathfrak{A}^7$. Then,

$$\begin{aligned} w_{\tau,3} &= w_{\tau^1(3)} w_{\tau^2(3)} w_{\tau^3(3)} w_{\tau^4(3)} \qquad (\text{since } \operatorname{ord}_\tau(3) = 4) \\ &= w_4 w_2 w_1 w_3 \\ &\qquad \left(\text{since } \tau^1(3) = 4 \text{ and } \tau^2(3) = \tau(4) = 2 \text{ and } \tau^3(3) = \tau(2) = 1 \text{ and } \tau^4(3) = \tau(1) = 3\right) \\ &= 2141. \end{aligned}$$

Likewise, we can check that $w_{\tau,1} = w_3 w_4 w_2 w_1 = 1214$ and $w_{\tau,5} = w_7 w_5 = 45$ and $w_{\tau,6} = w_6 = 2$.

[332] In other words, c rotates any n-tuple of elements of $\mathfrak{A}$ cyclically to the left. Thus, $c^n \in C$ acts trivially on $\mathfrak{A}^n$, and so this action of C on $\mathfrak{A}^n$ factors through $C / \langle c^n \rangle$ (a cyclic group of order n).

[333] See Exercise 6.1.34 for the motivation behind this word.

Notice that there are no 0-necklaces, because we required n to be positive in the definition of a necklace. This is intentional.

[334] by Exercise 6.1.34(a), because N is an n-necklace for some positive integer n

We begin the study of the words $w_{\tau,h}$ by stating some of their simplest properties:[335]

Proposition 6.6.5. *Let $w = (w_1, w_2, \ldots, w_n) \in \mathfrak{A}^n$ be a word. Let $\tau \in \mathfrak{S}_n$. Let $h \in \{1, 2, \ldots, n\}$. Then:*

(a) *The word $w_{\tau,h}$ is nonempty and has length $\operatorname{ord}_\tau(h)$.*
(b) *The first letter of the word $w_{\tau,h}$ is $w_{\tau(h)}$.*
(c) *The last letter of the word $w_{\tau,h}$ is w_h.*
(d) *We have $w_{\tau,\tau(h)} = c \cdot w_{\tau,h}$.*
(e) *We have $w_{\tau,\tau^i(h)} = c^i \cdot w_{\tau,h}$ for each $i \in \mathbb{Z}$.*

Recall that if $n \in \mathbb{N}$ and if $w \in \mathfrak{A}^n$ is a word, then a permutation $\operatorname{std} w \in \mathfrak{S}_n$ was defined in Definition 5.3.3. The words $w_{\tau,h}$ have particularly nice properties when $\tau = (\operatorname{std} w)^{-1}$:

Lemma 6.6.6. *Let $w = (w_1, w_2, \ldots, w_n) \in \mathfrak{A}^n$ be a word. Let τ be the permutation $(\operatorname{std} w)^{-1} \in \mathfrak{S}_n$. Let α and β be two elements of $\{1, 2, \ldots, n\}$ such that $\alpha < \beta$. Then:*

(a) *If $\tau^{-1}(\alpha) < \tau^{-1}(\beta)$, then $w_\alpha \leq w_\beta$.*
(b) *If $\tau^{-1}(\alpha) \geq \tau^{-1}(\beta)$, then $w_\alpha > w_\beta$.*
(c) *We have $w_{\tau(\alpha)} \leq w_{\tau(\beta)}$.*
(d) *If $\tau(\alpha) \geq \tau(\beta)$, then $w_{\tau(\alpha)} < w_{\tau(\beta)}$.*
(e) *If $w_{\tau(\alpha)} = w_{\tau(\beta)}$, then $\tau(\alpha) < \tau(\beta)$.*
(f) *If $w_{\tau,\alpha} = w_{\tau,\beta}$, then $\tau(\alpha) < \tau(\beta)$ and $w_{\tau,\tau(\alpha)} = w_{\tau,\tau(\beta)}$.*
(g) *If $w_{\tau,\alpha} = w_{\tau,\beta}$, then $\tau^i(\alpha) < \tau^i(\beta)$ for each $i \in \mathbb{N}$.*
(h) *Let $j \in \mathbb{N}$ be such that every $i \in \{0, 1, \ldots, j-1\}$ satisfies $w_{\tau^{i+1}(\alpha)} = w_{\tau^{i+1}(\beta)}$. Then, $w_{\tau^{j+1}(\alpha)} \leq w_{\tau^{j+1}(\beta)}$.*

Proposition 6.6.7. *Let $w \in \mathfrak{A}^n$ be a word. Let τ be the permutation $(\operatorname{std} w)^{-1} \in \mathfrak{S}_n$. Let z be a cycle of τ. Then:*

(a) *For each $h \in z$, we have $[w_{\tau,h}] = \{w_{\tau,i} \mid i \in z\}$.*
(b) *If α and β are two distinct elements of z, then $w_{\tau,\alpha} \neq w_{\tau,\beta}$.*
(c) *We have $|\{w_{\tau,i} \mid i \in z\}| = |z|$.*
(d) *The set $\{w_{\tau,i} \mid i \in z\}$ is an aperiodic necklace.*

Exercise 6.6.8. Prove Proposition 6.6.5, Lemma 6.6.6 and Proposition 6.6.7.

Definition 6.6.9. Let $w \in \mathfrak{A}^n$ be a word. Let τ be the permutation $(\operatorname{std} w)^{-1} \in \mathfrak{S}_n$. Let z be a cycle of τ. Then, we define an aperiodic necklace $[w]_z$ by $[w]_z = \{w_{\tau,i} \mid i \in z\}$. (This is indeed an aperiodic necklace, according to Proposition 6.6.7(d).)

Example 6.6.10. Let $\mathfrak{A}$ be the alphabet $\{1 < 2 < 3 < \cdots\}$, and let w be the word $2511321 \in \mathfrak{A}^7$. Let τ be the permutation $(\operatorname{std} w)^{-1} \in \mathfrak{S}_7$; this is the permutation 3471652 (in one-line notation). One cycle of τ is $z = \{1, 3, 7, 2, 4\}$. The corresponding aperiodic necklace $[w]_z$ is

$$\begin{aligned} [w]_z &= \{w_{\tau,i} \mid i \in z\} = \{w_{\tau,1}, w_{\tau,3}, w_{\tau,7}, w_{\tau,2}, w_{\tau,4}\} \qquad (\text{since } z = \{1,3,7,2,4\}) \\ &= \{11512, 15121, 51211, 12115, 21151\} = [11512]. \end{aligned}$$

Definition 6.6.11. We let $\mathfrak{N}$ be the set of all necklaces. We let $\mathfrak{N}^{\mathfrak{a}}$ be the set of all aperiodic necklaces. We let $\mathfrak{M}\mathfrak{N}^{\mathfrak{a}}$ be the set of all finite multisets of aperiodic necklaces.

Definition 6.6.12. We define a map $\operatorname{GR} : \mathfrak{A}^* \to \mathfrak{M}\mathfrak{N}^{\mathfrak{a}}$ as follows:

Let $w \in \mathfrak{A}^*$. Let $n = \ell(w)$ (so that $w \in \mathfrak{A}^n$). Let τ be the permutation $(\operatorname{std} w)^{-1} \in \mathfrak{S}_n$. Then, we define the multiset $\operatorname{GR} w \in \mathfrak{M}\mathfrak{N}^{\mathfrak{a}}$ by setting

$$\operatorname{GR} w = \{[w]_z \mid z \text{ is a cycle of } \tau\}_{\text{multiset}}.$$

(This multiset $\operatorname{GR} w$ is indeed a finite multiset of aperiodic necklaces[336], and thus belongs to $\mathfrak{M}\mathfrak{N}^{\mathfrak{a}}$.)

[335] See Exercise 6.6.8 below for the proof of Proposition 6.6.5, as well as for the proofs of all other propositions stated before Exercise 6.6.8.

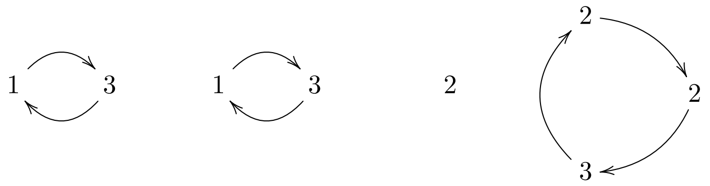
[336] Indeed, this multiset $\operatorname{GR} w$ is finite (since τ has only finitely many cycles), and its elements $[w]_z$ are aperiodic necklaces (as we have seen in the definition of $[w]_z$).

Example 6.6.13. Let $\mathfrak{A}$ be the alphabet $\{1 < 2 < 3 < \cdots\}$, and let $w = 33232112 \in \mathfrak{A}^8$.

To compute $\operatorname{GR} w$, we first notice that $\operatorname{std} w = 67384125$ (in one-line notation). Hence, the permutation τ from Definition 6.6.12 satisfies $\tau = (\operatorname{std} w)^{-1} = 67358124$. The cycles of τ are $\{1,6\}$, $\{2,7\}$, $\{3\}$ and $\{4,5,8\}$. Thus,

$$\begin{aligned}\operatorname{GR} w &= \{[w]_z \mid z \text{ is a cycle of } \tau\}_{\text{multiset}} = \left\{[w]_{\{1,6\}}, [w]_{\{2,7\}}, [w]_{\{3\}}, [w]_{\{4,5,8\}}\right\}_{\text{multiset}} \\ &= \{[31], [31], [2], [322]\}_{\text{multiset}} = \{[13], [13], [2], [223]\}_{\text{multiset}}\end{aligned}$$

(since $[31] = [13]$ and $[322] = [223]$ as necklaces). Drawn on the plane, the necklaces in $\operatorname{GR} w$ look as follows:



The map GR is called the *Gessel-Reutenauer bijection*. In order to show that it indeed is a bijection, we shall construct its inverse. First, we introduce some further objects.

Definition 6.6.14. A nonempty word w is said to be *aperiodic* if there exist no $m \geq 2$ and $u \in \mathfrak{A}^*$ satisfying $w = u^m$.

Let $\mathfrak{A}^{\mathfrak{a}}$ be the set of all aperiodic words in $\mathfrak{A}^*$.

For example, the word 132231 is aperiodic, but the word 132132 is not (since $132132 = u^m$ for $u = 132$ and $m = 2$).

Aperiodic words are directly connected to aperiodic necklaces, as the following facts show:[337]

Proposition 6.6.15. *Let $w \in \mathfrak{A}^*$ be a nonempty word. Then, the word w is aperiodic if and only if the necklace $[w]$ is aperiodic.*

Corollary 6.6.16. *Let $w \in \mathfrak{A}^*$ be an aperiodic word. Then, the word $c \cdot w$ is aperiodic.*[338]

Corollary 6.6.17. *Each aperiodic necklace is a set of aperiodic words.*

Let us now introduce a new total order on the set $\mathfrak{A}^{\mathfrak{a}}$ of all aperiodic words:

Definition 6.6.18. Let u and v be two aperiodic words. Then, we write $u \leq_\omega v$ if and only if $uv \leq vu$. Thus, we have defined a binary relation $\leq_\omega$ on the set $\mathfrak{A}^{\mathfrak{a}}$ of all aperiodic words.

Proposition 6.6.19. *The relation $\leq_\omega$ on the set $\mathfrak{A}^{\mathfrak{a}}$ is the smaller-or-equal relation of a total order.*

For the next proposition, we should recall Definition 6.6.1 (and, in particular, the meaning of c and its action on words).

Proposition 6.6.20. *Let u and v be two aperiodic words.*

(a) *We have $u \leq_\omega v$ if and only if either $u_1 < v_1$ or ($u_1 = v_1$ and $c \cdot u \leq_\omega c \cdot v$).* [339]

(b) *If $u \neq v$, then there exists some $i \in \mathbb{N}$ satisfying $\left(c^i \cdot u\right)_1 \neq \left(c^i \cdot v\right)_1$.*

(c) *We have $u \leq_\omega v$ if and only if the smallest $i \in \mathbb{N}$ satisfying $\left(c^i \cdot u\right)_1 \neq \left(c^i \cdot v\right)_1$* ***either*** *does not exist* ***or*** *satisfies $\left(c^i \cdot u\right)_1 < \left(c^i \cdot v\right)_1$.*

(d) *Let n and m be positive integers such that $n\ell(u) = m\ell(v)$. We have $u \leq_\omega v$ if and only if $u^n \leq v^m$.*

Remark 6.6.21. We are avoiding the use of infinite words here; if we didn't, we could restate the relation $\leq_\omega$ in a simpler way (which is easily seen to be equivalent to Proposition 6.6.20(c)): Two aperiodic words u and v satisfy $u \leq_\omega v$ if and only if $u^\infty \leq v^\infty$. Here, for any nonempty word w, we are letting w^∞ denote the infinite word

$$\left(w_1, w_2, \ldots, w_{\ell(w)}, w_1, w_2, \ldots, w_{\ell(w)}, w_1, w_2, \ldots, w_{\ell(w)}, \ldots\right)$$

[337] See Exercise 6.6.23 for the proofs of all unproved statements made until Exercise 6.6.23.

[338] See Definition 6.6.1 for the definition of c and its action on words.

[339] The relation "$c \cdot u \leq_\omega c \cdot v$" here makes sense because the words $c \cdot u$ and $c \cdot v$ are aperiodic (by Corollary 6.6.16).

(that is, the word w repeated endlessly), and the symbol "$\leq$" in "$u^\infty \leq v^\infty$" refers to the lexicographic order on $\mathfrak{A}^\infty$.

Other equivalent descriptions of the relation $\leq_\omega$ (or, more precisely, of the "strictly less" relation corresponding to it) can be found in [54, Corollary 11].

Proposition 6.6.22. *Let $w \in \mathfrak{A}^n$ be a word. Let τ be the permutation $(\operatorname{std} w)^{-1} \in \mathfrak{S}_n$. Then:*

(a) *The words $w_{\tau,1}, w_{\tau,2}, \ldots, w_{\tau,n}$ are aperiodic.*
(b) *We have $w_{\tau,1} \leq_\omega w_{\tau,2} \leq_\omega \cdots \leq_\omega w_{\tau,n}$.*

Exercise 6.6.23. Prove Proposition 6.6.15, Corollary 6.6.16, Corollary 6.6.17, Proposition 6.6.19, Proposition 6.6.20 and Proposition 6.6.22.

We need two more notations about multisets:

Definition 6.6.24. Let T be a totally ordered set, and let $\leq_T$ be the smaller-or-equal relation of T. Let M be a finite multiset of elements of T. Then, there is a unique list $(m_1, m_2, \ldots, m_n)$ such that

$$\{m_1, m_2, \ldots, m_n\}_{\text{multiset}} = M \qquad \text{and} \qquad m_1 \leq_T m_2 \leq_T \cdots \leq_T m_n.$$

This list $(m_1, m_2, \ldots, m_n)$ is obtained by listing all elements of M (with their multiplicities) in increasing order (increasing with respect to $\leq_T$). We shall refer to this list $(m_1, m_2, \ldots, m_n)$ as the *$\leq_T$-increasing list* of M.

(For example, the $\leq_\mathbb{Z}$-increasing list of $\{1, 2, 3, 2, 1\}_{\text{multiset}}$ is $(1, 1, 2, 2, 3)$.)

Definition 6.6.25. Let S be a finite multiset.

(a) The *support* $\operatorname{Supp} S$ is defined to be the set of all elements of S. Thus, if $S = \{m_1, m_2, \ldots, m_n\}_{\text{multiset}}$, then $\operatorname{Supp} S = \{m_1, m_2, \ldots, m_n\}$.
(b) For each $s \in S$, let M_s be a finite multiset. Then, we define the *multiset union* $\biguplus_{s \in S} M_s$ to be the finite multiset M with the following property: For any object x, we have

$$(\text{multiplicity of } x \text{ in } M) = \sum_{s \in \operatorname{Supp} S} (\text{multiplicity of } s \text{ in } S) \cdot (\text{multiplicity of } x \text{ in } M_s).$$

For example:

- If $S = \{1, 2, 3\}_{\text{multiset}}$ and $M_s = \{s, s+1\}_{\text{multiset}}$ for each $s \in \operatorname{Supp} S$, then $\biguplus_{s \in S} M_s = \{1, 2, 2, 3, 3, 4\}_{\text{multiset}}$.
- If $S = \{1, 1, 2\}_{\text{multiset}}$ and $M_s = \{s, s+1\}_{\text{multiset}}$ for each $s \in \operatorname{Supp} S$, then $\biguplus_{s \in S} M_s = \{1, 1, 2, 2, 2, 3\}_{\text{multiset}}$.

We regard each set as a multiset; thus, the multiset union $\biguplus_{s \in S} M_s$ is also defined when the M_s are sets.

Now, we can construct the inverse of the Gessel-Reutenauer bijection:

Definition 6.6.26. We define a map $\operatorname{RG} : \mathfrak{MN}^\mathfrak{a} \to \mathfrak{A}^*$ as follows:

Let $M \in \mathfrak{MN}^\mathfrak{a}$ be a finite multiset of aperiodic necklaces. Let $M' = \biguplus_{N \in M} N$. (We are here using the fact that each necklace $N \in M$ is a finite set, thus a finite multiset.) Notice that M' is a finite multiset of aperiodic words[340]. Let $(m_1, m_2, \ldots, m_n)$ be the $\leq_\omega$-increasing list of M'. For each $i \in \{1, 2, \ldots, n\}$, let ℓ_i be the last letter of the nonempty word m_i. Then, $\operatorname{RG}(M)$ is defined to be the word $(\ell_1, \ell_2, \ldots, \ell_n) \in \mathfrak{A}^*$.

Example 6.6.27. Let $\mathfrak{A}$ be the alphabet $\{1 < 2 < 3 < \cdots\}$, and let $M = \{[13], [13], [2], [223]\}_{\text{multiset}}$. Clearly, $M \in \mathfrak{MN}^\mathfrak{a}$ (since M is a finite multiset of aperiodic necklaces). (Actually, M is the multiset of

[340] Indeed:

- Each $N \in M$ is an aperiodic necklace (since M is a multiset of aperiodic necklaces), and thus (by Corollary 6.6.17) a set of aperiodic words. Therefore, $\biguplus_{N \in M} N$ is a multiset of aperiodic words.
- Each $N \in M$ is a necklace, and thus is a finite set (since any necklace is a finite set). Since the multiset M is also finite, this shows that $\biguplus_{N \in M} N$ is finite.

Thus, $\biguplus_{N \in M} N$ is a finite multiset of aperiodic words. In other words, M' is a finite multiset of aperiodic words (since $M' = \biguplus_{N \in M} N$).

aperiodic necklaces drawn in Example 6.6.13.) In order to compute the word $\operatorname{RG}(M)$, let us first compute the multiset M' from Definition 6.6.26. Indeed, the definition of M' yields

$$\begin{aligned} M' &= \biguplus_{N\in M} N = \underbrace{[13]}_{=\{13,31\}} \uplus \underbrace{[13]}_{=\{13,31\}} \uplus \underbrace{[2]}_{=\{2\}} \uplus \underbrace{[223]}_{=\{223,232,322\}} \\ &\qquad \left(\text{where we are using the notation } M_1 \uplus M_2 \uplus \cdots \uplus M_k \text{ for a multiset union } \biguplus_{s\in\{1,2,\ldots,k\}} M_s\right) \\ &= \{13,31\} \uplus \{13,31\} \uplus \{2\} \uplus \{223,232,322\} \\ &= \{13,31,13,31,2,223,232,322\}_{\text{multiset}} . \end{aligned}$$

Hence, the $\leq_\omega$-increasing list of M' is $(13,13,2,223,232,31,31,322)$ (since $13 \leq_\omega 13 \leq_\omega 2 \leq_\omega 223 \leq_\omega 232 \leq_\omega 31 \leq_\omega 31 \leq_\omega 322$). The last letters of the words in this list are $3,3,2,3,2,1,1,2$ (in this order). Hence, Definition 6.6.26 shows that

$$\operatorname{RG}(M) = (3,3,2,3,2,1,1,2) = 33232112.$$

Remark 6.6.28. The $\leq_\omega$-increasing list of a multiset M' of aperiodic words is not always the same as its $\leq$-increasing list. For example, the $\leq_\omega$-increasing list of $\{2,21\}$ is $(21,2)$ (since $21 \leq_\omega 2$), whereas its $\leq$-increasing list is $(2,21)$ (since $2 \leq 21$).

A comparison of Examples 6.6.13 and 6.6.27 suggests that the maps GR and RG undo one another. This is indeed true, as the following theorem (due to Gessel and Reutenauer [82, Lemma 3.4 and Example 3.5]; also proved in [182, Theorem 7.20], [51, Theorem 3.1 and Proposition 3.1] and [81, §2]) shows:

Theorem 6.6.29. *The maps* $\operatorname{GR} : \mathfrak{A}^* \to \mathfrak{MN}^{\mathfrak{a}}$ *and* $\operatorname{RG} : \mathfrak{MN}^{\mathfrak{a}} \to \mathfrak{A}^*$ *are mutually inverse bijections.*

Exercise 6.6.30. Prove Theorem 6.6.29.

[**Hint:** First, use Proposition 6.6.22 to show that $\operatorname{RG} \circ \operatorname{GR} = \operatorname{id}$. Then recall the fact that any injective map between two finite sets of the same sizes is a bijection. This does not directly apply here, since the sets $\mathfrak{A}^*$ and $\mathfrak{MN}^{\mathfrak{a}}$ are usually not finite. However, GR can be restricted to a map between two appropriate finite subsets, obtained by focussing on a finite sub-alphabet of $\mathfrak{A}$ and fixing the length of the words; these subsets can be shown to have equal size using the Chen-Fox-Lyndon factorization (see the following paragraph for the connection).[341]]

Theorem 6.6.29 shows that the sets $\mathfrak{A}^*$ and $\mathfrak{MN}^{\mathfrak{a}}$ are in bijection. This bijection is in some sense similar to the Chen-Fox-Lyndon factorization[342], and preserves various quantities (for example, the number of times a given letter a appears in a word $w \in \mathfrak{A}^*$ equals the number of times this letter a appears in the words in the corresponding multiset $\operatorname{GR} w \in \mathfrak{MN}^{\mathfrak{a}}$, provided that we pick one representative of each necklace in $\operatorname{GR} w$), and predictably affects other quantities (for example, the cycles of the standardization $\operatorname{std} w$ of a word $w \in \mathfrak{A}^*$ have the same lengths as the aperiodic necklaces in the corresponding multiset $\operatorname{GR} w \in \mathfrak{MN}^{\mathfrak{a}}$); these properties have ample applications to enumerative questions (discussed in [82]).

Remark 6.6.31. The Gessel-Reutenauer bijection relates to the *Burrows-Wheeler transformation* (e.g., [45, §2]). Indeed, the latter sends an aperiodic word $w \in \mathfrak{A}^{\mathfrak{a}}$ to the word $\operatorname{RG}(\{[w]\}_{\text{multiset}})$ obtained by applying RG to the multiset consisting of the single aperiodic necklace $[w]$. This transformation is occasionally applied in (lossless) data compression, as the word $\operatorname{RG}(\{[w]\}_{\text{multiset}})$ tends to have many strings of consecutive equal letters when w has substrings occurring multiple times (for example, if $\mathfrak{A} = \{a < b < c < d < \cdots\}$ and $w = bananaban$, then $\operatorname{RG}(\{[w]\}_{\text{multiset}}) = nnbbnaaaa$), and strings of consecutive equal letters can easily be compressed. (In order to guarantee that w can be recovered from the result, one can add a new letter ζ – called a "sentinel symbol" – to the alphabet $\mathfrak{A}$, and apply the Burrows-Wheeler transformation to the word

[341] This argument roughly follows [81].

[342] The Chen-Fox-Lyndon factorization (Theorem 6.1.27) provides a bijection between words in $\mathfrak{A}^*$ and multisets of Lyndon words (because the factors in the CFL factorization of a word $w \in \mathfrak{A}^*$ can be stored in a multiset), whereas the Gessel-Reutenauer bijection $\operatorname{GR} : \mathfrak{A}^* \to \mathfrak{MN}^{\mathfrak{a}}$ is a bijection between words in $\mathfrak{A}^*$ and multisets of aperiodic necklaces. Since the Lyndon words are in bijection with the aperiodic necklaces (by Exercise 6.1.34(e)), we can thus view the two bijections as having the same targets (and the same domains). That said, they are not the same bijection.

$w\zeta$ instead of w. This also ensures that $w\zeta$ is an aperiodic word, so the Burrows-Wheeler transformation can be applied to $w\zeta$ even if it cannot be applied to w.)

Kufleitner, in [116, §4], suggests a bijective variant of the Burrows-Wheeler transformation. In our notations, it sends a word $w \in \mathfrak{A}^*$ to the word $\operatorname{RG}\left(\left\{[a_1], [a_2], \ldots, [a_k]\right\}_{\text{multiset}}\right)$, where $(a_1, a_2, \ldots, a_k)$ is the CFL factorization of w.

For variants and generalizations of the Gessel-Reutenauer bijection, see [116], [209], [200], [56] and [179].

6.6.2. *The Gessel-Reutenauer symmetric functions.* In this subsection, we shall study a certain family of symmetric functions. First, we recall that every word $w \in \mathfrak{A}^*$ has a unique CFL factorization (see Theorem 6.1.27). Based on this fact, we can make the following definition:

Definition 6.6.32. For the rest of Subsection 6.6.2, we let $\mathfrak{A}$ be the alphabet $\{1 < 2 < 3 < \cdots\}$.

Let $w \in \mathfrak{A}^*$ be a word. The *CFL type* of w is defined to be the partition whose parts are the positive integers $\ell(a_1), \ell(a_2), \ldots, \ell(a_k)$ (listed in decreasing order), where $(a_1, a_2, \ldots, a_k)$ is the CFL factorization of w. This CFL type is denoted by $\operatorname{CFLtype} w$.

Example 6.6.33. Let w be the word 213212412112. Then, the tuple $(2, 132, 124, 12, 112)$ is the CFL factorization of w. Hence, the CFL type of w is the partition whose parts are the positive integers $\ell(2), \ell(132), \ell(124), \ell(12), \ell(112)$ (listed in decreasing order). In other words, the CFL type of w is the partition $(3, 3, 3, 2, 1)$ (since the positive integers $\ell(2), \ell(132), \ell(124), \ell(12), \ell(112)$ are $1, 3, 3, 2, 3$).

Definition 6.6.34. For each word $w = (w_1, w_2, \ldots, w_n) \in \mathfrak{A}^*$, we define a monomial $\mathbf{x}_w$ in $\mathbf{k}[[\mathbf{x}]]$ by setting $\mathbf{x}_w = x_{w_1} x_{w_2} \cdots x_{w_n}$. (For example, $\mathbf{x}_{(1,3,2,1)} = x_1 x_3 x_2 x_1 = x_1^2 x_2 x_3$.)

For any partition λ, we define a power series $\mathbf{GR}_\lambda \in \mathbf{k}[[\mathbf{x}]]$ by

$$\mathbf{GR}_\lambda = \sum_{\substack{w \in \mathfrak{A}^*; \\ \operatorname{CFLtype} w = \lambda}} \mathbf{x}_w .$$

Example 6.6.35. Let us compute $\mathbf{GR}_{(2,1)}$. Indeed, the words $w \in \mathfrak{A}^*$ satisfying $\operatorname{CFLtype} w = (2,1)$ are the words whose CFL factorization consists of two words, one of which has length 1 and the other has length 2. In other words, these words $w \in \mathfrak{A}^*$ must have the form $w = a_1 a_2$ for two Lyndon words a_1 and a_2 satisfying $a_1 \geq a_2$ and $(\ell(a_1), \ell(a_2)) \in \{(1,2), (2,1)\}$. A straightforward analysis of possibilities reveals that these are precisely the 3-letter words $w = (w_1, w_2, w_3)$ satisfying either $(w_1 < w_2$ and $w_1 \geq w_3)$ or $(w_1 > w_2$ and $w_2 < w_3)$. Hence,

$$\begin{aligned}
\mathbf{GR}_{(2,1)} &= \sum_{\substack{w \in \mathfrak{A}^*; \\ \operatorname{CFLtype} w = (2,1)}} \mathbf{x}_w = \sum_{\substack{w \in \mathfrak{A}^*; \\ w_1 < w_2 \text{ and } w_1 \geq w_3}} \mathbf{x}_w + \sum_{\substack{w \in \mathfrak{A}^*; \\ w_1 > w_2 \text{ and } w_2 < w_3}} \mathbf{x}_w \\
&= \sum_{\substack{w \in \mathfrak{A}^*; \\ w_1 < w_2 \text{ and } w_1 \geq w_3}} \mathbf{x}_w + \sum_{\substack{w \in \mathfrak{A}^*; \\ w_1 > w_2 \text{ and } w_2 < w_3 \text{ and } w_1 \leq w_3}} \mathbf{x}_w + \sum_{\substack{w \in \mathfrak{A}^*; \\ w_1 > w_2 \text{ and } w_2 < w_3 \text{ and } w_1 > w_3}} \mathbf{x}_w \\
&\qquad \text{(here, we have split the second sum according to the relation between } w_1 \text{ and } w_3) \\
&= \sum_{\substack{w \in \mathfrak{A}^*; \\ w_3 \leq w_1 < w_2}} \mathbf{x}_w + \sum_{\substack{w \in \mathfrak{A}^*; \\ w_2 < w_1 \leq w_3}} \mathbf{x}_w + \sum_{\substack{w \in \mathfrak{A}^*; \\ w_2 < w_3 < w_1}} \mathbf{x}_w
\end{aligned}$$

(here, we rewrote the conditions under the summation signs). The three sums on the right hand side are clearly quasisymmetric functions. Using (5.2.3), we can rewrite them as $L_{(2,1)}$, $L_{(1,2)}$ and $L_{(1,1,1)}$, respectively. Thus, we obtain

$$\begin{aligned}
\mathbf{GR}_{(2,1)} &= L_{(2,1)} + L_{(1,2)} + L_{(1,1,1)} = 3M_{(1,1,1)} + M_{(1,2)} + M_{(2,1)} \\
&= 3m_{(1,1,1)} + m_{(2,1)}.
\end{aligned}$$

Thus, $\mathbf{GR}_{(2,1)}$ is actually a symmetric function! We shall soon (in Proposition 6.6.37) see that this is not a coincidence.

We shall now state various properties of the power series $\mathbf{GR}_\lambda$; their proofs are all part of Exercise 6.6.51.

Proposition 6.6.36. *Let n be a positive integer. Then:*

(a) *The partition* (n) *satisfies*

$$\mathbf{GR}_{(n)} = \sum_{\substack{w \in \mathfrak{A}^n; \\ w \text{ is Lyndon}}} \mathbf{x}_w .$$

(b) *Assume that* $\mathbf{k}$ *is a* $\mathbb{Q}$*-algebra. Then,*

$$\mathbf{GR}_{(n)} = \frac{1}{n} \sum_{d \mid n} \mu(d) \, p_d^{n/d} .$$

Here, μ *denotes the number-theoretical Möbius function (defined as in Exercise 2.9.6), and the summation sign "*$\sum_{d \mid n}$*" is understood to range over all* ***positive*** *divisors* d *of* n*.*

Proposition 6.6.37. *Let* λ *be a partition. Then, the power series* $\mathbf{GR}_\lambda$ *belongs to* Λ*.*

Thus, $(\mathbf{GR}_\lambda)_{\lambda \in \operatorname{Par}}$ is a family of symmetric functions.[343] Unlike many other such families we have studied, it is not a basis of Λ; it is not linearly independent (e.g., it satisfies $\mathbf{GR}_{(2,1,1)} = \mathbf{GR}_{(4)}$). Nevertheless, it satisfies a Cauchy-kernel-like identity[344]:

Proposition 6.6.38. *Consider two countable sets of indeterminates* $\mathbf{x} = (x_1, x_2, x_3, \ldots)$ *and* $\mathbf{y} = (y_1, y_2, y_3, \ldots)$*.*

(a) *In the power series ring* $\mathbf{k}[[\mathbf{x}, \mathbf{y}]] = \mathbf{k}[[x_1, x_2, x_3, \ldots, y_1, y_2, y_3, \ldots]]$*, we have*

$$\sum_{\lambda \in \operatorname{Par}} \mathbf{GR}_\lambda(\mathbf{x}) \, p_\lambda(\mathbf{y}) = \sum_{\lambda \in \operatorname{Par}} p_\lambda(\mathbf{x}) \, \mathbf{GR}_\lambda(\mathbf{y}) .$$

(b) *For each word* $w = (w_1, w_2, \ldots, w_n) \in \mathfrak{A}^*$*, we define a monomial* $\mathbf{y}_w$ *in* $\mathbf{k}[[\mathbf{y}]]$ *by setting* $\mathbf{y}_w = y_{w_1} y_{w_2} \cdots y_{w_n}$*. Then,*

$$\sum_{\lambda \in \operatorname{Par}} \mathbf{GR}_\lambda(\mathbf{x}) \, p_\lambda(\mathbf{y}) = \sum_{w \in \mathfrak{A}^*} \mathbf{x}_w p_{\operatorname{CFLtype} w}(\mathbf{y}) = \prod_{w \in \mathfrak{L}} \prod_{u \in \mathfrak{L}} \frac{1}{1 - \mathbf{x}_w^{\ell(u)} \mathbf{y}_u^{\ell(w)}}$$
$$= \sum_{\lambda \in \operatorname{Par}} p_\lambda(\mathbf{x}) \, \mathbf{GR}_\lambda(\mathbf{y}) .$$

The proof of this proposition rests upon the following simple equality[345]:

Proposition 6.6.39. *In the power series ring* $(\mathbf{k}[[\mathbf{x}]])[[t]]$*, we have*

$$\frac{1}{1 - p_1 t} = \prod_{w \in \mathfrak{L}} \frac{1}{1 - \mathbf{x}_w t^{\ell(w)}} .$$

We can furthermore represent the symmetric functions $\mathbf{GR}_\lambda$ in terms of the fundamental basis $(L_\alpha)_{\alpha \in \operatorname{Comp}}$ of QSym; here, the Gessel-Reutenauer bijection from Theorem 6.6.29 reveals its usefulness. We will use Definition 5.3.5.

Proposition 6.6.40. *Let* λ *be a partition. Let* $n = |\lambda|$*. Then,*

$$\mathbf{GR}_\lambda = \sum_{\substack{\sigma \in \mathfrak{S}_n; \\ \sigma \text{ has cycle type } \lambda}} L_{\gamma(\sigma)} .$$

The proof of this relies on Lemma 5.3.6 (see Exercise 6.6.51 below for the details).

Definition 6.6.41. Let $\mathfrak{S} = \bigsqcup_{n \in \mathbb{N}} \mathfrak{S}_n$ (an external disjoint union). For each $\sigma \in \mathfrak{S}$, we let $\operatorname{type} \sigma$ denote the cycle type of σ.

[343] Several sources, including [82], [206, Exercise 7.89] and [66], write L_λ for what we call $\mathbf{GR}_\lambda$. (So would we if L_α didn't already have another meaning here.)

[344] Recall that $\mathfrak{L}$ denotes the set of Lyndon words in $\mathfrak{A}^*$.

[345] Recall that $\mathfrak{L}$ denotes the set of Lyndon words in $\mathfrak{A}^*$. Also, recall that $\mathfrak{A} = \{1 < 2 < 3 < \cdots\}$. Thus, $p_1 = \sum_{i \geq 1} x_i = \sum_{a \in \mathfrak{A}} x_a$.

Proposition 6.6.42. *Consider two countable sets of indeterminates* $\mathbf{x} = (x_1, x_2, x_3, \ldots)$ *and* $\mathbf{y} = (y_1, y_2, y_3, \ldots)$. *In the power series ring* $\mathbf{k}[[\mathbf{x}, \mathbf{y}]]$, *we have*

$$\sum_{\lambda \in \operatorname{Par}} \mathbf{GR}_\lambda(\mathbf{x}) p_\lambda(\mathbf{y}) = \sum_{\lambda \in \operatorname{Par}} p_\lambda(\mathbf{x}) \mathbf{GR}_\lambda(\mathbf{y}) = \sum_{\sigma \in \mathfrak{S}} L_{\gamma(\sigma)}(\mathbf{x}) p_{\operatorname{type}\sigma}(\mathbf{y}).$$

Let us finally give two alternative descriptions of the $\mathbf{GR}_\lambda$ that do not rely on the notion of CFL factorization. First, we state a fact that is essentially trivial:

Proposition 6.6.43. *Let* N *be a necklace. Let* w *and* w' *be two elements of* N. *Then:*

(a) *There exist words* u *and* v *such that* $w = uv$ *and* $w' = vu$.
(b) *We have* $\mathbf{x}_w = \mathbf{x}_{w'}$.

Definition 6.6.44. Let $N \in \mathfrak{N}$ be a necklace. Then, we define a monomial $\mathbf{x}_N$ in $\mathbf{k}[[\mathbf{x}]]$ by setting $\mathbf{x}_N = \mathbf{x}_w$, where w is any element of N. (This is well-defined, because Proposition 6.6.43(b) shows that $\mathbf{x}_w$ does not depend on the choice of w.)

Definition 6.6.45. Let M be a finite multiset of necklaces. Then, we define a monomial $\mathbf{x}_M$ in $\mathbf{k}[[\mathbf{x}]]$ by setting $\mathbf{x}_M = \mathbf{x}_{N_1}\mathbf{x}_{N_2}\cdots\mathbf{x}_{N_k}$, where M is written in the form $M = \{N_1, N_2, \ldots, N_k\}_{\text{multiset}}$.

Definition 6.6.46. Let M be a finite multiset of necklaces. Then, we can obtain a partition by listing the sizes of the necklaces in M in decreasing order. This partition will be called the *type* of M, and will be denoted by $\operatorname{type} M$.

Example 6.6.47. If $M = \{[13], [13], [2], [223]\}_{\text{multiset}}$, then the type of M is $(3, 2, 2, 1)$ (because the sizes of the necklaces in M are $2, 2, 1, 3$).

Proposition 6.6.48. *Let* λ *be a partition. Then,*

$$\mathbf{GR}_\lambda = \sum_{\substack{M \in \mathfrak{MN}^{\mathfrak{a}}; \\ \operatorname{type} M = \lambda}} \mathbf{x}_M.$$

This was our first alternative description of $\mathbf{GR}_\lambda$. Note that it is used as a definition of $\mathbf{GR}_\lambda$ in [82, (2.1)] (where $\mathbf{GR}_\lambda$ is denoted by L_λ). Using the Gessel-Reutenauer bijection, we can restate it as follows:

Proposition 6.6.49. *Let* λ *be a partition. Then,*

$$\mathbf{GR}_\lambda = \sum_{\substack{w \in \mathfrak{A}^*; \\ \operatorname{type}(\operatorname{GR} w) = \lambda}} \mathbf{x}_w.$$

Let us finally give a second alternative description of $\mathbf{GR}_\lambda$:

Proposition 6.6.50. *Let* λ *be a partition. Then,*

$$\mathbf{GR}_\lambda = \sum_{\substack{w \in \mathfrak{A}^*; \\ \operatorname{type}(\operatorname{std} w) = \lambda}} \mathbf{x}_w.$$

Exercise 6.6.51. Prove all statements made in Subsection 6.6.2.

[**Hint:** Here is one way to proceed:

- First prove Proposition 6.6.39, by using the CFL factorization to argue that both sides equal $\sum_{w \in \mathfrak{A}^*} \mathbf{x}_w t^{\ell(w)}$.
- Use a similar argument to derive Proposition 6.6.38 (starting with part (b)).
- Proposition 6.6.43 is almost trivial.
- Derive Proposition 6.6.48 from the definition of $\mathbf{GR}_\lambda$ using the uniqueness of the CFL factorization.
- Derive Proposition 6.6.49 from Proposition 6.6.48 using the bijectivity of GR.
- Derive Proposition 6.6.50 from Proposition 6.6.49.
- Obtain Proposition 6.6.40 by combining Proposition 6.6.50 with Lemma 5.3.6.
- Derive Proposition 6.6.42 from Propositions 6.6.40 and 6.6.38.

- Derive Proposition 6.6.37 either from Proposition 6.6.48 or from Proposition 6.6.38. (In the latter case, make sure to work with $\mathbf{k} = \mathbb{Q}$ first, and then extend to all other $\mathbf{k}$, as the proof will rely on the $\mathbf{k}$-linear independence of $(p_\lambda)_{\lambda \in \operatorname{Par}}$, which doesn't hold for all $\mathbf{k}$.)
- Prove Proposition 6.6.36(a) directly using the definition of $\mathbf{GR}_{(n)}$.
- Show that each positive integer n satisfies $p_1^n = \sum_{d \mid n} d \cdot \mathbf{GR}_{(d)} \left(x_1^{n/d}, x_2^{n/d}, x_3^{n/d}, \ldots \right)$ by taking logarithms in Proposition 6.6.39. Use this and (2.9.7) to prove Proposition 6.6.36(b) recursively.

Other approaches are, of course, possible.]

Remark 6.6.52. Let n be a positive integer. The symmetric function $\mathbf{GR}_{(n)}$ has a few more properties:

(a) It is an $\mathbb{N}$-linear combination of Schur functions. To state the precise rule, we need a few more notations: A *standard tableau* can be defined as a column-strict tableau T with $\operatorname{cont}(T) = (1^m)$, where m is the number of boxes of T. (That is, each of the numbers $1, 2, \ldots, m$ appears exactly once in T, and no other numbers appear.) If T is a standard tableau with m boxes, then a *descent* of T means an $i \in \{1, 2, \ldots, m-1\}$ such that the entry $i+1$ appears in T in a row further down than i does. The *major index* $\operatorname{maj} T$ of a standard tableau T is defined to be the sum of its descents.[346] Now,

$$\mathbf{GR}_{(n)} = \sum_{\lambda \in \operatorname{Par}_n} a_{\lambda,1} s_\lambda,$$

where $a_{\lambda,1}$ is the number of standard tableaux T of shape λ satisfying $\operatorname{maj} T \equiv 1 \bmod n$. (See [206, Exercise 7.89 (c)].)

(b) Assume that $\mathbf{k} = \mathbb{C}$. Recall the map $\operatorname{ch} : A(\mathfrak{S}) \to \Lambda$ from Theorem 4.4.1. Embed the cyclic group $C_n = \mathbb{Z}/n\mathbb{Z}$ as a subgroup in the symmetric group $\mathfrak{S}_n$ by identifying some generator g of C_n with some n-cycle in $\mathfrak{S}_n$. Let ω be a primitive n-th root of unity in $\mathbb{C}$ (for instance, $\exp(2\pi i/n)$). Let $\gamma : C_n \to \mathbb{C}$ be the character of C_n that sends each $g^i \in C_n$ to ω^i. Then,

$$\mathbf{GR}_{(n)} = \operatorname{ch}\left(\operatorname{Ind}_{C_n}^{\mathfrak{S}_n} \gamma \right).$$

(See [206, Exercise 7.89 (b)].)

(c) The character $\operatorname{Ind}_{C_n}^{\mathfrak{S}_n} \gamma$ of $\mathfrak{S}_n$ is actually the character of a representation. To construct it, set $\mathbf{k} = \mathbb{C}$, and recall the notations from Exercise 6.1.41 (while keeping $\mathfrak{A} = \{1, 2, 3, \ldots\}$). Let $\mathfrak{m}_n$ be the $\mathbb{C}$-vector subspace of $T(V)$ spanned by the products $x_{\sigma(1)} x_{\sigma(2)} \cdots x_{\sigma(n)}$ with $\sigma \in \mathfrak{S}_n$. The symmetric group $\mathfrak{S}_n$ acts on $T(V)$ by algebra homomorphisms, with $\sigma \in \mathfrak{S}_n$ sending each x_i to $x_{\sigma(i)}$ when $i \leq n$ and to x_i otherwise. Both $\mathfrak{g}_n$ and $\mathfrak{m}_n$ are $\mathbb{C}\mathfrak{S}_n$-submodules of $T(V)$. Thus, so is the intersection $\mathfrak{g}_n \cap \mathfrak{m}_n$. It is not hard to see that this intersection is spanned by all "nested commutators" $\left[x_{\sigma(1)}, \left[x_{\sigma(2)}, \left[x_{\sigma(3)}, \ldots \right] \right] \right]$ (in $T(V)$) with $\sigma \in \mathfrak{S}_n$. The character of this $\mathbb{C}\mathfrak{S}_n$-module $\mathfrak{g}_n \cap \mathfrak{m}_n$ is precisely the $\operatorname{Ind}_{C_n}^{\mathfrak{S}_n} \gamma$ from Remark 6.6.52(b), so applying the Frobenius characteristic map ch to it yields the symmetric function $\mathbf{GR}_{(n)}$. (See [182, Theorem 9.41(i)]. There are similar ways to obtain $\mathbf{GR}_\lambda$ for all $\lambda \in \operatorname{Par}$.)

Exercise 6.6.53. Prove the claim of Remark 6.6.52(b).

[**Hint:** It helps to recall (or prove) that for any positive integer m, the sum of all primitive m-th roots of unity in $\mathbb{C}$ is $\mu(m)$.]

The symmetric functions $\mathbf{GR}_\lambda$ for more general partitions λ can be expressed in terms of the symmetric functions $\mathbf{GR}_{(n)}$ (which, as we recall from Proposition 6.6.36(b), have a simple expression in terms of the p_m) using the concept of *plethysm*; see [82, Theorem 3.6].

In [82], Gessel and Reutenauer apply the symmetric functions $\mathbf{GR}_\lambda$ to questions of permutation enumeration via the following result[347]:

[346] For example, the tableau

$$\begin{array}{cccc} 1 & 3 & 4 & 8 \\ 2 & 5 & 6 & 9 \\ 7 & & & \end{array}$$

is standard and has descents $1, 4, 6, 8$ and major index $1 + 4 + 6 + 8 = 19$.

[347] Proposition 6.6.54(a) is [82, Corollary 2.2]; Proposition 6.6.54(b) is [82, Theorem 2.1].

Proposition 6.6.54. *Let $n \in \mathbb{N}$. Let $\lambda \in \operatorname{Par}_n$ and $\beta = (\beta_1, \beta_2, \ldots, \beta_k) \in \operatorname{Comp}_n$. We shall use the notations introduced in Definition 5.1.10. Definition 5.3.5 and Definition 6.6.41.*

(a) *Let $\mu \in \operatorname{Par}_n$ be the partition obtained by sorting the entries of β into decreasing order. Then,*

$$\begin{aligned}
&(\text{the number of permutations } \sigma \in \mathfrak{S}_n \text{ satisfying } \operatorname{type}\sigma = \lambda \text{ such that } \beta \text{ refines } \gamma(\sigma)) \\
&= (\text{the number of permutations } \sigma \in \mathfrak{S}_n \text{ satisfying } \operatorname{type}\sigma = \lambda \text{ and } \operatorname{Des}\sigma \subset D(\beta)) \\
&= \left(\text{the coefficient of } x_1^{\beta_1} x_2^{\beta_2} \cdots x_k^{\beta_k} \text{ in } \mathbf{GR}_\lambda\right) = (\text{the coefficient of } \mathbf{x}^\mu \text{ in } \mathbf{GR}_\lambda) \\
&= (\mathbf{GR}_\lambda, h_\mu) \qquad (\text{this is the Hall inner product of } \mathbf{GR}_\lambda \in \Lambda \text{ and } h_\mu \in \Lambda).
\end{aligned}$$

(b) *Recall the ribbon diagram $\operatorname{Rib}(\beta)$ corresponding to the composition β (defined as in Definition 5.1.10). Then,*

$$\begin{aligned}
&(\text{the number of permutations } \sigma \in \mathfrak{S}_n \text{ satisfying } \operatorname{type}\sigma = \lambda \text{ and } \beta = \gamma(\sigma)) \\
&= (\text{the number of permutations } \sigma \in \mathfrak{S}_n \text{ satisfying } \operatorname{type}\sigma = \lambda \text{ and } \operatorname{Des}\sigma = D(\beta)) \\
&= \left(\mathbf{GR}_\lambda, s_{\operatorname{Rib}(\beta)}\right) \qquad \left(\text{this is the Hall inner product of } \mathbf{GR}_\lambda \in \Lambda \text{ and } s_{\operatorname{Rib}(\beta)} \in \Lambda\right).
\end{aligned}$$

Exercise 6.6.55. Prove Proposition 6.6.54.

[**Hint:** Use Proposition 6.6.40, Theorem 5.4.10, the equality (5.4.3) and the adjointness between π and i in Corollary 5.4.3.]

By strategic application of Proposition 6.6.54, Gessel and Reutenauer arrive at several enumerative consequences, such as the following:

- ([82, Theorem 8.3]) If A is a proper subset of $\{1, 2, \ldots, n-1\}$, then

$$\begin{aligned}
&(\text{the number of permutations } \sigma \in \mathfrak{S}_n \text{ satisfying } |\operatorname{Fix}\sigma| = 0 \text{ and } \operatorname{Des}\sigma = A) \\
&= (\text{the number of permutations } \sigma \in \mathfrak{S}_n \text{ satisfying } |\operatorname{Fix}\sigma| = 1 \text{ and } \operatorname{Des}\sigma = A),
\end{aligned}$$

  where $\operatorname{Fix}\sigma$ denotes the set of all fixed points of a permutation σ. This can also be proved bijectively; such a bijective proof can be obtained by combining [50, Theorems 5.1 and 6.1].
- ([82, Theorem 9.4]) If $i \in \{1, 2, \ldots, n-1\}$, then

$$(\text{the number of } n\text{-cycles } \sigma \in \mathfrak{S}_n \text{ satisfying } \operatorname{Des}\sigma = \{i\}) = \sum_{d \mid \gcd(n,i)} \mu(d) \binom{n/d}{i/d}.$$

  Note that this also equals the number of necklaces $[(w_1, w_2, \ldots, w_n)]$ (or, equivalently, Lyndon words $(w_1, w_2, \ldots, w_n)$) with $w_1, w_2, \ldots, w_n \in \{0, 1\}$ and $w_1 + w_2 + \cdots + w_n = i$. This suggests that there should be a bijection between $\{n\text{-cycles } \sigma \in \mathfrak{S}_n \text{ satisfying } \operatorname{Des}\sigma = \{i\}\}$ and the set of such necklaces; and indeed, such a bijection can be found in [45, Theorem 1].

See [82] and [66] for more such applications.

7. Aguiar-Bergeron-Sottile character theory Part I: QSym as a terminal object

It turns out that the universal mapping property of NSym as a free associative algebra leads via duality to a universal property for its dual QSym, elegantly explaining several combinatorial invariants that take the form of quasisymmetric or symmetric functions:

- Ehrenborg's quasisymmetric function of a *ranked poset* [64],
- Stanley's *chromatic* symmetric function of a *graph* [205],
- the quasisymmetric function of a *matroid* considered in [21].

7.1. Characters and the universal property.

Definition 7.1.1. Given a Hopf algebra A over $\mathbf{k}$, a *character* is an algebra morphism $A \xrightarrow{\zeta} \mathbf{k}$, that is,

- $\zeta(1_A) = 1_{\mathbf{k}}$,
- ζ is $\mathbf{k}$-linear, and
- $\zeta(ab) = \zeta(a)\zeta(b)$ for a, b in A.

Example 7.1.2. A particularly important character for $A = \mathrm{QSym}$ is defined as follows:[348]

$$\begin{array}{rcl} \mathrm{QSym} & \xrightarrow{\zeta_Q} & \mathbf{k}, \\ f(\mathbf{x}) & \longmapsto & f(1,0,0,\ldots) = [f(\mathbf{x})]_{x_1=1, x_2=x_3=\cdots=0} . \end{array}$$

Hence,

$$\zeta_Q(M_\alpha) = \zeta_Q(L_\alpha) = \begin{cases} 1, & \text{if } \alpha = (n) \text{ for some } n; \\ 0, & \text{otherwise.} \end{cases}$$

In other words, the restriction $\zeta_Q|_{\mathrm{QSym}_n}$ coincides with the functional H_n in $\mathrm{NSym}_n = \mathrm{Hom}_{\mathbf{k}}(\mathrm{QSym}_n, \mathbf{k})$: one has for f in QSym_n that

$$\zeta_Q(f) = (H_n, f). \tag{7.1.1}$$

It is worth remarking that there is nothing special about setting $x_1 = 1$ and $x_2 = x_3 = \cdots = 0$: for quasisymmetric f, we could have defined the same character ζ_Q by picking any variable, say x_n, and sending

$$f(\mathbf{x}) \longmapsto [f(\mathbf{x})]_{\substack{x_n=1, \text{ and} \\ x_m=0 \text{ for } m\neq n}} .$$

This character $\mathrm{QSym} \xrightarrow{\zeta_Q} \mathbf{k}$ has a certain universal property, known as the *Aguiar-Bergeron-Sottile universality theorem* (part of [4, Theorem 4.1]):

Theorem 7.1.3. *Let A be a connected graded Hopf algebra, and let $A \xrightarrow{\zeta} \mathbf{k}$ be a character. Then, there is a unique graded Hopf morphism $A \xrightarrow{\Psi} \mathrm{QSym}$ making the following diagram commute:*

$$\begin{tikzcd} A \arrow[rr, "\Psi"] \arrow[dr, "\zeta"'] & & \mathrm{QSym} \arrow[dl, "\zeta_Q"] \\ & \mathbf{k} & \end{tikzcd} \tag{7.1.2}$$

Furthermore, Ψ is given by the following formula on homogeneous elements:

$$\Psi(a) = \sum_{\alpha \in \mathrm{Comp}_n} \zeta_\alpha(a) M_\alpha \qquad \text{for all } n \in \mathbb{N} \text{ and } a \in A_n, \tag{7.1.3}$$

where for $\alpha = (\alpha_1, \ldots, \alpha_\ell)$, the map ζ_α is the composite

$$A_n \xrightarrow{\Delta^{(\ell-1)}} A^{\otimes \ell} \xrightarrow{\pi_\alpha} A_{\alpha_1} \otimes \cdots \otimes A_{\alpha_\ell} \xrightarrow{\zeta^{\otimes \ell}} \mathbf{k}$$

in which $A^{\otimes \ell} \xrightarrow{\pi_\alpha} A_{\alpha_1} \otimes \cdots \otimes A_{\alpha_\ell}$ is the canonical projection.

[348] We are using the notation of Proposition 5.1.9 here, and we are still identifying QSym with $\mathrm{QSym}(\mathbf{x})$, where $\mathbf{x}$ denotes the infinite chain $(x_1 < x_2 < \cdots)$.

Proof. One argues that Ψ is unique, and has formula (7.1.3), using only that ζ is $\mathbf{k}$-linear and sends 1 to 1 and that Ψ is a graded $\mathbf{k}$-*coalgebra* map making (7.1.2) commute. Equivalently, consider the adjoint $\mathbf{k}$-*algebra* map[349]

$$\mathrm{NSym} = \mathrm{QSym}^o \xrightarrow{\Psi^*} A^o.$$

Commutativity of (7.1.2) implies that for a in A_n,

$$(\Psi^*(H_n), a) = (H_n, \Psi(a)) \overset{(7.1.1)}{=} \zeta_Q(\Psi(a)) = \zeta(a),$$

whereas gradedness of Ψ^* yields that $(\Psi^*(H_m), a) = 0$ whenever $a \in A_n$ and $m \neq n$. In other words, $\Psi^*(H_n)$ is the element of A^o defined as the following functional on A:

$$\Psi^*(H_n)(a) = \begin{cases} \zeta(a), & \text{if } a \in A_n; \\ 0, & \text{if } a \in A_m \text{ for some } m \neq n. \end{cases} \tag{7.1.4}$$

By the universal property for $\mathrm{NSym} \cong \mathbf{k}\langle H_1, H_2, \ldots\rangle$ as free associative $\mathbf{k}$-algebra, we see that any choice of a $\mathbf{k}$-linear map $A \xrightarrow{\zeta} \mathbf{k}$ uniquely produces a $\mathbf{k}$-algebra morphism $\Psi^* : \mathrm{QSym}^o \to A^o$ which satisfies (7.1.4) for all $n \geq 1$. It is easy to see that this Ψ^* then automatically satisfies (7.1.4) for $n = 0$ as well if ζ sends 1 to 1 (it is here that we use $\zeta(1) = 1$ and the connectedness of A). Hence, any given $\mathbf{k}$-linear map $A \xrightarrow{\zeta} \mathbf{k}$ sending 1 to 1 uniquely produces a $\mathbf{k}$-algebra morphism $\Psi^* : \mathrm{QSym}^o \to A^o$ which satisfies (7.1.4) for all $n \geq 0$. Formula (7.1.3) follows as

$$\Psi(a) = \sum_{\alpha \in \mathrm{Comp}} (H_\alpha, \Psi(a))\ M_\alpha$$

and for a composition $\alpha = (\alpha_1, \ldots, \alpha_\ell)$, one has

$$\begin{aligned} (H_\alpha, \Psi(a)) = (\Psi^*(H_\alpha), a) &= (\Psi^*(H_{\alpha_1}) \cdots \Psi^*(H_{\alpha_\ell}), a) \\ &= \left(\Psi^*(H_{\alpha_1}) \otimes \cdots \otimes \Psi^*(H_{\alpha_\ell}), \Delta^{(\ell-1)}(a)\right) \\ &\overset{(7.1.4)}{=} \left(\zeta^{\otimes \ell} \circ \pi_\alpha\right) \left(\Delta^{(\ell-1)}(a)\right) = \zeta_\alpha(a), \end{aligned}$$

where the definition of ζ_α was used in the last equality.

We wish to show that if, in addition, A is a Hopf algebra and $A \xrightarrow{\zeta} \mathbf{k}$ is a character (i.e., an algebra morphism), then $A \xrightarrow{\Psi} \mathrm{QSym}$ will be an algebra morphism, that is, the two maps $A \otimes A \longrightarrow \mathrm{QSym}$ given by $\Psi \circ m$ and $m \circ (\Psi \otimes \Psi)$ coincide. To see this, consider these two diagrams having the two maps in question as the composites of their top rows:

$$\begin{tikzcd} A \otimes A \arrow[r, "m"] \arrow[dr, "\zeta\otimes\zeta"'] & A \arrow[r, "\Psi"] \arrow[d, "\zeta"'] & \mathrm{QSym} \arrow[dl, "\zeta_Q"] \\ & \mathbf{k} & \end{tikzcd} \qquad \begin{tikzcd} A \otimes A \arrow[r, "\Psi\otimes\Psi"] \arrow[dr, "\zeta\otimes\zeta"'] & \mathrm{QSym}^{\otimes 2} \arrow[r, "m"] \arrow[d, "\zeta_Q\otimes\zeta_Q"'] & \mathrm{QSym} \arrow[dl, "\zeta_Q"] \\ & \mathbf{k} & \end{tikzcd} \tag{7.1.5}$$

The fact that ζ, ζ_Q are algebra morphisms makes the above diagrams commute, so that applying the uniqueness in the first part of the proof to the character $A \otimes A \xrightarrow{\zeta\otimes\zeta} \mathbf{k}$ proves the desired equality $\Psi \circ m = m \circ (\Psi \otimes \Psi)$. □

Remark 7.1.4. When one assumes in addition that A is cocommutative, it follows that the image of Ψ will lie in the subalgebra $\Lambda \subset \mathrm{QSym}$, e.g. from the explicit formula (7.1.3) and the fact that one will have $\zeta_\alpha = \zeta_\beta$ whenever β is a rearrangement of α. In other words, the character $\Lambda \xrightarrow{\zeta_\Lambda} \mathbf{k}$ defined by restricting ζ_Q to Λ, or by

$$\zeta_\Lambda(m_\lambda) = \begin{cases} 1, & \text{if } \lambda = (n) \text{ for some } n; \\ 0, & \text{otherwise,} \end{cases}$$

has a universal property as terminal object with respect to characters on cocommutative Hopf algebras.

[349] Here we are using the fact that there is a 1-to-1 correspondence between graded $\mathbf{k}$-linear maps $A \to \mathrm{QSym}$ and graded $\mathbf{k}$-linear maps $\mathrm{QSym}^o \to A^o$ given by $f \mapsto f^*$, and this correspondence has the property that a given graded map $f : A \to \mathrm{QSym}$ is a $\mathbf{k}$-coalgebra morphism if and only if f^* is a $\mathbf{k}$-algebra morphism. This is a particular case of Exercise 1.6.1(f).

The graded Hopf morphism Ψ in Theorem 7.1.3 will be called the *map* $A \to \operatorname{QSym}$ *induced by the character* ζ.

We close this section by discussing a well-known polynomiality and reciprocity phenomenon; see, e.g., Humpert and Martin [103, Prop. 2.2], Stanley [205, §4].

Definition 7.1.5. The *binomial Hopf algebra* (over the commutative ring $\mathbf{k}$) is the polynomial algebra $\mathbf{k}[m]$ in a single variable m, with a Hopf algebra structure transported from the symmetric algebra $\operatorname{Sym}\left(\mathbf{k}^1\right)$ (which is a Hopf algebra by virtue of Example 1.3.14, applied to $V = \mathbf{k}^1$) along the isomorphism $\operatorname{Sym}\left(\mathbf{k}^1\right) \to \mathbf{k}[m]$ which sends the standard basis element of $\mathbf{k}^1$ to m. Thus the element m is primitive; that is, $\Delta m = 1 \otimes m + m \otimes 1$ and $S(m) = -m$. As S is an algebra anti-endomorphism by Proposition 1.4.10 and $\mathbf{k}[m]$ is commutative, one has $S(g)(m) = g(-m)$ for all polynomials $g(m)$ in $\mathbf{k}[m]$.

Definition 7.1.6. For an element $f(\mathbf{x})$ in QSym and a nonnegative integer m, let $\operatorname{ps}^1(f)(m)$ denote the element of $\mathbf{k}$ obtained by *principal specialization at* $q = 1$

$$\begin{aligned}\operatorname{ps}^1(f)(m) &= [f(\mathbf{x})]_{\substack{x_1=x_2=\cdots=x_m=1,\\ x_{m+1}=x_{m+2}=\cdots=0}} \\ &= f(\underbrace{1,1,\ldots,1}_{m \text{ ones}},0,0,\ldots).\end{aligned}$$

Proposition 7.1.7. *Assume that $\mathbb{Q}$ is a subring of $\mathbf{k}$. The map ps^1 has the following properties.*

(i) *Let $f \in \operatorname{QSym}$. There is a unique polynomial in $\mathbf{k}[m]$ which agrees for each nonnegative integer m with $\operatorname{ps}^1(f)(m)$, and which, by abuse of notation, we will also denote $\operatorname{ps}^1(f)(m)$. If f lies in QSym_n, then $\operatorname{ps}^1(f)(m)$ is a polynomial of degree at most n, taking these values on M_α, L_α for $\alpha = (\alpha_1, \ldots, \alpha_\ell)$ in Comp_n:*

$$\begin{aligned}\operatorname{ps}^1(M_\alpha)(m) &= \binom{m}{\ell},\\ \operatorname{ps}^1(L_\alpha)(m) &= \binom{m-\ell+n}{n}.\end{aligned}$$

(ii) *The map $\operatorname{QSym} \xrightarrow{\operatorname{ps}^1} \mathbf{k}[m]$ is a Hopf morphism into the binomial Hopf algebra.*

(iii) *For all m in $\mathbb{Z}$ and f in QSym one has*

$$\zeta_Q^{\star m}(f) = \operatorname{ps}^1(f)(m).$$

In particular, one also has

$$\zeta_Q^{\star(-m)}(f) = \operatorname{ps}^1(S(f))(m) = \operatorname{ps}^1(f)(-m).$$

(iv) *For a graded Hopf algebra A with a character $A \xrightarrow{\zeta} \mathbf{k}$, and any element a in A_n, the polynomial $\operatorname{ps}^1(\Psi(a))(m)$ in $\mathbf{k}[m]$ has degree at most n, and when specialized to m in $\mathbb{Z}$ satisfies*

$$\zeta^{\star m}(a) = \operatorname{ps}^1(\Psi(a))(m).$$

Proof. To prove assertion (i), note that one has

$$\begin{aligned}\operatorname{ps}^1(M_\alpha)(m) = M_\alpha(1,1,\ldots,1,0,0,\ldots) &= \sum_{1 \le i_1 < \cdots < i_\ell \le m} \left[x_{i_1}^{\alpha_1} \cdots x_{i_\ell}^{\alpha_\ell}\right]_{x_j=1} = \binom{m}{\ell},\\ \operatorname{ps}^1(L_\alpha)(m) = L_\alpha(1,1,\ldots,1,0,0,\ldots) &= \sum_{\substack{1 \le i_1 \le \cdots \le i_n \le m:\\ i_k < i_{k+1} \text{ if } k \in D(\alpha)}} \left[x_{i_1} \cdots x_{i_n}\right]_{x_j=1} \\ &= \left|\{1 \le j_1 \le j_2 \le \cdots \le j_n \le m-\ell+1\}\right| = \binom{m-\ell+n}{n}.\end{aligned}$$

As $\{M_\alpha\}_{\alpha \in \operatorname{Comp}_n}$ form a basis for QSym_n, and $\binom{m}{\ell}$ is a polynomial function in m of degree $\ell (\le n)$, one concludes that for f in QSym_n one has that $\operatorname{ps}^1(f)(m)$ is a polynomial function in m of degree at most n. The polynomial giving rise to this function is unique, since infinitely many of its values are fixed.

To prove assertion (ii), note that ps^1 is an algebra morphism because it is an evaluation homomorphism. To check that it is a coalgebra morphism, it suffices to check $\Delta \circ \mathrm{ps}^1 = (\mathrm{ps}^1 \otimes \mathrm{ps}^1) \circ \Delta$ on each M_α for $\alpha = (\alpha_1, \ldots, \alpha_\ell)$ in Comp_n. Using the Vandermonde summation $\binom{A+B}{\ell} = \sum_k \binom{A}{k}\binom{B}{\ell-k}$, one has

$$(\Delta \circ \mathrm{ps}^1)(M_\alpha) = \Delta \binom{m}{\ell} = \binom{m \otimes 1 + 1 \otimes m}{\ell} = \sum_{k=0}^{\ell} \binom{m \otimes 1}{k} \binom{1 \otimes m}{\ell - k} = \sum_{k=0}^{\ell} \binom{m}{k} \otimes \binom{m}{\ell - k}$$

while at the same time

$$\left((\mathrm{ps}^1 \otimes \mathrm{ps}^1) \circ \Delta\right)(M_\alpha) = \sum_{k=0}^{\ell} \mathrm{ps}^1(M_{(\alpha_1,\ldots,\alpha_k)}) \otimes \mathrm{ps}^1(M_{(\alpha_{k+1},\ldots,\alpha_\ell)}) = \sum_{k=0}^{\ell} \binom{m}{k} \otimes \binom{m}{\ell - k}.$$

Thus ps^1 is a bialgebra morphism, and hence also a Hopf morphism, by Corollary 1.4.27.

For assertion (iii), first assume m lies in $\{0, 1, 2, \ldots\}$. Since $\zeta_Q(f) = f(1, 0, 0, \ldots)$, one has

$$\begin{aligned}
\zeta_Q^{\star m}(f) &= \zeta_Q^{\otimes m} \circ \Delta^{(m-1)} f(\mathbf{x}) = \zeta_Q^{\otimes m}\left(f(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(m)})\right) \\
&= \left[f(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(m)})\right]_{\substack{x_1^{(1)} = x_1^{(2)} = \cdots = x_1^{(m)} = 1, \\ x_2^{(j)} = x_3^{(j)} = \cdots = 0 \text{ for all } j}} \\
&= f(1, 0, 0, \ldots, 1, 0, 0, \ldots, \cdots, 1, 0, 0, \ldots) = f(\underbrace{1, 1, \ldots, 1}_{m \text{ ones}}, 0, 0, \ldots) = \mathrm{ps}^1(f)(m).
\end{aligned}$$

[350] But then Proposition 1.4.26(a) also implies

$$\begin{aligned}
\zeta_Q^{\star(-m)}(f) &= \left(\zeta_Q^{\star(-1)}\right)^{\star m}(f) = (\zeta_Q \circ S)^{\star m}(f) = \zeta_Q^{\star m}(S(f)) \\
&= \mathrm{ps}^1(S(f))(m) = S(\mathrm{ps}^1(f))(m) = \mathrm{ps}^1(f)(-m).
\end{aligned}$$

For assertion (iv), note that

$$\zeta^{\star m}(a) = (\zeta_Q \circ \Psi)^{\star m}(a) = (\zeta_Q^{\star m})(\Psi(a)) = \mathrm{ps}^1(\Psi(a))(m),$$

where the three equalities come from (7.1.2), Proposition 1.4.26(a), and assertion (iii) above, respectively. □

Remark 7.1.8. Aguiar, Bergeron and Sottile give a very cute (third) proof of the QSym antipode formula Theorem 5.1.11, via Theorem 7.1.3, in [4, Example 4.8]. They apply Theorem 7.1.3 to the *coopposite coalgebra* QSym^{cop} and its character $\zeta_Q^{\star(-1)}$. One can show that the map $\mathrm{QSym}^{cop} \xrightarrow{\Psi} \mathrm{QSym}$ induced by $\zeta_Q^{\star(-1)}$ is $\Psi = S$, the antipode of QSym, because $S : \mathrm{QSym} \to \mathrm{QSym}$ is a coalgebra anti-endomorphism (by Exercise 1.4.28) satisfying $\zeta_Q^{\star(-1)} = \zeta_Q \circ S$. They then use the formula (7.1.3) for $\Psi = S$ (together with the polynomiality Proposition 7.1.7) to derive Theorem 5.1.11.

Exercise 7.1.9. Show that $\zeta_Q^{\star m}(f) = \mathrm{ps}^1(f)(m)$ for all $f \in \mathrm{QSym}$ and $m \in \{0, 1, 2, \ldots\}$. (This was already proven in Proposition 7.1.7(iii); give an alternative proof using Proposition 5.1.7.)

7.2. **Example: Ehrenborg's quasisymmetric function of a ranked poset.**

Here we consider incidence algebras, coalgebras and Hopf algebras generally, and then particularize to the case of graded posets, to recover Ehrenborg's interesting quasisymmetric function invariant via Theorem 7.1.3.

7.2.1. *Incidence algebras, coalgebras, Hopf algebras.*

Definition 7.2.1. Given a family $\mathcal{P}$ of finite partially ordered sets P, let $\mathbf{k}[\mathcal{P}]$ denote the free $\mathbf{k}$-module whose basis consists of symbols $[P]$ corresponding to isomorphism classes of posets P in $\mathcal{P}$.

We will assume throughout that each P in $\mathcal{P}$ is *bounded*, that is, it has a unique minimal element $\hat{0} := \hat{0}_P$ and a unique maximal element $\hat{1} := \hat{1}_P$. In particular, $P \neq \varnothing$, although it is allowed that $|P| = 1$, so that $\hat{0} = \hat{1}$; denote this isomorphism class of posets with one element by $[o]$.

If $\mathcal{P}$ is closed under taking intervals

$$[x, y] := [x, y]_P := \{z \in P : x \leq_P z \leq_P y\},$$

[350] See Exercise 7.1.9 for an alternative way to prove this, requiring less thought to verify its soundness.

then one can easily see that the following coproduct and counit endow $\mathbf{k}[\mathcal{P}]$ with the structure of a coalgebra, called the *(reduced) incidence coalgebra*:

$$\Delta[P] := \sum_{x \in P} [\hat{0}, x] \otimes [x, \hat{1}],$$

$$\epsilon[P] := \begin{cases} 1, & \text{if } |P| = 1; \\ 0, & \text{otherwise.} \end{cases}$$

The dual algebra $\mathbf{k}[\mathcal{P}]^*$ is generally called the *reduced incidence algebra (modulo isomorphism)* for the family $\mathcal{P}$ (see, e.g., [192]). It contains the important element $\mathbf{k}[\mathcal{P}] \xrightarrow{\zeta} \mathbf{k}$, called the *$\zeta$-function* that takes the value $\zeta[P] = 1$ for all P.

If $\mathcal{P}$ (is not empty and) satisfies the further property of being *hereditary* in the sense that for every P_1, P_2 in $\mathcal{P}$, the *Cartesian product poset* $P_1 \times P_2$ with componentwise partial order is also in $\mathcal{P}$, then one can check that the following product and unit endow $\mathbf{k}[\mathcal{P}]$ with the structure of a (commutative) algebra:

$$[P_1] \cdot [P_2] := m([P_1] \otimes [P_2]) := [P_1 \times P_2],$$

$$1_{\mathbf{k}[\mathcal{P}]} := [o].$$

Proposition 7.2.2. *For any hereditary family $\mathcal{P}$ of finite posets, $\mathbf{k}[\mathcal{P}]$ is a bialgebra, and even a Hopf algebra with antipode S given as in* (1.4.7) *(Takeuchi's formula):*

$$S[P] = \sum_{k \geq 0} (-1)^k \sum_{\hat{0} = x_0 < \cdots < x_k = \hat{1}} [x_0, x_1] \cdots [x_{k-1}, x_k].$$

Proof. Checking the commutativity of the pentagonal diagram in (1.3.4) amounts to the fact that, for any $(x_1, x_2) <_{P_1 \times P_2} (y_1, y_2)$, one has a poset isomorphism

$$[(x_1, x_2)\ ,\ (y_1, y_2)]_{P_1 \times P_2} \cong [x_1, y_1]_{P_1} \times [x_2, y_2]_{P_2}.$$

Commutativity of the remaining diagrams in (1.3.4) is straightforward, and so $\mathbf{k}[\mathcal{P}]$ is a bialgebra. But then Remark 1.4.25 implies that it is a Hopf algebra, with antipode S as in (1.4.7), because the map $f := \mathrm{id}_{\mathbf{k}[\mathcal{P}]} - u\epsilon$ (sending the class $[o]$ to 0, and fixing all other $[P]$) is locally $\star$-nilpotent:

$$f^{\star k}[P] = \sum_{\hat{0} = x_0 < \cdots < x_k = \hat{1}} [x_0, x_1] \cdots [x_{k-1}, x_k]$$

will vanish due to an empty sum whenever k exceeds the maximum length of a chain in the finite poset P. □

It is perhaps worth remarking how this generalizes the Möbius function formula of P. Hall. Note that the zeta function $\mathbf{k}[\mathcal{P}] \xrightarrow{\zeta} \mathbf{k}$ is a *character*, that is, an algebra morphism. Proposition 1.4.26(a) then tells us that ζ should have a convolutional inverse $\mathbf{k}[\mathcal{P}] \overset{\mu = \zeta^{\star -1}}{\longrightarrow} \mathbf{k}$, traditionally called the *Möbius function*, with the formula $\mu = \zeta^{\star -1} = \zeta \circ S$. Rewriting this via the antipode formula for S given in Proposition 7.2.2 yields P. Hall's formula.

Corollary 7.2.3. *For a finite bounded poset P, one has*

$$\mu[P] = \sum_{k \geq 0} (-1)^k |\{\textit{chains } \hat{0} = x_0 < \cdots < x_k = \hat{1} \textit{ in } P\}|.$$

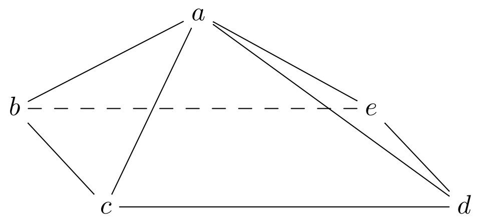
We can also notice that S is an algebra anti-endomorphism (by Proposition 1.4.10), thus an algebra endomorphism (since $\mathbf{k}[\mathcal{P}]$ is commutative, so Exercise 1.5.8(a) shows that the algebra anti-endomorphisms of $\mathbf{k}[\mathcal{P}]$ are the same as the algebra endomorphisms of $\mathbf{k}[\mathcal{P}]$). Hence, $\mu = \zeta \circ S$ is a composition of two algebra homomorphisms, thus an algebra homomorphism itself. We therefore obtain the following classical fact:

Corollary 7.2.4. *For two finite bounded posets P and Q, we have $\mu[P \times Q] = \mu[P] \cdot \mu[Q]$.*

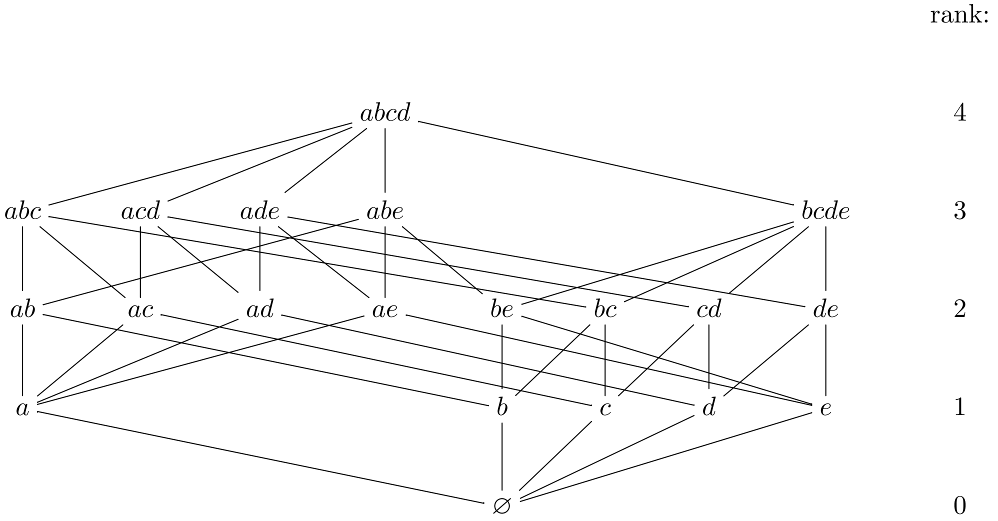
7.2.2. *The incidence Hopf algebras for ranked posets and Ehrenborg's function.*

Definition 7.2.5. Take $\mathcal{P}$ to be the class of bounded *ranked* finite posets P, that is, those for which all maximal chains from $\hat{0}$ to $\hat{1}$ have the same length $r(P)$. This is a hereditary class, as it implies that any interval is $[x, y]_P$ is also ranked, and the product of two bounded ranked posets is also bounded and ranked. It also uniquely defines a *rank function* $P \xrightarrow{r} \mathbb{N}$ in which $r(\hat{0}) = 0$ and $r(x)$ is the length of any maximal chain from $\hat{0}$ to x.

Example 7.2.6. Consider a pyramid with apex vertex a over a square base with vertices b, c, d, e:



Ordering its faces by inclusion gives a bounded ranked poset P, where the rank of an element is one more than the dimension of the face it represents:



Definition 7.2.7. *Ehrenborg's quasisymmetric function* $\Psi[P]$ for a bounded ranked poset P is the image of $[P]$ under the map $\mathbf{k}[\mathcal{P}] \xrightarrow{\Psi} \mathrm{QSym}$ induced by the zeta function $\mathbf{k}[\mathcal{P}] \xrightarrow{\zeta} \mathbf{k}$ as a character, via Theorem 7.1.3.

The quasisymmetric function $\Psi[P]$ captures several interesting combinatorial invariants of P; see Stanley [206, Chap. 3] for more background on these notions.

Definition 7.2.8. Let P be a bounded ranked poset P of rank $r(P) := r(\hat{1})$. Define its *rank-generating function*

$$RGF(P, q) := \sum_{p \in P} q^{r(p)} \in \mathbb{Z}[q],$$

its *characteristic polynomial*

$$\chi(P, q) := \sum_{p \in P} \mu(\hat{0}, p) q^{r(p)} \in \mathbb{Z}[q]$$

(where $\mu(u, v)$ is shorthand for $\mu([u, v])$), and its *zeta polynomial*

$$Z(P, m) = |\{\text{multichains } \hat{0} \leq_P p_1 \leq_P \cdots \leq_P p_{m-1} \leq_P \hat{1}\}| \tag{7.2.1}$$

$$= \sum_{s=0}^{r(P)-1} \binom{m}{s+1} |\{\text{chains } \hat{0} < p_1 < \cdots < p_s < \hat{1}\}| \in \mathbb{Q}[m] \tag{7.2.2}$$

[351]. Also, for each subset $S \subset \{1, 2, \ldots, r(P) - 1\}$, define the *flag number* f_S of P by

$$f_S = |\{\text{chains } \hat{0} <_P p_1 <_P \cdots <_P p_s <_P \hat{1} \text{ with } \{r(p_1), \ldots, r(p_s)\} = S\}|.$$

These flag numbers are the components of the *flag f-vector* $(f_S)_{S \subset [r-1]}$ of P. Further define the *flag h-vector* $(h_T)_{T \subset [r-1]}$ of P, whose entries h_T are given by $f_S = \sum_{T \subset S} h_T$, or, equivalently[352], by $h_S = \sum_{T \subset S} (-1)^{|S \setminus T|} f_T$.

Example 7.2.9. For the poset P in Example 7.2.6, one has $RGF(P, q) = 1 + 5q + 8q^2 + 5q^3 + q^4$. Since P is the poset of faces of a polytope, the Möbius function values for its intervals are easily predicted: $\mu(x, y) = (-1)^{r[x,y]}$, that is, P is an *Eulerian ranked poset*; see Stanley [206, §3.16]. Hence its characteristic polynomial is trivially related to the rank generating function, sending $q \mapsto -q$, that is,

$$\chi(P, q) = RGF(P, -q) = 1 - 5q + 8q^2 - 5q^3 + q^4.$$

Its flag f-vector and h-vector entries are given in the following table.

S	f_S		h_S
$\varnothing$	1		1
$\{1\}$	5	$5 - 1 =$	4
$\{2\}$	8	$8 - 1 =$	7
$\{3\}$	5	$5 - 1 =$	4
$\{1, 2\}$	16	$16 - (5 + 8) + 1 =$	4
$\{1, 3\}$	16	$16 - (5 + 5) + 1 =$	7
$\{2, 3\}$	16	$16 - (5 + 8) + 1 =$	4
$\{1, 2, 3\}$	32	$32 - (16 + 16 + 16) + (5 + 8 + 5) - 1 =$	1

and using (7.2.2), its zeta polynomial is

$$Z(P, m) = 1\binom{m}{1} + (5 + 8 + 5)\binom{m}{2} + (16 + 16 + 16)\binom{m}{3} + 32\binom{m}{4} = \frac{m^2(2m-1)(2m+1)}{3}.$$

Theorem 7.2.10. *Assume that $\mathbb{Q}$ is a subring of $\mathbf{k}$. Ehrenborg's quasisymmetric function $\Psi[P]$ for a bounded ranked poset P encodes*

(i) *the flag f-vector entries f_S and flag h-vector entries h_S as its M_α and L_α expansion coefficients*[353] *:*

$$\Psi[P] = \sum_\alpha f_{D(\alpha)}(P)\ M_\alpha = \sum_\alpha h_{D(\alpha)}(P)\ L_\alpha,$$

(ii) *the zeta polynomial as the specialization from Definition 7.1.6*

$$Z(P, m) = \mathrm{ps}^1(\Psi[P])(m) = [\Psi[P]]_{\substack{x_1 = x_2 = \cdots = x_m = 1, \\ x_{m+1} = x_{m+2} = \cdots = 0}},$$

(iii) *the rank-generating function as the specialization*

$$RGF(P, q) = [\Psi[P]]_{\substack{x_1 = q, x_2 = 1, \\ x_3 = x_4 = \cdots = 0}},$$

(iv) *the characteristic polynomial as the convolution*

$$\chi(P, q) = ((\psi_q \circ S) \star \zeta_Q) \circ \Psi[P],$$

where $\mathrm{QSym} \xrightarrow{\psi_q} \mathbf{k}[q]$ *maps* $f(\mathbf{x}) \longmapsto f(q, 0, 0, \ldots)$.

Proof. In assertion (i), the expansion $\Psi[P] = \sum_\alpha f_{D(\alpha)}(P)\, M_\alpha$ is (7.1.3), since $\zeta_\alpha[P] = f_{D(\alpha)}(P)$. The L_α expansion follows from this, as $L_\alpha = \sum_{\beta : D(\beta) \supset D(\alpha)} M_\beta$ and $f_S(P) = \sum_{T \subset S} h_T$.

Assertion (ii) is immediate from Proposition 7.1.7(iv), since $Z(P, m) = \zeta^{\star m}[P]$.

[351] Actually, (7.2.2) is false if $|P| = 1$ (but only then). We use (7.2.1) to define $Z(P, m)$ in this case.

[352] The equivalence follows from inclusion-exclusion (more specifically, from the converse of Lemma 5.2.6(a)).

[353] In fact, Ehrenborg *defined* $\Psi[P]$ in [64, Defn. 4.1] via this M_α expansion, and then showed that it gave a Hopf morphism.

Assertion (iii) can be deduced from assertion (i), but it is perhaps more fun and in the spirit of things to proceed as follows. Note that $\psi_q(M_\alpha) = q^n$ for $\alpha = (n)$, and $\psi_q(M_\alpha)$ vanishes for all other $\alpha \neq (n)$ in Comp_n. Hence for a bounded ranked poset P one has

$$(\psi_q \circ \Psi)[P] = q^{r(P)}. \tag{7.2.3}$$

But if we treat $\zeta_Q : \mathrm{QSym} \to \mathbf{k}$ as a map $\mathrm{QSym} \to \mathbf{k}[q]$, then (1.4.2) (applied to $\mathbf{k}[\mathcal{P}]$, QSym, $\mathbf{k}[q]$, $\mathbf{k}[q]$, Ψ, $\mathrm{id}_{\mathbf{k}[q]}$, ψ_q and ζ_Q instead of C, C', A, A', γ, α, f and g) shows that

$$(\psi_q \star \zeta_Q) \circ \Psi = (\psi_q \circ \Psi) \star (\zeta_Q \circ \Psi), \tag{7.2.4}$$

since $\Psi : \mathbf{k}[\mathcal{P}] \to \mathrm{QSym}$ is a $\mathbf{k}$-coalgebra homomorphism. Consequently, one can compute

$$\begin{aligned} RGF(P,q) &= \sum_{p \in P} q^{r(p)} \cdot 1 = \sum_{p \in P} q^{r([\hat{0},p])} \cdot \zeta[p,\hat{1}] \overset{\substack{(7.2.3),\\(7.1.2)}}{=} \sum_{p \in P} (\psi_q \circ \Psi)[\hat{0},p] \cdot (\zeta_Q \circ \Psi)[p,\hat{1}] \\ &= \left((\psi_q \circ \Psi) \star (\zeta_Q \circ \Psi) \right)[P] \overset{(7.2.4)}{=} (\psi_q \star \zeta_Q)(\Psi[P]) = (\psi_q \otimes \zeta_Q)(\Delta \Psi[P]) \\ &= [\Psi[P](\mathbf{x},\mathbf{y})]_{\substack{x_1=q, x_2=x_3=\cdots=0 \\ y_1=1, y_2=y_3=\cdots=0}} = [\Psi[P](\mathbf{x})]_{\substack{x_1=q, x_2=1, \\ x_3=x_4=\cdots=0}}. \end{aligned}$$

Similarly, for assertion (iv) first note that

$$((\psi_q \circ S) \star \zeta_Q) \circ \Psi = (\psi_q \circ S \circ \Psi) \star (\zeta_Q \circ \Psi), \tag{7.2.5}$$

(this is proven similarly to (7.2.4), but now using the map $\psi_q \circ S$ instead of ψ_q). Now, Proposition 7.2.2 and Corollary 7.2.3 let one calculate that

$$\begin{aligned} (\psi_q \circ \Psi \circ S)[P] &= \sum_k (-1)^k \sum_{\hat{0}=x_0<\cdots<x_k=\hat{1}} (\psi_q \circ \Psi)([x_0,x_1]) \cdots (\psi_q \circ \Psi)([x_{k-1},x_k]) \\ &\overset{(7.2.3)}{=} \sum_k (-1)^k \sum_{\hat{0}=x_0<\cdots<x_k=\hat{1}} q^{r(P)} = \mu(\hat{0},\hat{1}) q^{r(P)}. \end{aligned}$$

This is used in the penultimate equality here:

$$\begin{aligned} ((\psi_q \circ S) \star \zeta_Q) \circ \Psi[P] &\overset{(7.2.5)}{=} ((\psi_q \circ S \circ \Psi) \star (\zeta_Q \circ \Psi))[P] = ((\psi_q \circ \Psi \circ S) \star \zeta)[P] \\ &= \sum_{p \in P} (\psi_q \circ \Psi \circ S)[\hat{0},p] \cdot \zeta[p,\hat{1}] = \sum_{p \in P} \mu[\hat{0},p] q^{r(p)} = \chi(P,q). \end{aligned}$$

□

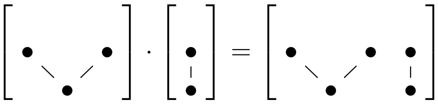
7.3. **Example: Stanley's chromatic symmetric function of a graph.** We introduce the *chromatic Hopf algebra of graphs* and an associated character ζ so that the map Ψ from Theorem 7.1.3 sends a graph G to Stanley's *chromatic symmetric function* of G. Then principal specialization ps^1 sends this to the *chromatic polynomial* of the graph.

7.3.1. *The chromatic Hopf algebra of graphs.*

Definition 7.3.1. The *chromatic Hopf algebra* (see Schmitt [194, §3.2]) $\mathcal{G}$ is a free $\mathbf{k}$-module whose $\mathbf{k}$-basis elements $[G]$ are indexed by isomorphism classes of (finite) simple graphs $G = (V, E)$. Define for $G_1 = (V_1, E_1), G_2 = (V_2, E_2)$ the multiplication

$$[G_1] \cdot [G_2] := [G_1 \sqcup G_2]$$

where $[G_1 \sqcup G_2]$ denote the isomorphism class of the disjoint union, on vertex set $V = V_1 \sqcup V_2$ which is a disjoint union of copies of their vertex sets V_1, V_2, with edge set $E = E_1 \sqcup E_2$. For example,



Thus the class $[\varnothing]$ of the empty graph $\varnothing$ having $V = \varnothing, E = \varnothing$ is a unit element.

Given a graph $G = (V, E)$ and a subset $V' \subset V$, the *subgraph induced on vertex set* V' is defined as the graph $G|_{V'} := (V', E')$ with edge set $E' = \{e \in E : e = \{v_1, v_2\} \subset V'\}$. This lets one define a comultiplication $\Delta : \mathcal{G} \to \mathcal{G} \otimes \mathcal{G}$ by setting

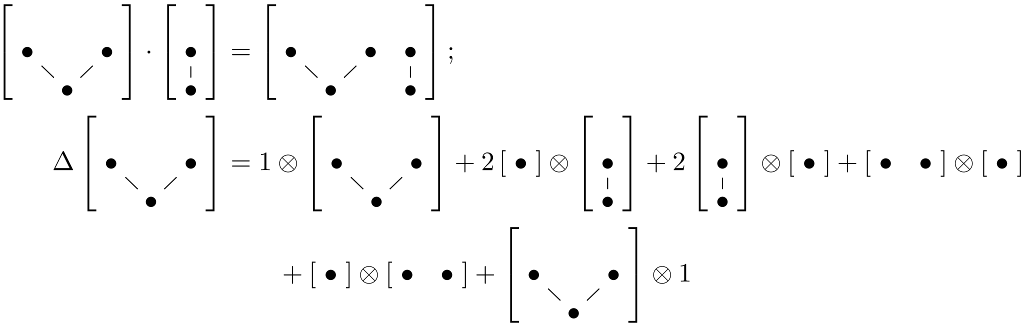
$$\Delta[G] := \sum_{(V_1,V_2):V_1 \sqcup V_2 = V} [G|_{V_1}] \otimes [G|_{V_2}].$$

Define a counit $\epsilon : \mathcal{G} \to \mathbf{k}$ by

$$\epsilon[G] := \begin{cases} 1, & \text{if } G = \varnothing; \\ 0, & \text{otherwise.} \end{cases}$$

Proposition 7.3.2. *The above maps endow $\mathcal{G}$ with the structure of a connected graded finite type Hopf algebra over* $\mathbf{k}$, *which is both commutative and cocommutative.*

Example 7.3.3. Here are some examples of these structure maps:



Proof of Proposition 7.3.2. The associativity of the multiplication and comultiplication should be clear as

$$m^{(2)}([G_1] \otimes [G_2] \otimes [G_3]) = [G_1 \sqcup G_2 \sqcup G_3],$$

$$\Delta^{(2)}[G] = \sum_{\substack{(V_1,V_2,V_3): \\ V = V_1 \sqcup V_2 \sqcup V_3}} [G|_{V_1}] \otimes [G|_{V_2}] \otimes [G|_{V_3}].$$

Checking the unit and counit conditions are straightforward. Commutativity of the pentagonal bialgebra diagram in (1.3.4) comes down to check that, given graphs G_1, G_2 on disjoint vertex sets V_1, V_2 , when one applies to $[G_1] \otimes [G_2]$ either the composite $\Delta \circ m$ or the composite $(m \otimes m) \circ (\mathrm{id} \otimes T \otimes \mathrm{id}) \circ (\Delta \otimes \Delta)$, the result is the same:

$$\sum_{\substack{(V_{11},V_{12},V_{21},V_{22}): \\ V_1 = V_{11} \sqcup V_{12} \\ V_2 = V_{21} \sqcup V_{22}}} [G_1|_{V_{11}} \sqcup G_2|_{V_{21}}] \otimes [G_1|_{V_{12}} \sqcup G_2|_{V_{22}}].$$

Letting $\mathcal{G}_n$ be the $\mathbf{k}$-span of $[G]$ having n vertices makes $\mathcal{G}$ a bialgebra which is graded and connected, and hence also a Hopf algebra by Proposition 1.4.16. Cocommutativity should be clear, and commutativity follows from the graph isomorphism $G_1 \sqcup G_2 \cong G_2 \sqcup G_1$. Finally, $\mathcal{G}$ is of finite type since there are only finitely many isomorphism classes of simple graphs on n vertices for every given n. □

Remark 7.3.4. Humpert and Martin [103, Theorem 3.1] gave the following expansion for the antipode in the chromatic Hopf algebra, containing fewer terms than Takeuchi's general formula (1.4.7): given a graph $G = (V, E)$, one has

$$S[G] = \sum_F (-1)^{|V| - \mathrm{rank}(F)} \operatorname{acyc}(G/F)[G_{V,F}]. \tag{7.3.1}$$

Here F runs over all subsets of edges that form *flats* in the graphic matroid for G, meaning that if $e = \{v, v'\}$ is an edge in E for which one has a path of edges in F connecting v to v', then e also lies in F. Here G/F denotes the quotient graph in which all of the edges of F have been *contracted*, while $\operatorname{acyc}(G/F)$ denotes its number of *acyclic orientations*, and $G_{V,F} := (V, F)$ as a simple graph.[354]

[354] The notation $\mathrm{rank}(F)$ denotes the *rank* of F in the graphic matroid of G. We can define it without reference to matroid theory as the maximum cardinality of a subset F' of F such that the graph $G_{V,F'}$ is acyclic. Equivalently, $\mathrm{rank}(F)$ is $|V| - c(F)$,

Remark 7.3.5. In [14], Benedetti, Hallam and Machacek define a Hopf algebra of simplicial complexes, which contains $\mathcal{G}$ as a Hopf subalgebra (and also has $\mathcal{G}$ as a quotient Hopf algebra). They compute a formula for its antipode similar to (and generalizing) (7.3.1).

Remark 7.3.6. The chromatic Hopf algebra $\mathcal{G}$ is used in [122] and [39, §14.4] to study *Vassiliev invariants of knots*. In fact, a certain quotient of $\mathcal{G}$ (named $\mathcal{F}$ in [122] and $\mathcal{L}$ in [39, §14.4]) is shown to naturally host invariants of *chord diagrams* and therefore Vassiliev invariants of knots.

Remark 7.3.7. The $\mathbf{k}$-algebra $\mathcal{G}$ is isomorphic to a polynomial algebra (in infinitely many indeterminates) over $\mathbf{k}$. Indeed, every finite graph can be uniquely written as a disjoint union of finitely many connected finite graphs (up to order). Therefore, the basis elements $[G]$ of $\mathcal{G}$ corresponding to connected finite graphs G are algebraically independent in $\mathcal{G}$ and generate the whole $\mathbf{k}$-algebra $\mathcal{G}$ (indeed, the disjoint unions of connected finite graphs are precisely the monomials in these elements). Thus, $\mathcal{G}$ is isomorphic to a polynomial $\mathbf{k}$-algebra with countably many generators (one for each isomorphism class of connected finite graphs). As a consequence, for example, we see that $\mathcal{G}$ is an integral domain if $\mathbf{k}$ is an integral domain.

7.3.2. *A "ribbon basis" for $\mathcal{G}$ and self-duality.* In this subsection, we shall explore a second basis of $\mathcal{G}$ and a bilinear form on $\mathcal{G}$. This material will not be used in the rest of these notes (except in Exercise 7.3.25), but it is of some interest and provides an example of how a commutative cocommutative Hopf algebra can be studied.

First, let us define a second basis of $\mathcal{G}$, which is obtained by Möbius inversion (in an appropriate sense) from the standard basis $([G])_{[G] \text{ is an isomorphism class of finite graphs}}$:

Definition 7.3.8. For every finite graph $G = (V, E)$, set

$$[G]^\sharp = \sum_{\substack{H=(V,E'); \\ E' \supset E^c}} (-1)^{\left|E' \setminus E^c\right|} [H] \in \mathcal{G},$$

where E^c denotes the complement of the subset E in the set of all two-element subsets of V. Clearly, $[G]^\sharp$ depends only on the isomorphism class $[G]$ of G, not on G itself.

Proposition 7.3.9. (a) *Every finite graph $G = (V, E)$ satisfies*

$$[G] = \sum_{\substack{H=(V,E'); \\ E' \cap E = \varnothing}} [H]^\sharp .$$

(b) *The elements $[G]^\sharp$, where $[G]$ ranges over all isomorphism classes of finite graphs, form a basis of the $\mathbf{k}$-module $\mathcal{G}$.*

(c) *For any graph $H = (V, E)$, we have*

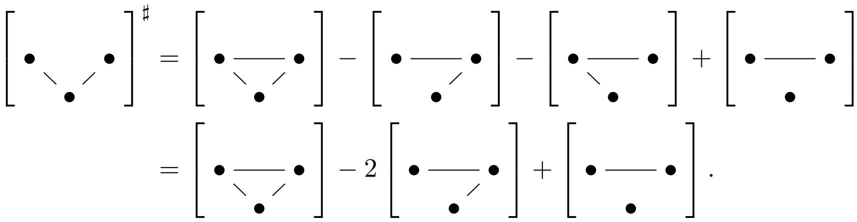
$$\Delta [H]^\sharp = \sum_{\substack{(V_1,V_2); \\ V=V_1 \sqcup V_2; \\ H=H|_{V_1} \sqcup H|_{V_2}}} [H|_{V_1}]^\sharp \otimes [H|_{V_2}]^\sharp . \tag{7.3.2}$$

(d) *For any two graphs $H_1 = (V_1, E_1)$ and $H_2 = (V_2, E_2)$, we have*

$$[H_1]^\sharp [H_2]^\sharp = \sum_{\substack{H=(V_1 \sqcup V_2, E); \\ H|_{V_1}=H_1; \\ H|_{V_2}=H_2}} [H]^\sharp . \tag{7.3.3}$$

where $c(F)$ denotes the number of connected components of the graph $G_{V,F}$. Thus, the equality (7.3.1) can be rewritten as $S[G] = \sum_F (-1)^{c(F)} \operatorname{acyc}(G/F)[G_{V,F}]$. In this form, this equality is also proven in [15, Thm. 7.1].

For example,



Proving Proposition 7.3.9 is part of Exercise 7.3.14 further below.

The equalities that express the elements $[G]^\sharp$ in terms of the elements $[H]$ (as in Definition 7.3.8), and vice versa (Proposition 7.3.9(a)), are reminiscent of the relations (5.4.10) and (5.4.9) between the bases (R_α) and (H_α) of NSym. In this sense, we can call the basis of $\mathcal{G}$ formed by the $[G]^\sharp$ a "ribbon basis" of $\mathcal{G}$.

We now define a $\mathbf{k}$-bilinear form on $\mathcal{G}$:

Definition 7.3.10. For any two graphs G and H, let $\operatorname{Iso}(G,H)$ denote the set of all isomorphisms from G to H [355]. Let us now define a $\mathbf{k}$-bilinear form $(\cdot,\cdot) : \mathcal{G} \times \mathcal{G} \to \mathbf{k}$ on $\mathcal{G}$ by setting

$$\left([G]^\sharp, [H]\right) = \left|\operatorname{Iso}(G,H)\right|.$$

[356]

Proposition 7.3.11. *The form $(\cdot,\cdot) : \mathcal{G} \times \mathcal{G} \to \mathbf{k}$ is symmetric.*

Again, we refer to Exercise 7.3.14 for a proof of Proposition 7.3.11.

The basis of $\mathcal{G}$ constructed in Proposition 7.3.9(b) and the bilinear form $(\cdot,\cdot)$ defined in Definition 7.3.10 can be used to construct a Hopf algebra homomorphism from $\mathcal{G}$ to its graded dual $\mathcal{G}^o$:

Definition 7.3.12. For any finite graph G, let $\operatorname{aut}(G)$ denote the number $\left|\operatorname{Iso}(G,G)\right|$. Notice that this is a positive integer, since the set $\operatorname{Iso}(G,G)$ is nonempty (it contains id_G).

Now, recall that the Hopf algebra $\mathcal{G}$ is a connected graded Hopf algebra of finite type. The n-th homogeneous component is spanned by the $[G]$ where G ranges over the graphs with n vertices. Since $\mathcal{G}$ is of finite type, its graded dual $\mathcal{G}^o$ is defined. Let $\left([G]^*\right)_{[G] \text{ is an isomorphism class of finite graphs}}$ be the basis of $\mathcal{G}^o$ dual to the basis $([G])_{[G] \text{ is an isomorphism class of finite graphs}}$ of $\mathcal{G}$. Define a $\mathbf{k}$-linear map $\psi : \mathcal{G} \to \mathcal{G}^o$ by

$$\psi\left([G]^\sharp\right) = \operatorname{aut}(G) \cdot [G]^* \qquad \text{for every finite graph } G.$$

[357]

Proposition 7.3.13. *Consider the map $\psi : \mathcal{G} \to \mathcal{G}^o$ defined in Definition 7.3.12.*

(a) *This map ψ satisfies $(\psi(a))(b) = (a,b)$ for all $a \in \mathcal{G}$ and $b \in \mathcal{G}$.*
(b) *The map $\psi : \mathcal{G} \to \mathcal{G}^o$ is a Hopf algebra homomorphism.*
(c) *If $\mathbb{Q}$ is a subring of $\mathbf{k}$, then the map ψ is a Hopf algebra isomorphism $\mathcal{G} \to \mathcal{G}^o$.*

Exercise 7.3.14. Prove Proposition 7.3.9, Proposition 7.3.11 and Proposition 7.3.13.

[355] We recall that if $G = (V,E)$ and $H = (W,F)$ are two graphs, then an *isomorphism* from G to H means a bijection $\varphi : V \to W$ such that $\varphi_*(E) = F$. Here, φ_* denotes the map from the powerset of V to the powerset of W which sends every $T \subset V$ to $\varphi(T) \subset W$.

[356] This is well-defined, because:

- the number $\left|\operatorname{Iso}(G,H)\right|$ depends only on the isomorphism classes $[G]$ and $[H]$ of G and H, but not on G and H themselves;
- the elements $[G]^\sharp$, where $[G]$ ranges over all isomorphism classes of finite graphs, form a basis of the $\mathbf{k}$-module $\mathcal{G}$ (because of Proposition 7.3.9(b));
- the elements $[G]$, where $[G]$ ranges over all isomorphism classes of finite graphs, form a basis of the $\mathbf{k}$-module $\mathcal{G}$.

[357] This is well-defined, since $\left([G]^\sharp\right)_{[G] \text{ is an isomorphism class of finite graphs}}$ is a basis of the $\mathbf{k}$-module $\mathcal{G}$ (because of Proposition 7.3.9(b)).

Remark 7.3.15. Proposition 7.3.13(c) shows that the Hopf algebra $\mathcal{G}$ is self-dual when $\mathbb{Q}$ is a subring of $\mathbf{k}$. On the other hand, if $\mathbf{k}$ is a field of positive characteristic, then $\mathcal{G}$ is never self-dual. Here is a quick way to see this: The elements $[G]^*$ of $\mathcal{G}^o$ defined in Definition 7.3.12 have the property that

$$\left([\circ]^*\right)^n = n! \cdot \sum_{\substack{[G] \text{ is an isomorphism} \\ \text{class of finite graphs on} \\ n \text{ vertices}}} [G]^*$$

for every $n \in \mathbb{N}$, where $\circ$ denotes the graph with one vertex.[358] Thus, if p is a prime and $\mathbf{k}$ is a field of characteristic p, then $\left([\circ]^*\right)^p = 0$. Hence, the $\mathbf{k}$-algebra $\mathcal{G}^o$ has nilpotents in this situation. However, the $\mathbf{k}$-algebra $\mathcal{G}$ does not (indeed, Remark 7.3.7 shows that it is an integral domain whenever $\mathbf{k}$ is an integral domain). Thus, when $\mathbf{k}$ is a field of characteristic p, then $\mathcal{G}$ and $\mathcal{G}^o$ are not isomorphic as $\mathbf{k}$-algebras (let alone as Hopf algebras).

7.3.3. *Stanley's chromatic symmetric function of a graph.*

Definition 7.3.16. *Stanley's chromatic symmetric function* $\Psi[G]$ for a simple graph $G = (V, E)$ is the image of $[G]$ under the map $\mathcal{G} \xrightarrow{\Psi} \mathrm{QSym}$ induced via Theorem 7.1.3 from the *edge-free character* $\mathcal{G} \xrightarrow{\zeta} \mathbf{k}$ defined by

$$\zeta[G] = \begin{cases} 1, & \text{if } G \text{ has no edges, that is, } G \text{ is an independent/stable set of vertices;} \\ 0, & \text{otherwise.} \end{cases} \tag{7.3.4}$$

Note that, because $\mathcal{G}$ is cocommutative, $\Psi[G]$ is symmetric and not just quasisymmetric; see Remark 7.1.4.

Recall that for a graph $G = (V, E)$, a (vertex-)coloring $f : V \to \{1, 2, \ldots\}$ is called *proper* if no edge $e = \{v, v'\}$ in E has $f(v) = f(v')$.

Proposition 7.3.17. *For a graph $G = (V, E)$, the symmetric function $\Psi[G]$ has the expansion* [359]

$$\Psi[G] = \sum_{\substack{\text{proper colorings} \\ f : V \to \{1,2,\ldots\}}} \mathbf{x}_f$$

where $\mathbf{x}_f := \prod_{v \in V} x_{f(v)}$. In particular, its specialization from Proposition 7.1.6 gives the chromatic polynomial of G:

$$\mathrm{ps}^1 \Psi[G](m) = \chi_G(m) = \left|\{\text{proper colorings } f : V \to \{1, 2, \ldots, m\}\}\right|.$$

Proof. The iterated coproduct $\mathcal{G} \xrightarrow{\Delta^{(\ell-1)}} \mathcal{G}^{\otimes \ell}$ sends

$$[G] \longmapsto \sum_{\substack{(V_1,\ldots,V_\ell): \\ V = V_1 \sqcup \cdots \sqcup V_\ell}} [G|_{V_1}] \otimes \cdots \otimes [G|_{V_\ell}]$$

and the map $\zeta^{\otimes \ell}$ sends each addend on the right to 1 or 0, depending upon whether each $V_i \subset V$ is a stable set or not, that is, whether the assignment of color i to the vertices in V_i gives a proper coloring of G. Thus formula (7.1.3) shows that the coefficient ζ_α of $x_1^{\alpha_1} \cdots x_\ell^{\alpha_\ell}$ in $\Psi[G]$ counts the proper colorings f in which $|f^{-1}(i)| = \alpha_i$ for each i. □

Example 7.3.18. For the complete graph K_n on n vertices, one has

$$\Psi[K_n] = n! e_n, \qquad \text{thus}$$

$$\begin{aligned} \mathrm{ps}^1(\Psi[K_n])(m) &= n! e_n(\underbrace{1, 1, \ldots, 1}_{m \text{ ones}}) = n! \binom{m}{n} \\ &= m(m-1)\cdots(m-(n-1)) = \chi_{K_n}(m). \end{aligned}$$

In particular, the single vertex graph K_1 has $\Psi[K_1] = e_1$, and since the Hopf morphism Ψ is in particular an algebra morphism, a graph $K_1^{\sqcup n}$ having n isolated vertices and no edges will have $\Psi[K_1^{\sqcup n}] = e_1^n$.

[358] To see this, observe that the tensor $[\circ]^{\otimes n}$ appears in the iterated coproduct $\Delta^{(n-1)}([G])$ exactly $n!$ times whenever G is a graph on n vertices.

[359] In fact, Stanley *defined* $\Psi[G]$ in [205, Defn. 2.1] via this expansion.

As a slightly more interesting example, the graph P_3 which is a path having three vertices and two edges will have

$$\Psi[P_3] = m_{(2,1)} + 6m_{(1,1,1)} = e_2 e_1 + 3e_3.$$

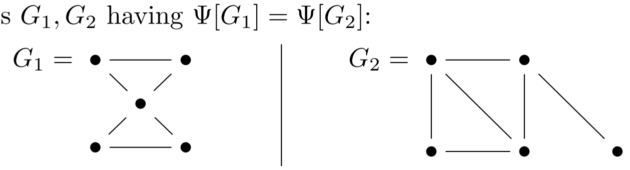
One might wonder, based on the previous examples, when $\Psi[G]$ is *e-positive*, that is, when does its unique expansion in the $\{e_\lambda\}$ basis for Λ have nonnegative coefficients? This is an even stronger assertion than *s-positivity*, that is, having nonnegative coefficients for the expansion in terms of Schur functions $\{s_\lambda\}$, since each e_λ is s-positive. This weaker property fails, starting with the *claw graph* $K_{3,1}$, which has

$$\Psi[K_{3,1}] = s_{(3,1)} - s_{(2,2)} + 5s_{(2,1,1)} + 8s_{(1,1,1,1)}.$$

On the other hand, a result of Gasharov [75, Theorem 2] shows that one at least has s-positivity for $\Psi[\operatorname{inc}(P)]$ where $\operatorname{inc}(P)$ is the *incomparability graph* of a poset which is $(\mathbf{3}+\mathbf{1})$-free; we refer the reader to Stanley [205, §5] for a discussion of the following conjecture, which remains open[360]:

Conjecture 7.3.19. *For any $(\mathbf{3}+\mathbf{1})$-free poset P, the incomparability graph* $\operatorname{inc}(P)$ *has* $\Psi[\operatorname{inc}(P)]$ *an e-positive symmetric function.*

Here is another question about $\Psi[G]$: how well does it distinguish nonisomorphic graphs? Stanley gave this example of two graphs G_1, G_2 having $\Psi[G_1] = \Psi[G_2]$:



At least $\Psi[G]$ appears to do better at distinguishing *trees*, much better than its specialization, the chromatic polynomial $\chi_G(m)$, which takes the same value $m(m-1)^{n-1}$ on all trees with n vertices.

Question 7.3.20. Does the chromatic symmetric function (for $\mathbf{k} = \mathbb{Z}$) distinguish trees?

It has been checked that the answer is affirmative for trees on 23 vertices or less. There are also interesting partial results on this question by Martin, Morin and Wagner [161].

We close this section with a few other properties of $\Psi[G]$ proven by Stanley which follow easily from the theory we have developed. For example, his work makes no explicit mention of the chromatic Hopf algebra $\mathcal{G}$, and the fact that Ψ is a Hopf morphism (although he certainly notes the trivial algebra morphism property $\Psi[G_1 \sqcup G_2] = \Psi[G_1]\Psi[G_2]$). One property he proves is implicitly related to Ψ as a coalgebra morphism: he considers (in the case when $\mathbb{Q}$ is a subring of $\mathbf{k}$) the effect on Ψ of the operator $\frac{\partial}{\partial p_1} : \Lambda_{\mathbb{Q}} \longrightarrow \Lambda_{\mathbb{Q}}$ which acts by first expressing a symmetric function $f \in \Lambda_{\mathbb{Q}}$ as a polynomial in the power sums $\{p_n\}$, and then applies the partial derivative operator $\frac{\partial}{\partial p_1}$ of the polynomial ring $\mathbb{Q}[p_1, p_2, p_3, \ldots]$. It is not hard to see that $\frac{\partial}{\partial p_1}$ is the same as the skewing operator $s_{(1)}^{\perp} = p_1^{\perp}$: both act as derivations on $\Lambda_{\mathbb{Q}} = \mathbb{Q}[p_1, p_2, \ldots]$ (since $p_1 \in \Lambda_{\mathbb{Q}}$ is primitive), and agree in their effect on each p_n, in that both send $p_1 \mapsto 1$, and both annihilate $p_2, p_3, \ldots$.

Proposition 7.3.21. *(Stanley [205, Cor. 2.12(a)]) For any graph $G = (V, E)$, one has*

$$\frac{\partial}{\partial p_1} \Psi[G] = \sum_{v \in V} \Psi[G|_{V \setminus v}].$$

Proof. Since Ψ is a coalgebra homomorphism, we have

$$\Delta \Psi[G] = (\Psi \otimes \Psi) \Delta[G] = \sum_{\substack{(V_1, V_2): \\ V = V_1 \sqcup V_2}} \Psi[G|_{V_1}] \otimes \Psi[G|_{V_2}].$$

Using this expansion (and the equality $\frac{\partial}{\partial p_1} = s_{(1)}^{\perp}$), we now compute

$$\frac{\partial}{\partial p_1} \Psi[G] = s_{(1)}^{\perp} \Psi[G] = \sum_{\substack{(V_1, V_2): \\ V = V_1 \sqcup V_2}} (s_{(1)}, \Psi[G|_{V_1}]) \cdot \Psi[G|_{V_2}] = \sum_{v \in V} \Psi[G|_{V \setminus v}]$$

(since degree considerations force $(s_{(1)}, \Psi[G|_{V_1}]) = 0$ unless $|V_1| = 1$, in which case $\Psi[G|_{V_1}] = s_{(1)}$). □

[360] A recent refinement for incomparability graphs of posets which are both $(\mathbf{3}+\mathbf{1})$- and $(\mathbf{2}+\mathbf{2})$-free, also known as *unit interval orders* is discussed by Shareshian and Wachs [198].

Definition 7.3.22. Given a graph $G = (V, E)$, an acyclic orientation Ω of the edges E (that is, an orientation of each edge such that the resulting directed graph has no cycles), and a vertex-coloring $f : V \to \{1, 2, \ldots\}$, say that the pair (Ω, f) are *weakly compatible* if whenever Ω orients an edge $\{v, v'\}$ in E as $v \to v'$, one has $f(v) \leq f(v')$. Note that a *proper* vertex-coloring f of a graph $G = (V, E)$ is weakly compatible with a unique acyclic orientation Ω.

Proposition 7.3.23. *(Stanley [205, Prop. 4.1, Thm. 4.2]) The involution ω of Λ sends $\Psi[G]$ to $\omega(\Psi[G]) = \sum_{(\Omega,f)} \mathbf{x}_f$ in which the sum runs over weakly compatible pairs (Ω, f) of an acyclic orientation Ω and vertex-coloring f.*

Furthermore, the chromatic polynomial $\chi_G(m)$ has the property that $(-1)^{|V|}\chi_G(-m)$ counts all such weakly compatible pairs (Ω, f) in which $f : V \to \{1, 2, \ldots, m\}$ is a vertex-m-coloring.

Proof. As observed above, a proper coloring f is weakly compatible with a unique acyclic orientation Ω of G. Denote by P_Ω the poset on V which is the transitive closure of Ω, endowed with a *strict labelling* by integers, that is, every $i \in P_\Omega$ and $j \in P_\Omega$ satisfying $i <_{P_\Omega} j$ must satisfy $i >_{\mathbb{Z}} j$. Then proper colorings f that induce Ω are the same as P_Ω-partitions, so that

$$\Psi[G] = \sum_{\Omega} F_{P_\Omega}(\mathbf{x}). \tag{7.3.5}$$

Applying the antipode S and using Corollary 5.2.20 gives

$$\omega(\Psi[G]) = (-1)^{|V|} S(\Psi[G]) = \sum_{\Omega} F_{P_\Omega^{\text{opp}}}(\mathbf{x}) = \sum_{(\Omega,f)} \mathbf{x}_f$$

where in the last line one sums over weakly compatible pairs as in the proposition. The last equality comes from the fact that since each P_Ω has been given a strict labelling, P_Ω^{opp} acquires a *weak (or natural) labelling*, that is, every $i \in P_\Omega$ and $j \in P_\omega$ satisfying $i <_{P_\Omega^{\text{opp}}} j$ must satisfy $i <_{\mathbb{Z}} j$.

The last assertion follows from Proposition 7.1.7(iii). □

Remark 7.3.24. The interpretation of $\chi_G(-m)$ in Proposition 7.3.23 is a much older result of Stanley [204]. The special case interpreting $\chi_G(-1)$ as $(-1)^{|V|}$ times the number of acyclic orientations of G has sometimes been called Stanley's *(-1)-color theorem*. It also follows (via Proposition 7.1.7) from Humpert and Martin's antipode formula for $\mathcal{G}$ discussed in Remark 7.3.4: taking ζ to be the character of $\mathcal{G}$ given in (7.3.4),

$$\chi_G(-1) = \zeta^{\star(-1)}[G] = \zeta(S[G]) = \sum_F (-1)^{|V|-\text{rank}(F)} \operatorname{acyc}(G/F)\zeta[G_{V,F}] = (-1)^{|V|} \operatorname{acyc}(G)$$

where the last equality uses the vanishing of ζ on graphs that have edges, so only the $F = \varnothing$ term survives.

Exercise 7.3.25. If V and X are two sets, and if $f : V \to X$ is any map, then $\operatorname{eqs} f$ will denote the set

$$\{\{u, u'\} \mid u \in V,\ u' \in V,\ u \neq u' \text{ and } f(u) = f(u')\}.$$

This is a subset of the set of all two-element subsets of V.

If $G = (V, E)$ is a finite graph, then show that the map Ψ introduced in Definition 7.3.16 satisfies

$$\Psi\left([G]^\sharp\right) = \sum_{\substack{f:V\to\{1,2,3,\ldots\};\\ \operatorname{eqs} f = E}} \mathbf{x}_f,$$

where $\mathbf{x}_f := \prod_{v \in V} x_{f(v)}$. Here, $[G]^\sharp$ is defined as in Definition 7.3.8.

7.4. **Example: The quasisymmetric function of a matroid.** We introduce the *matroid-minor Hopf algebra* of Schmitt [191], and studied extensively by Crapo and Schmitt [41, 42, 43]. A very simple character ζ on this Hopf algebra will then give rise, via the map Ψ from Theorem 7.1.3, to the quasisymmetric function invariant of matroids from the work of Billera, Jia and the second author [21].

7.4.1. *The matroid-minor Hopf algebra.* We begin by reviewing some notions from matroid theory; see Oxley [164] for background, undefined terms and unproven facts.

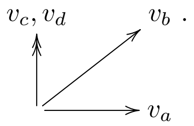
Definition 7.4.1. A *matroid* M of rank r on a (finite) ground set E is specified by a nonempty collection $\mathcal{B}(M)$ of r-element subsets of E with the following *exchange property*:

For any B, B' in $\mathcal{B}(M)$ and b in B, there exists b' in B' with $(B \setminus \{b\}) \cup \{b'\}$ in $\mathcal{B}(M)$.

The elements of $\mathcal{B}(M)$ are called the *bases* of the matroid M.

Example 7.4.2. A matroid M with ground set E is *represented* by a family of vectors $S = (v_e)_{e \in E}$ in a vector space if $\mathcal{B}(M)$ is the collection of subsets $B \subset E$ having the property that the subfamily $(v_e)_{e \in B}$ is a basis for the span of all of the vectors in S.

For example, if M is the matroid with $\mathcal{B}(M) = \{\{a,b\},\{a,c\},\{a,d\},\{b,c\},\{b,d\}\}$ on the ground set $E = \{a,b,c,d\}$, then M is represented by the family $S = (v_a, v_b, v_c, v_d)$ of the four vectors $v_a = (1,0), v_b = (1,1), v_c = (0,1) = v_d$ in $\mathbb{R}^2$ depicted here



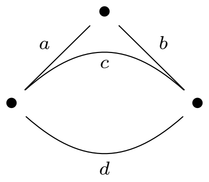
Conversely, whenever E is a finite set and $S = (v_e)_{e \in E}$ is a family of vectors in a vector space, then the set

$$\left\{B \subset E : \text{the subfamily } (v_e)_{e \in B} \text{ is a basis for the span of all of the vectors in } S\right\}$$

is a matroid on the ground set E.

A matroid is said to be *linear* if there exists a family of vectors in a vector space representing it. Not all matroids are linear, but many important ones are.

Example 7.4.3. A special case of matroids M represented by vectors are *graphic matroids*, coming from a graph $G = (V, E)$, with parallel edges and self-loops allowed. One represents these by vectors in $\mathbb{R}^V$ with standard basis $\{\epsilon_v\}_{v \in V}$ by associating the vector $\epsilon_v - \epsilon_{v'}$ to any edge connecting a vertex v with a vertex v'. One can check (or see [164, §1.2]) that the bases B in $\mathcal{B}(M)$ correspond to the edge sets of *spanning forests* for G, that is, edge sets which are acyclic and contain one spanning tree for each connected component of G. For example, the matroid $\mathcal{B}(M)$ corresponding to the graph $G = (V, E)$ shown below:



is exactly the matroid represented by the vectors in Example 7.4.2; indeed, the spanning forests of this graph G are the edge sets $\{a,b\},\{a,c\},\{a,d\},\{b,c\},\{b,d\}$. (In this example, spanning forests are the same as spanning trees, since G is connected.)

To define the matroid-minor Hopf algebra one needs the basic matroid operations of *deletion* and *contraction*. These model the operations of deleting or contracting an edge in a graph. For configurations of vectors they model the deletion of a vector, or the passage to images in the quotient space modulo the span of a vector.

Definition 7.4.4. Given a matroid M of rank r and an element e of its ground set E, say that e is *loop* (resp. *coloop*) of M if e lies in no basis (resp. every basis) B in $\mathcal{B}(M)$. If e is not a coloop, the *deletion* $M \setminus e$ is a matroid of rank r on ground set $E \setminus \{e\}$ having bases

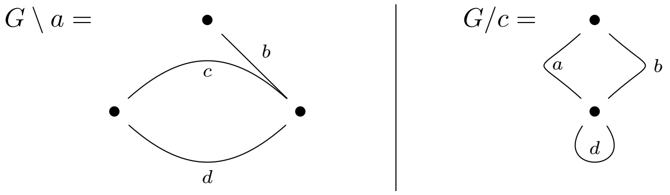
$$\mathcal{B}(M \setminus e) := \{B \in \mathcal{B}(M) : e \notin B\}. \tag{7.4.1}$$

If e is not a loop, the *contraction* M/e is a matroid of rank $r-1$ on ground set $E \setminus \{e\}$ having bases

$$\mathcal{B}(M/e) := \{B \setminus \{e\} : e \in B \in \mathcal{B}(M)\}. \tag{7.4.2}$$

When e is a loop of M, then M/e has rank r instead of $r-1$ and one defines its bases as in (7.4.1) rather than (7.4.2); similarly, if e is a coloop of M then $M \setminus e$ has rank $r-1$ instead of r and one defines its bases as in (7.4.2) rather than (7.4.1).

Example 7.4.5. Starting with the graph G and its graphic matroid M from Example 7.4.3, the deletion $M \setminus a$ and contraction M/c correspond to the graphs $G \setminus a$ and G/c shown here:



One has

- $\mathcal{B}(M \setminus a) = \{\{b,c\},\{b,d\}\}$, so that b has become a coloop in $M \setminus a$, and
- $\mathcal{B}(M/c) = \{\{a\},\{b\}\}$, so that d has become a loop in M/c.

Definition 7.4.6. Deletions and contractions commute with each other. Thus, given a matroid M with ground set E, and a subset $A \subset E$, two well-defined matroids can be constructed:

- the *restriction* $M|_A$, which is a matroid on ground set A, obtained from M by deleting all $e \in E \setminus A$ in any order, and
- the *quotient/contraction* M/A, which is a matroid on ground set $E \setminus A$, obtained from M by contracting all $e \in A$ in any order.

We will also need the *direct sum* $M_1 \oplus M_2$ of two matroids M_1 and M_2. This is the matroid whose ground set $E = E_1 \sqcup E_2$ is the disjoint union of a copy of the ground sets E_1, E_2 for M_1, M_2, and whose bases are

$$\mathcal{B}(M_1 \oplus M_2) := \{B_1 \sqcup B_2 : B_i \in \mathcal{B}(M_i) \text{ for } i = 1, 2\}.$$

Lastly, say that two matroids M_1, M_2 are *isomorphic* if there is a bijection of their ground sets $E_1 \xrightarrow{\varphi} E_2$ having the property that $\varphi\mathcal{B}(M_1) = \mathcal{B}(M_2)$.

Now one can define the matroid-minor Hopf algebra, originally introduced by Schmitt [191, §15], and studied further by Crapo and Schmitt [41, 42, 43].

Definition 7.4.7. Let $\mathcal{M}$ have $\mathbf{k}$-basis elements $[M]$ indexed by isomorphism classes of matroids. Define the multiplication via

$$[M_1] \cdot [M_2] := [M_1 \oplus M_2],$$

so that the class $[\varnothing]$ of the *empty matroid* $\varnothing$ having empty ground set gives a unit. Define the comultiplication for M a matroid on ground set E via

$$\Delta[M] := \sum_{A \subset E} [M|_A] \otimes [M/A],$$

and a counit

$$\epsilon[M] := \begin{cases} 1, & \text{if } M = \varnothing; \\ 0, & \text{otherwise.} \end{cases}$$

Proposition 7.4.8. *The above maps endow $\mathcal{M}$ with the structure of a connected graded finite type Hopf algebra over $\mathbf{k}$, which is commutative.*

Proof. Checking the unit and counit conditions are straightforward. Associativity and commutativity of the multiplication follow because the direct sum operation $\oplus$ for matroids is associative and commutative up to isomorphism. Coassociativity follows because for a matroid M on ground set E, one has the following equality between the two candidates for $\Delta^{(2)}[M]$:

$$\sum_{\varnothing \subset A_1 \subset A_2 \subset E} [M|_{A_1}] \otimes [(M|_{A_2})/A_1] \otimes [M/A_2]$$
$$= \sum_{\varnothing \subset A_1 \subset A_2 \subset E} [M|_{A_1}] \otimes [(M/A_1)|_{A_2 \setminus A_1}] \otimes [M/A_2]$$

due to the matroid isomorphism $(M|_{A_2})/A_1 \cong (M/A_1)|_{A_2 \setminus A_1}$. Commutativity of the bialgebra diagram in (1.3.4) amounts to the fact that for a pair of matroids M_1, M_2 and subsets A_1, A_2 of their (disjoint) ground sets E_1, E_2, one has isomorphisms

$$M_1|_{A_1} \oplus M_2|_{A_2} \cong (M_1 \oplus M_2)\,|_{A_1 \sqcup A_2},$$
$$M_1/A_1 \oplus M_2/A_2 \cong (M_1 \oplus M_2)\,/(A_1 \sqcup A_2).$$

Letting $\mathcal{M}_n$ be the $\mathbf{k}$-span of $[M]$ for matroids whose ground set E has cardinality $|E| = n$, one can then easily check that $\mathcal{M}$ becomes a bialgebra which is graded, connected, and of finite type, hence also a Hopf algebra by Proposition 1.4.16. □

See [59] for an application of $\mathcal{M}$ (and the operator $\exp^\star$ from Section 1.7) to proving the *Tutte recipe theorem*, a "universal" property of the Tutte polynomial of a matroid.

7.4.2. *A quasisymmetric function for matroids.*

Definition 7.4.9. Define a character $\mathcal{M} \xrightarrow{\zeta} \mathbf{k}$ by

$$\zeta[M] = \begin{cases} 1, & \text{if } M \text{ has only one basis;} \\ 0, & \text{otherwise.} \end{cases}$$

It is easily checked that this is a character, that is, an algebra morphism $\mathcal{M} \xrightarrow{\zeta} \mathbf{k}$. Note that if M has only one basis, say $\mathcal{B}(M) = \{B\}$, then $B := \operatorname{coloops}(M)$ is the set of coloops of M, and $E \setminus B = \operatorname{loops}(M)$ is the set of loops of M. Equivalently, $M = \bigoplus_{e \in E} M|_{\{e\}}$ is the direct sum of matroids each having one element, each a coloop or loop.

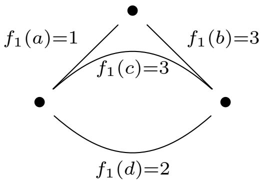
Define $\Psi[M]$ for a matroid M to be the image of $[M]$ under the map $\mathcal{M} \xrightarrow{\Psi} \mathrm{QSym}$ induced via Theorem 7.1.3 from the above character ζ.

It turns out that $\Psi[M]$ is intimately related with greedy algorithms and finding minimum cost bases. A fundamental property of matroids (and one that characterizes them, in fact; see [164, §1.8]) is that no matter how one assigns costs $f : E \to \mathbb{R}$ to the elements of E, the following *greedy algorithm* (generalizing *Kruskal's algorithm* for finding minimum cost spanning trees) always succeeds in finding one basis B in $\mathcal{B}(M)$ achieving the minimum *total cost* $f(B) := \sum_{b \in B} f(b)$:

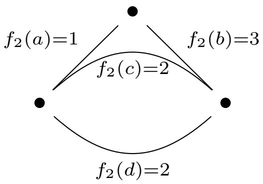
Algorithm 7.4.10. Start with the empty subset $I_0 = \varnothing$ of E. For $j = 1, 2, \ldots, r$, having already defined the set I_{j-1}, let e be the element of $E \setminus I_{j-1}$ having the lowest cost $f(e)$ among all those for which $I_{j-1} \cup \{e\}$ is *independent*, that is, still a subset of at least one basis B in $\mathcal{B}(M)$. Then define $I_j := I_{j-1} \cup \{e\}$. Repeat this until $j = r$, and $B = I_r$ will be among the bases that achieve the minimum cost.

Definition 7.4.11. Say that a cost function $f : E \to \{1, 2, \ldots\}$ is *M-generic* if there is a *unique* basis B in $\mathcal{B}(M)$ achieving the minimum cost $f(B)$.

Example 7.4.12. For the graphic matroid M of Example 7.4.3, this cost function $f_1 : E \to \{1, 2, \ldots\}$



is M-generic, as it minimizes uniquely on the basis $\{a, d\}$, whereas this cost function $f_2 : E \to \{1, 2, \ldots\}$



is *not* M-generic, as it achieves its minimum value on the two bases $\{a,c\},\{a,d\}$.

Proposition 7.4.13. *For a matroid M on ground set E, one has this expansion*[361]

$$\Psi[M] = \sum_{\substack{M\text{-generic} \\ f:E\to\{1,2,\ldots\}}} \mathbf{x}_f$$

where $\mathbf{x}_f := \prod_{e\in E} x_{f(e)}$. In particular, for $m \geq 0$, its specialization ps^1 from Definition 7.1.6 has this interpretation:

$$\mathrm{ps}^1 \Psi[M](m) = |\{M\text{-generic } f : E \to \{1,2,\ldots,m\}\}|.$$

Proof. The iterated coproduct $\mathcal{M} \xrightarrow{\Delta^{(\ell-1)}} \mathcal{M}^{\otimes \ell}$ sends

$$[M] \longmapsto \sum [M|_{A_1}] \otimes [(M|_{A_2})/A_1] \otimes \cdots \otimes [(M|_{A_\ell})/A_{\ell-1}]$$

where the sum is over flags of nested subsets

$$\varnothing = A_0 \subset A_1 \subset \cdots \subset A_{\ell-1} \subset A_\ell = E. \tag{7.4.3}$$

The map $\zeta^{\otimes \ell}$ sends each summand to 1 or 0, depending upon whether each $(M|_{A_j})/A_{j-1}$ has a unique basis or not. Thus formula (7.1.3) shows that the coefficient ζ_α of $x_{i_1}^{\alpha_1} \cdots x_{i_\ell}^{\alpha_\ell}$ in $\Psi[M]$ counts the flags of subsets in (7.4.3) for which $|A_j \setminus A_{j-1}| = \alpha_j$ and $(M|_{A_j})/A_{j-1}$ has a unique basis, for each j.

Given a flag as in (7.4.3), associate the cost function $f : E \to \{1,2,\ldots\}$ whose value on each element of $A_j \setminus A_{j-1}$ is i_j; conversely, given any cost function f, say whose distinct values are $i_1 < \cdots < i_\ell$, one associates the flag having $A_j \setminus A_{j-1} = f^{-1}(i_j)$ for each j.

Now, apply the greedy algorithm (Algorithm 7.4.10) to find a minimum-cost basis of M for such a cost function f. At each step of the greedy algorithm, one new element is added to the independent set; these elements weakly increase in cost as the algorithm progresses[362]. Thus, the algorithm first adds some elements of cost i_1, then adds some elements of cost i_2, then adds some elements of cost i_3, and so on. We can therefore subdivide the execution of the algorithm into phases $1,2,\ldots,\ell$, where each phase consists of some finite number of steps, such that all elements added in phase k have cost i_k. (A phase may be empty.) For each $k \in \{1,2,\ldots,\ell\}$, we let β_k be the number of steps in phase k; in other words, β_k is the number of elements of elements of cost i_k added during the algorithm.

We will prove below, using induction on $s = 0,1,2,\ldots,\ell$ the following **claim**: After having completed phases $1,2,\ldots,s$ in the greedy algorithm (Algorithm 7.4.10), there is *a unique choice* for the independent set produced thus far, namely

$$I_{\beta_1+\beta_2+\cdots+\beta_s} = \bigsqcup_{j=1}^{s} \mathrm{coloops}((M|_{A_j})/A_{j-1}), \tag{7.4.4}$$

if and only if each of the matroids $(M|_{A_j})/A_{j-1}$ for $j = 1,2,\ldots,s$ *has a unique basis.*

The case $s = \ell$ in this claim would show what we want, namely that f is M-generic, minimizing uniquely on the basis shown in (7.4.4) with $s = \ell$, if and only if each $(M|_{A_j})/A_{j-1}$ has a unique basis.

The assertion of the claim is trivially true for $s = 0$. In the inductive step, one may assume that

- the independent set $I_{\beta_1+\beta_2+\cdots+\beta_{s-1}}$ takes the form in (7.4.4), replacing s by $s-1$,
- it is the unique f-minimizing basis for $M|_{A_{s-1}}$, and
- $(M|_{A_j})/A_{j-1}$ has a unique basis for $j = 1,2,\ldots,s-1$.

Since A_{s-1} exactly consists of all of the elements e of E whose costs $f(e)$ lie in the range $\{i_1,i_2,\ldots,i_{s-1}\}$, in phase s the algorithm will work in the quotient matroid M/A_{s-1} and attempt to augment $I_{\beta_1+\beta_2+\cdots+\beta_{s-1}}$ using the next-cheapest elements, namely the elements of $A_s \setminus A_{s-1}$, which all have cost f equal to i_s. Thus the algorithm will have no choices about how to do this augmentation if and only if $(M|_{A_s})/A_{s-1}$ has a unique basis, namely its set of coloops, in which case the algorithm will choose to add all of these coloops, giving $I_{\beta_1+\beta_2+\cdots+\beta_s}$ as described in (7.4.4). This completes the induction.

The last assertion follows from Proposition 7.1.7. □

[361] In fact, this expansion was the original definition of $\Psi[M]$ in [21, Defn. 1.1].

[362] *Proof.* Let e be the element added at step i, and let e' be the element added at step $i+1$. We want to show that $f(e) \leq f(e')$. But the element e' could already have been added at step i. Since it wasn't, we thus conclude that the element e that was added instead must have been cheaper or equally expensive. In other words, $f(e) \leq f(e')$, qed.

Example 7.4.14. If M has one basis then every function $f : E \to \{1, 2, \ldots\}$ is M-generic, and

$$\Psi[M] = \sum_{f:E\to\{1,2,\ldots\}} \mathbf{x}_f = (x_1 + x_2 + \cdots)^{|E|} = M_{(1)}^{|E|}.$$

Example 7.4.15. Let $U_{r,n}$ denote the *uniform matroid* of rank r on n elements E, having $\mathcal{B}(U_{r,n})$ equal to all of the r-element subsets of E.

As $U_{1,2}$ has $E = \{1, 2\}$ and $\mathcal{B} = \{\{1\}, \{2\}\}$, genericity means $f(1) \neq f(2)$, so

$$\Psi[U_{1,2}] = \sum_{\substack{(f(1),f(2)):\\ f(1)\neq f(2)}} x_{f(1)}x_{f(2)} = x_1x_2 + x_2x_1 + x_1x_3 + x_3x_1 + \cdots = 2M_{(1,1)}.$$

Similarly $U_{1,3}$ has $E = \{1, 2, 3\}$ with $\mathcal{B} = \{\{1\}, \{2\}, \{3\}\}$, and genericity means either that $f(1), f(2), f(3)$ are all distinct, or that two of them are the same and the third is smaller. This shows

$$\begin{aligned}\Psi[U_{1,3}] &= 3\sum_{i<j} x_i x_j^2 + 6\sum_{i<j<k} x_i x_j x_k \\ &= 3M_{(1,2)} + 6M_{(1,1,1)}; \\ \mathrm{ps}^1\Psi[U_{1,3}](m) &= 3\binom{m}{2} + 6\binom{m}{3} = \frac{m(m-1)(2m-1)}{2}.\end{aligned}$$

One can similarly analyze $U_{2,3}$ and check that

$$\begin{aligned}\Psi[U_{2,3}] &= 3M_{(2,1)} + 6M_{(1,1,1)}; \\ \mathrm{ps}^1\Psi[U_{2,3}](m) &= 3\binom{m}{2} + 6\binom{m}{3} = \frac{m(m-1)(2m-1)}{2}.\end{aligned}$$

These last examples illustrate the behavior of Ψ under the duality operation on matroids.

Definition 7.4.16. Given a matroid M of rank r on ground set E, its *dual* or *orthogonal matroid* $M^\perp$ is a matroid of rank $|E| - r$ on the same ground set E, having

$$\mathcal{B}(M^\perp) := \{E \setminus B\}_{B\in\mathcal{B}(M)}.$$

See [164, Theorem 2.1.1] or [34, Section 4] for a proof of the fact that this is well-defined (i.e., that the collection $\{E \setminus B\}_{B\in\mathcal{B}(M)}$ really satisfies the exchange property). Here are a few examples of dual matroids.

Example 7.4.17. The dual of a uniform matroid is another uniform matroid:

$$U_{r,n}^\perp = U_{n-r,n}.$$

Example 7.4.18. If M is matroid of rank r represented by family of vectors $\{e_1, \ldots, e_n\}$ in a vector space over some field $\mathbf{k}$, one can find a family of vectors $\{e_1^\perp, \ldots, e_n^\perp\}$ that represent $M^\perp$ in the following way. Pick a basis for the span of the vectors $\{e_i\}_{i=1}^n$, and create a matrix A in $\mathbf{k}^{r\times n}$ whose columns express the e_i in terms of this basis. Then pick any matrix $A^\perp$ whose row space is the null space of A, and one finds that the columns $\{e_i^\perp\}_{i=1}^n$ of $A^\perp$ represent $M^\perp$. See Oxley [164, §2.2].

Example 7.4.19. Let $G = (V, E)$ be a graph embedded in the plane with edge set E, giving rise to a graphic matroid M on ground set E. Let $G^\perp$ be a planar dual of G, so that, in particular, for each edge e in E, the graph $G^\perp$ has one edge $e^\perp$, crossing e transversely. Then the graphic matroid of $G^\perp$ is $M^\perp$. See Oxley [164, §2.3].

Proposition 7.4.20. *If* $\Psi[M] = \sum_\alpha c_\alpha M_\alpha$ *then* $\Psi[M^\perp] = \sum_\alpha c_\alpha M_{\mathrm{rev}(\alpha)}$.

Consequently, $\mathrm{ps}^1\Psi[M](m) = \mathrm{ps}^1\Psi[M^\perp](m)$.

Proof. First, let us prove that if $\Psi[M] = \sum_\alpha c_\alpha M_\alpha$ then $\Psi[M^\perp] = \sum_\alpha c_\alpha M_{\mathrm{rev}(\alpha)}$. In other words, let us show that for any given composition α, the coefficient of M_α in $\Psi[M]$ (when $\Psi[M]$ is expanded in the basis $(M_\beta)_{\beta\in\mathrm{Comp}}$ of QSym) equals the coefficient of $M_{\mathrm{rev}(\alpha)}$ in $\Psi[M^\perp]$. This amounts to showing that for any composition $\alpha = (\alpha_1, \ldots, \alpha_\ell)$, the cardinality of the set of M-generic f having $\mathbf{x}_f = \mathbf{x}^\alpha$ is the same as the cardinality of the set of $M^{\perp}$-generic $f^{\perp}$ having $\mathbf{x}_{f^{\perp}} = \mathbf{x}^{\mathrm{rev}(\alpha)}$. We claim that the map $f \longmapsto f^{\perp}$ in which $f^{\perp}(e) = \ell + 1 - f(e)$ gives a bijection between these sets. To see this, note that any basis B of M satisfies

$$f(B) + f(E \setminus B) = \sum_{e \in E} f(e), \tag{7.4.5}$$

$$f(E \setminus B) + f^{\perp}(E \setminus B) = (\ell + 1)(|E| - r), \tag{7.4.6}$$

where r denotes the rank of M. Thus B is f-minimizing if and only if $E \setminus B$ is f-maximizing (by (7.4.5)) if and only if $E \setminus B$ is $f^{\perp}$-minimizing (by (7.4.6)). Consequently f is M-generic if and only if $f^{\perp}$ is $M^{\perp}$-generic.

The last assertion follows, for example, from the calculation in Proposition 7.1.7(i) that $\mathrm{ps}^1(M_\alpha)(m) = \binom{m}{\ell(\alpha)}$ together with the fact that $\ell(\mathrm{rev}(\alpha)) = \ell(\alpha)$. □

Just as (7.3.5) showed that Stanley's chromatic symmetric function of a graph has an expansion as a sum of P-partition enumerators for certain strictly labelled posets[363] P, the same holds for $\Psi[M]$.

Definition 7.4.21. Given a matroid M on ground set E, and a basis B in $\mathcal{B}(M)$, define the *base-cobase poset* P_B to have $b < b'$ whenever b lies in B and b' lies in $E \setminus B$ and $(B \setminus \{b\}) \cup \{b'\}$ is in $\mathcal{B}(M)$.

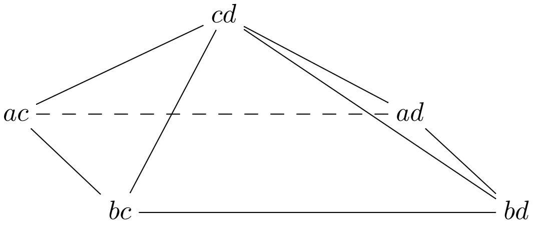
Proposition 7.4.22. *For any matroid M, one has $\Psi[M] = \sum_{B \in \mathcal{B}(M)} F_{(P_B, \mathrm{strict})}(\mathbf{x})$ where $F_{(P,\, \mathrm{strict})}(\mathbf{x})$ for a poset P means the P-partition enumerator for any strict labelling of P, i.e. a labelling such that the P-partitions satisfy $f(i) < f(j)$ whenever $i <_P j$.*

In particular, $\Psi[M]$ expands nonnegatively in the $\{L_\alpha\}$ basis.

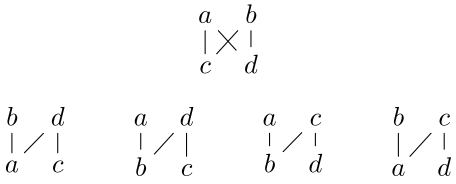
Proof. A basic result about matroids, due to Edmonds [62], describes the *edges* in the *matroid base polytope* which is the convex hull of all vectors $\{\sum_{b \in B} \epsilon_b\}_{B \in \mathcal{B}(M)}$ inside $\mathbb{R}^E$ with standard basis $\{\epsilon_e\}_{e \in E}$. He shows that all such edges connect two bases B, B' that differ by a single *basis exchange*, that is, $B' = (B \setminus \{b\}) \cup \{b'\}$ for some b in B and b' in $E \setminus B$.

Polyhedral theory then says that a cost function f on E will minimize uniquely at B if and only if one has a strict increase $f(B) < f(B')$ along each such edge $B \to B'$ emanating from B, that is, if and only if $f(b) < f(b')$ whenever $b <_{P_B} b'$ in the base-cobase poset P_B, that is, f lies in $\mathcal{A}(P_B, \mathrm{strict})$. □

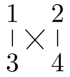
Example 7.4.23. The graphic matroid from Example 7.4.3 has this matroid base polytope, with the bases B in $\mathcal{B}(M)$ labelling the vertices:



The base-cobase posets P_B for its five vertices B are as follows:



One can label the first of these five strictly as

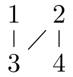


and compute its strict P-partition enumerator from the linear extensions $\{3412, 3421, 4312, 4321\}$ as

$$L_{(2,2)} + L_{(2,1,1)} + L_{(1,1,2)} + L_{(1,1,1,1)},$$

[363] A labelled poset P is said to be *strictly labelled* if every two elements i and j of P satisfying $i <_P j$ satisfy $i >_{\mathbb{Z}} j$.

while any of the last four can be labelled strictly as



and they each have an extra linear extension 3142 giving their strict P-partition enumerators as

$$L_{(2,2)} + L_{(2,1,1)} + L_{(1,1,2)} + L_{(1,1,1,1)} + L_{(1,2,1)}.$$

Hence one has

$$\Psi[M] = 5L_{(2,2)} + 5L_{(1,1,2)} + 4L_{(1,2,1)} + 5L_{(2,1,1)} + 5L_{(1,1,1,1)}.$$

As M is a graphic matroid for a self-dual planar graph, one has a matroid isomorphism $M \cong M^{\perp}$ (see Example 7.4.19), reflected in the fact that $\Psi[M]$ is invariant under the symmetry swapping $M_\alpha \leftrightarrow M_{\mathrm{rev}(\alpha)}$ (and simultaneously swapping $L_\alpha \leftrightarrow L_{\mathrm{rev}(\alpha)}$).

This P-partition expansion for $\Psi[M]$ also allows us to identify its image under the antipode of QSym.

Proposition 7.4.24. *For a matroid M on ground set E, one has*

$$S(\Psi[M]) = (-1)^{|E|} \sum_{f:E\to\{1,2,\ldots\}} |\{f\text{-maximizing bases } B\}| \cdot \mathbf{x}_f$$

and

$$\mathrm{ps}^1\Psi[M](-m) = (-1)^{|E|} \sum_{f:E\to\{1,2,\ldots,m\}} |\{f\text{-maximizing bases } B\}|.$$

In particular, the expected number of f-maximizing bases among all cost functions $f : E \to \{1,2,\ldots,m\}$ is $(-m)^{-|E|}\mathrm{ps}^1\Psi[M](-m)$.

Proof. Corollary 5.2.20 implies

$$S(\Psi[M]) = \sum_{B\in\mathcal{B}(M)} S(F_{(P_B,\mathrm{strict})}(\mathbf{x})) = (-1)^{|E|} \sum_{B\in\mathcal{B}(M)} F_{(P_B^{\mathrm{opp}},\mathrm{natural})}(\mathbf{x}),$$

where $F_{(P,\mathrm{natural})}(\mathbf{x})$ is the enumerator for P-partitions in which P has been *naturally* labelled, so that they satisfy $f(i) \leq f(j)$ whenever $i <_P j$. When $P = P_B^{\mathrm{opp}}$, this is exactly the condition for f to achieve its maximum value at $f(B)$ (possibly not uniquely), that is, for f to lie in the *closed* normal cone to the vertex indexed by B in the matroid base polytope; compare this with the discussion in the proof of Proposition 7.4.22. Thus one has

$$S(\Psi[M]) = (-1)^{|E|} \sum_{\substack{(B,f):\\ B\in\mathcal{B}(M)\\ f \text{ maximizing at } B}} \mathbf{x}_f,$$

which agrees with the statement of the proposition, after reversing the order of the summation.

The rest follows from Proposition 7.1.7. □

Example 7.4.25. We saw in Example 7.4.23 that the matroid M from Example 7.4.3 has

$$\Psi[M] = 5L_{(2,2)} + 5L_{(1,1,2)} + 4L_{(1,2,1)} + 5L_{(2,1,1)} + 5L_{(1,1,1,1)},$$

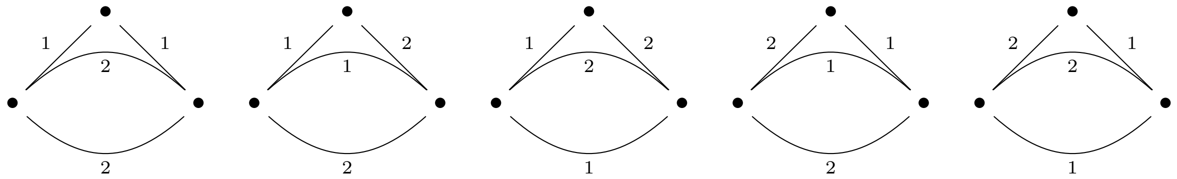
and therefore will have

$$\mathrm{ps}^1\Psi[M](m) = 5\binom{m-2+4}{4} + (5+4+5)\binom{m-3+4}{4} + 5\binom{m-4+4}{4} = \frac{m(m-1)(2m^2-2m+1)}{2}$$

using $\mathrm{ps}^1(L_\alpha)(m) = \binom{m-\ell+|\alpha|}{|\alpha|}$ from Proposition 7.1.7 (i). Let us first do a reality-check on a few of its values with $m \geq 0$ using Proposition 7.4.13, and for negative m using Proposition 7.4.24:

m	-1	0	1	2
$\mathrm{ps}^1\Psi[M](m)$	5	0	0	5

When $m = 0$, interpreting the set of cost functions $f : E \to \{1,2,\ldots,m\}$ as being empty explains why the value shown is 0. When $m = 1$, there is only one function $f : E \to \{1\}$, and it is not M-generic; any of the 5 bases in $\mathcal{B}(M)$ will minimize $f(B)$, explaining both why the value for $m = 1$ is 0, but also explaining the value of 5 for $m = -1$. The value of 5 for $m = 2$ counts these M-generic cost functions $f : E \to \{1, 2\}$:



Lastly, Proposition 7.4.24 predicts the expected number of f-minimizing bases for $f : E \to \{1, 2, \ldots, m\}$ as

$$(-m)^{-|E|}\mathrm{ps}^1\Psi[M](-m) = (-m)^{-4}\frac{m(m+1)(2m^2+2m+1)}{2} = \frac{(m+1)(2m^2+2m+1)}{2m^3},$$

whose limit as $m \to \infty$ is 1, consistent with the notion that "most" cost functions should be generic with respect to the bases of M, and maximize/minimize on a unique basis.

Remark 7.4.26. It is not coincidental that there is a similarity of results for Stanley's chromatic symmetric function of a graph $\Psi[G]$ and for the matroid quasisymmetric function $\Psi[M]$, such as the P-partition expansions (7.3.5) versus Proposition 7.4.22, and the reciprocity results Proposition 7.3.23 versus Proposition 7.4.24. It was noted in [21, §9] that one can associate a similar quasisymmetric function invariant to any *generalized permutohedra* in the sense of Postnikov [173]. Furthermore, recent work of Ardila and Aguiar [3] has shown that there is a Hopf algebra of such generalized permutohedra, arising from a *Hopf monoid* in the sense of Aguiar and Mahajan [6]. This Hopf algebra generalizes the chromatic Hopf algebra of graphs[364] and the matroid-minor Hopf algebra, and its quasisymmetric function invariant derives as usual from Theorem 7.1.3. Their work [3] also provides a generalization of the chromatic Hopf algebra antipode formula of Humpert and Martin [103] discussed in Remark 7.3.4 above.

[364] Aguiar and Ardila actually work with a larger Hopf algebra of graphs. Namely, their concept of graphs allows parallel edges, and it also allows "half-edges", which have only one endpoint. If $G = (V, E)$ is such a graph (where E is the set of its edges and its half-edges), and if V' is a subset of V, then they define $G/_{V'}$ to be the graph on vertex set V' obtained from G by

- removing all vertices that are not in V',
- removing all edges that have no endpoint in V', and all half-edges that have no endpoint in V', and
- replacing all edges that have only one endpoint in V' by half-edges.

(This is to be contrasted with the induced subgraph $G\mid_{V'}$, which is constructed in the same way but with the edges that have only one endpoint in V' getting removed as well.) The comultiplication they define on the Hopf algebra of such graphs sends the isomorphism class $[G]$ of a graph $G = (V, E)$ to $\sum_{(V_1,V_2):V_1\sqcup V_2=V} [G\mid_{V_1}] \otimes [G/_{V_2}]$. This is no longer a cocommutative Hopf algebra; our Hopf algebra $\mathcal{G}$ is a quotient of it. In [3, Corollary 13.10], Ardila and Aguiar compute the antipode of the Hopf monoid of such graphs; this immediately leads to a formula for the antipode of the corresponding Hopf algebra, because what they call the Fock functor $\overline{\mathcal{K}}$ preserves antipodes [3, Theorem 2.18].

8. The Malvenuto-Reutenauer Hopf algebra of permutations

Like so many Hopf algebras we have seen, the *Malvenuto-Reutenauer Hopf algebra* FQSym can be thought of fruitfully in more than one way. One is that it gives a natural noncommutative lift of the quasisymmetric P-partition enumerators and the fundamental basis $\{L_\alpha\}$ of QSym, rendering their product and coproduct formulas even more natural.

8.1. **Definition and Hopf structure.**

Definition 8.1.1. We shall regard permutations as words (over the alphabet $\{1, 2, 3, \ldots\}$) by identifying every permutation $\pi \in \mathfrak{S}_n$ with the word $(\pi(1), \pi(2), \ldots, \pi(n))$.

Define $\operatorname{FQSym} = \bigoplus_{n\geq 0} \operatorname{FQSym}_n$ to be a graded $\mathbf{k}$-module in which FQSym_n has $\mathbf{k}$-basis $\{F_w\}_{w\in\mathfrak{S}_n}$ indexed by the permutations $w = (w_1, \ldots, w_n)$ in $\mathfrak{S}_n$.

We first attempt to lift the product and coproduct formulas (5.2.6), (5.2.5) in the $\{L_\alpha\}$ basis of QSym. We attempt to define a product for $u \in \mathfrak{S}_k$ and $v \in \mathfrak{S}_\ell$ as follows[365]:

$$F_u F_v := \sum_{w\in u \sqcup\!\sqcup\, v[k]} F_w, \tag{8.1.1}$$

where for any word $v = (v_1, \ldots, v_\ell)$ we set $v[k] := (k + v_1, \ldots, k + v_\ell)$. Note that the multiset $u \sqcup\!\sqcup\, v[k]$ is an actual set in this situation (i.e., has each element appear only once) and is a subset of $\mathfrak{S}_{k+\ell}$.

The coproduct will be defined using the notation of standardization of $\operatorname{std}(w)$ a word w in some linearly ordered alphabet (see Definition 5.3.3).

Example 8.1.2. Considering words in the Roman alphabet $a < b < c < \cdots$, we have

$$\begin{array}{rccccccccccc} \operatorname{std}(b & a & c & c & b & a & a & b & a & c & b) \\ = (5 & 1 & 9 & 10 & 6 & 2 & 3 & 7 & 4 & 11 & 8). \end{array}$$

Using this, define for $w = (w_1, \ldots, w_n)$ in $\mathfrak{S}_n$ the element $\Delta F_w \in \operatorname{FQSym}\otimes\operatorname{FQSym}$ by

$$\Delta F_w := \sum_{k=0}^{n} F_{\operatorname{std}(w_1,w_2,\ldots,w_k)} \otimes F_{\operatorname{std}(w_{k+1},w_{k+2},\ldots,w_n)}. \tag{8.1.2}$$

It is possible to check directly that the maps defined in (8.1.1) and (8.1.2) endow FQSym with the structure of a connected graded finite type Hopf algebra; see Hazewinkel, Gubareni, Kirichenko [93, Thm. 7.1.8]. However in justifying this here, we will follow the approach of Duchamp, Hivert and Thibon [58, §3], which exhibits FQSym as a subalgebra of a larger ring of (noncommutative) power series of bounded degree in a totally ordered alphabet.

Definition 8.1.3. Given a totally ordered set I, create a totally ordered variable set $\{X_i\}_{i\in I}$, and the ring $R\langle\{X_i\}_{i\in I}\rangle$ of *noncommutative power series of bounded degree* in this alphabet[366]. Many times, we will use a variable set $\mathbf{X} := (X_1 < X_2 < \cdots)$, and call the ring $R\langle\mathbf{X}\rangle$.

[365] Recall that we regard permutations as words.

[366] Let us recall the definition of $R\langle\{X_i\}_{i\in I}\rangle$.

Let N denote the free monoid on the alphabet $\{X_i\}_{i\in I}$; it consists of words $X_{i_1}X_{i_2}\cdots X_{i_k}$. We define a topological $\mathbf{k}$-module $\mathbf{k}\langle\langle\{X_i\}_{i\in I}\rangle\rangle$ to be the Cartesian product $\mathbf{k}^N$ (equipped with the product topology), but we identify its element $(\delta_{w,u})_{u\in N}$ with the word w for every $w \in N$. Thus, every element $(\lambda_w)_{w\in N} \in \mathbf{k}^N = \mathbf{k}\langle\langle\{X_i\}_{i\in I}\rangle\rangle$ can be rewritten as the convergent sum $\sum_{w\in N} \lambda_w w$. We call λ_w the *coefficient of* w in this element (or the *coefficient of this element before* w). The elements of $\mathbf{k}\langle\langle\{X_i\}_{i\in I}\rangle\rangle$ will be referred to as *noncommutative power series*. We define a multiplication on $\mathbf{k}\langle\langle\{X_i\}_{i\in I}\rangle\rangle$ by the formula

$$\left(\sum_{w\in N} \lambda_w w\right)\left(\sum_{w\in N} \mu_w w\right) = \sum_{w\in N}\left(\sum_{(u,v)\in N^2;\ w=uv} \lambda_u\mu_v\right) w.$$

(This is well-defined thanks to the fact that, for each $w \in N$, there are only finitely many $(u, v) \in N^2$ satisfying $w = uv$.) Thus, $\mathbf{k}\langle\langle\{X_i\}_{i\in I}\rangle\rangle$ becomes a $\mathbf{k}$-algebra with unity 1 (the empty word). (It is similar to the monoid algebra $\mathbf{k}N$ of N over $\mathbf{k}$, with the only difference that infinite sums are allowed.)

Now, we define $R\langle\{X_i\}_{i\in I}\rangle$ to be the $\mathbf{k}$-subalgebra of $\mathbf{k}\langle\langle\{X_i\}_{i\in I}\rangle\rangle$ consisting of all noncommutative power series $\sum_{w\in N} \lambda_w w \in \mathbf{k}\langle\langle\{X_i\}_{i\in I}\rangle\rangle$ *of bounded degree* (i.e., such that all words $w \in N$ of sufficiently high length satisfy $\lambda_w = 0$).

We first identify the algebra structure for FQSym as the subalgebra of finite type within $R\langle \{X_i\}_{i\in I}\rangle$ spanned by the elements

$$F_w = F_w(\{X_i\}_{i\in I}) := \sum_{\substack{\mathbf{i}=(i_1,\ldots,i_n):\\ \operatorname{std}(\mathbf{i})=w^{-1}}} \mathbf{X}_{\mathbf{i}}, \tag{8.1.3}$$

where $\mathbf{X}_{\mathbf{i}} := X_{i_1}\cdots X_{i_n}$, as w ranges over $\bigcup_{n\geq 0}\mathfrak{S}_n$.

Example 8.1.4. For the alphabet $\mathbf{X} = (X_1 < X_2 < \cdots)$, in $R\langle\mathbf{X}\rangle$ one has

$$\begin{aligned}
F_1 &= \sum_{1\leq i} X_i = X_1 + X_2 + \cdots ,\\
F_{12} &= \sum_{1\leq i\leq j} X_iX_j = X_1^2 + X_2^2 + \cdots + X_1X_2 + X_1X_3 + X_2X_3 + X_1X_4 + \cdots ,\\
F_{21} &= \sum_{1\leq i<j} X_jX_i = X_2X_1 + X_3X_1 + X_3X_2 + X_4X_1 + \cdots ,\\
F_{312} &= \sum_{\mathbf{i}:\operatorname{std}(\mathbf{i})=231} \mathbf{X}_{\mathbf{i}} = \sum_{1\leq i<j\leq k} X_jX_kX_i\\
&= X_2^2X_1 + X_3^2X_1 + X_3^2X_2 + \cdots + X_2X_3X_1 + X_2X_4X_1 + \cdots .
\end{aligned}$$

Proposition 8.1.5. *For any totally ordered infinite set I, the elements $\{F_w\}$ as w ranges over $\bigcup_{n\geq 0}\mathfrak{S}_n$ form a $\mathbf{k}$-basis for a subalgebra* $\operatorname{FQSym}(\{X_i\}_{i\in I})$ *of $R\langle\mathbf{X}\rangle$, which is connected graded and of finite type, having multiplication defined $\mathbf{k}$-linearly by* (8.1.1).

Consequently all such algebras are isomorphic to a single algebra FQSym, *having basis $\{F_w\}$ and multiplication given by the rule* (8.1.1), *with the isomorphism mapping $F_w \longmapsto F_w(\{X_i\}_{i\in I})$.*

For example,

$$\begin{aligned}
F_1F_{21} &= (X_1 + X_2 + X_3 + \cdots)(X_2X_1 + X_3X_1 + X_3X_2 + X_4X_1 + \cdots)\\
&= X_1\cdot X_3X_2 + X_1\cdot X_4X_2 + \cdots + X_1\cdot X_2X_1 + X_2\cdot X_3X_2 + X_2\cdot X_4X_2 + \cdots\\
&\quad + X_2\cdot X_3X_1 + X_2\cdot X_4X_1 + \cdots + X_2\cdot X_2X_1 + X_3\cdot X_3X_1 + X_3\cdot X_3X_2 + \cdots\\
&\quad + X_3\cdot X_2X_1 + X_4\cdot X_2X_1 + \cdots\\
&= \sum_{\mathbf{i}:\operatorname{std}(\mathbf{i})=132} \mathbf{X}_{\mathbf{i}} + \sum_{\mathbf{i}:\operatorname{std}(\mathbf{i})=231} \mathbf{X}_{\mathbf{i}} + \sum_{\mathbf{i}:\operatorname{std}(\mathbf{i})=321} \mathbf{X}_{\mathbf{i}} = F_{132} + F_{312} + F_{321} = \sum_{w\in 1\,\sqcup\!\sqcup\, 32} F_w.
\end{aligned}$$

Proof of Proposition 8.1.5. The elements $\{F_w(\{X_i\}_{i\in I})\}$ are linearly independent as they are supported on disjoint monomials, and so form a $\mathbf{k}$-basis for their span. The fact that they multiply via rule (8.1.1) is the equivalence of conditions (i) and (iii) in the following Lemma 8.1.6, from which all the remaining assertions follow. □

Lemma 8.1.6. *For a triple of permutations*

$$\begin{aligned}
u &= (u_1,\ldots,u_k) \text{ in } \mathfrak{S}_k,\\
v &= (v_1,\ldots,v_{n-k}) \text{ in } \mathfrak{S}_{n-k},\\
w &= (w_1,\ldots,w_n) \text{ in } \mathfrak{S}_n,
\end{aligned}$$

the following conditions are equivalent:

(i) *w^{-1} lies in the set $u^{-1} \sqcup\!\sqcup\, v^{-1}[k]$.*
(ii) *$u = \operatorname{std}(w_1,\ldots,w_k)$ and $v = \operatorname{std}(w_{k+1},\ldots,w_n)$,*
(iii) *for some word $\mathbf{i} = (i_1,\ldots,i_n)$ with $\operatorname{std}(\mathbf{i}) = w$ one has $u = \operatorname{std}(i_1,\ldots,i_k)$ and $v = \operatorname{std}(i_{k+1},\ldots,i_n)$.*

Proof. The implication (ii) $\Rightarrow$ (iii) is clear since $\operatorname{std}(w) = w$. The reverse implication (iii) $\Rightarrow$ (ii) is best illustrated by example, e.g. considering Example 8.1.2 as concatenated, with $n = 11$ and $k = 6$ and $n-k = 5$:

$$\begin{array}{rcccccccc|ccccccc}
w = \operatorname{std} & (b & a & c & c & b & a & \mid & & a & b & a & c & b) \\
= & (5 & 1 & 9 & 10 & 6 & 2 & \mid & & 3 & 7 & 4 & 11 & 8) \\
\hline
u = \operatorname{std} & (5 & 1 & 9 & 10 & 6 & 2) & & v = \operatorname{std} & (3 & 7 & 4 & 11 & 8) \\
= & (3 & 1 & 5 & 6 & 4 & 2) & & = & (1 & 3 & 2 & 5 & 4) \\
= \operatorname{std} & (b & a & c & c & b & a) & & = \operatorname{std} & (a & b & a & c & b)
\end{array}$$

The equivalence of (i) and (ii) is a fairly standard consequence of unique parabolic factorization $W = W^J W_J$ where $W = \mathfrak{S}_n$ and $W_J = \mathfrak{S}_k \times \mathfrak{S}_{n-k}$, so that W^J are the minimum-length coset representatives for cosets xW_J (that is, the permutations $x \in \mathfrak{S}_n$ satisfying $x_1 < \cdots < x_k$ and $x_{k+1} < \cdots < x_n$). One can uniquely express any w in W as $w = xy$ with x in W^J and y in W_J, which here means that $y = u \cdot v[k] = v[k] \cdot u$ for some u in $\mathfrak{S}_k$ and v in $\mathfrak{S}_{n-k}$. Therefore $w = xuv[k]$, if and only if $w^{-1} = u^{-1}v^{-1}[k]x^{-1}$, which means that w^{-1} is the shuffle of the sequences u^{-1} in positions $\{x_1, \ldots, x_k\}$ and $v^{-1}[k]$ in positions $\{x_{k+1}, \ldots, x_n\}$. $\square$

Example 8.1.7. To illustrate the equivalence of (i) and (ii) and the parabolic factorization in the preceding proof, let $n = 9$ and $k = 5$ with

$$\begin{aligned}
w &= \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & \mid & 6 & 7 & 8 & 9 \\ 4 & 9 & 6 & 1 & 5 & \mid & 8 & 2 & 3 & 7 \end{pmatrix} \\
&= \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & \mid & 6 & 7 & 8 & 9 \\ 1 & 4 & 5 & 6 & 9 & \mid & 2 & 3 & 7 & 8 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 5 & 4 & 1 & 3 \end{pmatrix} \begin{pmatrix} 6 & 7 & 8 & 9 \\ 9 & 6 & 7 & 8 \end{pmatrix} \\
&= x \cdot u \cdot v[k];
\end{aligned}$$

then

$$\begin{aligned}
w^{-1} &= \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 4 & 9 & 6 & 1 & 5 & 8 & 2 & 3 & 7 \end{pmatrix} \\
&= \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 1 & 5 & 3 & 2 \end{pmatrix} \begin{pmatrix} 6 & 7 & 8 & 9 \\ 7 & 8 & 9 & 6 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ \underline{1} & \underline{\underline{6}} & \underline{\underline{7}} & \underline{2} & \underline{3} & \underline{4} & \underline{\underline{8}} & \underline{\underline{9}} & \underline{5} \end{pmatrix} \\
&= u^{-1} \cdot v^{-1}[k] \cdot x^{-1}.
\end{aligned}$$

Proposition 8.1.5 yields that FQSym is isomorphic to the $\mathbf{k}$-subalgebra $\operatorname{FQSym}(\mathbf{X})$ of the $\mathbf{k}$-algebra $R\langle \mathbf{X} \rangle$ when $\mathbf{X}$ is the variable set $(X_1 < X_2 < \cdots)$. We identify FQSym with $\operatorname{FQSym}(\mathbf{X})$ along this isomorphism. For any infinite alphabet $\{X_i\}_{i \in I}$ and any $f \in \operatorname{FQSym}$, we denote by $f\left(\{X_i\}_{i\in I}\right)$ the image of f under the algebra isomorphism $\operatorname{FQSym} \to \operatorname{FQSym}\left(\{X_i\}_{i\in I}\right)$ defined in Proposition 8.1.5.

One can now use this to define a coalgebra structure on FQSym. Roughly speaking, one wants to first evaluate an element f in $\operatorname{FQSym} \cong \operatorname{FQSym}(\mathbf{X}) \cong \operatorname{FQSym}(\mathbf{X}, \mathbf{Y})$ as $f(\mathbf{X}, \mathbf{Y})$, using the linearly ordered variable set $(\mathbf{X}, \mathbf{Y}) := (X_1 < X_2 < \cdots < Y_1 < Y_2 < \cdots)$. Then one should take the image of $f(\mathbf{X}, \mathbf{Y})$ after imposing the partial commutativity relations

$$X_i Y_j = Y_j X_i \text{ for every pair } (X_i, Y_j) \in \mathbf{X} \times \mathbf{Y}, \tag{8.1.4}$$

and hope that this image lies in a subalgebra isomorphic to

$$\operatorname{FQSym}(\mathbf{X}) \otimes \operatorname{FQSym}(\mathbf{Y}) \cong \operatorname{FQSym} \otimes \operatorname{FQSym}.$$

We argue this somewhat carefully. Start by considering the canonical monoid epimorphism

$$F\langle \mathbf{X}, \mathbf{Y} \rangle \overset{\rho}{\twoheadrightarrow} M, \tag{8.1.5}$$

where $F\langle \mathbf{X}, \mathbf{Y} \rangle$ denotes the *free monoid* on the alphabet $(\mathbf{X}, \mathbf{Y})$ and M denotes its quotient monoid imposing the partial commutativity relations (8.1.4). Let $\mathbf{k}^M$ denote the $\mathbf{k}$-module of all functions $f : M \to \mathbf{k}$, with pointwise addition and scalar multiplication; similarly define $\mathbf{k}^{F\langle \mathbf{X}, \mathbf{Y} \rangle}$. As both monoids $F\langle \mathbf{X}, \mathbf{Y} \rangle$ and M enjoy the property that an element m has only finitely many factorizations as $m = m_1 m_2$, one can define a convolution algebra structure on both $\mathbf{k}^{F\langle \mathbf{X}, \mathbf{Y} \rangle}$ and $\mathbf{k}^M$ via

$$(f_1 \star f_2)(m) = \sum_{\substack{(m_1, m_2) \in N \times N: \\ m = m_1 m_2}} f_1(m_1) f_2(m_2),$$

where N is respectively $F\langle \mathbf{X}, \mathbf{Y}\rangle$ or M. As fibers of the map ρ in (8.1.5) are finite, it induces a map of convolution algebras, which we also call ρ:

$$\mathbf{k}^{F\langle \mathbf{X},\mathbf{Y}\rangle} \overset{\rho}{\twoheadrightarrow} \mathbf{k}^{M}. \tag{8.1.6}$$

Now recall that $R\langle \mathbf{X}\rangle$ denotes the algebra of noncommutative formal power series in the variable set $\mathbf{X}$, of bounded degree, with coefficients in $\mathbf{k}$. One similarly has the ring $R\langle \mathbf{X}, \mathbf{Y}\rangle$, which can be identified with the subalgebra of $\mathbf{k}^{F\langle \mathbf{X},\mathbf{Y}\rangle}$ consisting of the functions $f : F\langle \mathbf{X}, \mathbf{Y}\rangle \to \mathbf{k}$ having a bound on the length of the words in their support (the value of f on a word in $(\mathbf{X}, \mathbf{Y})$ gives its power series coefficient corresponding to said word). We let $R\langle M\rangle$ denote the analogous subalgebra of $\mathbf{k}^{M}$; this can be thought of as the algebra of bounded degree "partially commutative power series" in the variable sets $\mathbf{X}$ and $\mathbf{Y}$. Note that ρ restricts to a map

$$R\langle \mathbf{X}, \mathbf{Y}\rangle \overset{\rho}{\to} R\langle M\rangle. \tag{8.1.7}$$

Finally, we claim (and see Proposition 8.1.9 below for a proof) that this further restricts to a map

$$\operatorname{FQSym}(\mathbf{X}, \mathbf{Y}) \overset{\rho}{\to} \operatorname{FQSym}(\mathbf{X}) \otimes \operatorname{FQSym}(\mathbf{Y}) \tag{8.1.8}$$

in which the target is identified with its image under the (injective[367]) multiplication map

$$\begin{array}{rcl} \operatorname{FQSym}(\mathbf{X}) \otimes \operatorname{FQSym}(\mathbf{Y}) & \hookrightarrow & R\langle M\rangle, \\ f(\mathbf{X}) \otimes g(\mathbf{Y}) & \mapsto & f(\mathbf{X})g(\mathbf{Y}). \end{array}$$

Using the identification of FQSym with all three of $\operatorname{FQSym}(\mathbf{X}), \operatorname{FQSym}(\mathbf{Y}), \operatorname{FQSym}(\mathbf{X}, \mathbf{Y})$, the map ρ in (8.1.8) will then define a coproduct structure on FQSym. Abusing notation, for f in FQSym, we will simply write $\Delta(f) = f(\mathbf{X}, \mathbf{Y})$ instead of $\rho(f(\mathbf{X}, \mathbf{Y}))$.

Example 8.1.8. Recall from Example 8.1.4 that one has

$$F_{312} = \sum_{\mathbf{i}:\operatorname{std}(\mathbf{i})=231} \mathbf{X_i} = \sum_{1\leq i<j\leq k} X_j X_k X_i,$$

and therefore its coproduct is

$$\begin{aligned} \Delta F_{312} &= F_{312}(X_1, X_2, \ldots, Y_1, Y_2, \ldots) \qquad\qquad \text{(by our abuse of notation)} \\ &= \sum_{i<j\leq k} X_j X_k X_i + \sum_{\substack{i<j,\\ k}} X_j Y_k X_i + \sum_{\substack{i,\\ j\leq k}} Y_j Y_k X_i + \sum_{i<j\leq k} Y_j Y_k Y_i \\ &= \sum_{i<j\leq k} X_j X_k X_i \cdot 1 + \sum_{\substack{i<j,\\ k}} X_j X_i \cdot Y_k + \sum_{\substack{i,\\ j\leq k}} X_i \cdot Y_j Y_k + \sum_{i<j\leq k} 1 \cdot Y_j Y_k Y_i \\ &= F_{312}(\mathbf{X}) \cdot 1 + F_{21}(\mathbf{X}) \cdot F_1(\mathbf{Y}) + F_1(\mathbf{X}) \cdot F_{12}(\mathbf{Y}) + 1 \cdot F_{312}(\mathbf{Y}) \\ &= F_{312} \otimes 1 + F_{21} \otimes F_1 + F_1 \otimes F_{12} + 1 \otimes F_{312}. \end{aligned}$$

Proposition 8.1.9. *The map ρ in (8.1.7) does restrict as claimed to a map as in (8.1.8), and hence defines a coproduct on* FQSym*, acting on the $\{F_w\}$ basis by the rule (8.1.2). This endows* FQSym *with the structure of a connected graded finite type Hopf algebra.*

Proof. Let I be the totally ordered set $\{1 < 2 < 3 < \cdots\}$. Let J be the totally ordered set $\left\{1 < 2 < 3 < \cdots < \widetilde{1} < \widetilde{2} < \widetilde{3} < \cdots\right\}$. We set $X_{\widetilde{i}} = Y_i$ for every positive integer i. Then, the alphabet $(\mathbf{X}, \mathbf{Y})$ can be written as $\{X_i\}_{i\in J}$.

If $\mathbf{i}$ is a word over the alphabet $I = \{1 < 2 < 3 < \cdots\}$, then we denote by $\widetilde{\mathbf{i}}$ the word over J obtained from $\mathbf{i}$ by replacing every letter i by $\widetilde{i}$.

[367] as images of the basis $F_u(\mathbf{X}) \otimes F_v(\mathbf{Y})$ of $\operatorname{FQSym}(\mathbf{X}) \otimes \operatorname{FQSym}(\mathbf{Y})$ are supported on disjoint monomials in $R\langle M\rangle$, so linearly independent.

For the first assertion of Proposition 8.1.9, it suffices to check that F_w indeed has the image under Δ claimed in (8.1.2). Let $n \in \mathbb{N}$ and $w \in \mathfrak{S}_n$. Then,

$$\begin{aligned} \Delta F_w &= F_w(\mathbf{X}, \mathbf{Y}) \qquad \text{(by our abuse of notation)} \\ &= \sum_{\mathbf{i} \in J^n : \operatorname{std}(\mathbf{i}) = w^{-1}} (\mathbf{X}, \mathbf{Y})_{\mathbf{i}} = \sum_{\mathbf{t} \in J^n : \operatorname{std}(\mathbf{t}) = w^{-1}} (\mathbf{X}, \mathbf{Y})_{\mathbf{t}} \\ &= \sum_{k=0}^{n} \sum_{(\mathbf{i}, \mathbf{j}) \in I^k \times I^{n-k}} \sum_{\substack{\mathbf{t} \in J^n : \\ \operatorname{std}(\mathbf{t}) = w^{-1}; \\ \mathbf{t} \in \mathbf{i} \sqcup\!\sqcup \widetilde{\mathbf{j}}}} (\mathbf{X}, \mathbf{Y})_{\mathbf{t}} \end{aligned} \tag{8.1.9}$$

(since for every $\mathbf{t} \in J^n$, there exists exactly one choice of $k \in \{0, 1, \ldots, n\}$ and $(\mathbf{i}, \mathbf{j}) \in I^k \times I^{n-k}$ satisfying $\mathbf{t} \in \mathbf{i} \sqcup\!\sqcup \widetilde{\mathbf{j}}$; namely, $\mathbf{i}$ is the restriction of $\mathbf{t}$ to the subalphabet I of J, whereas $\mathbf{j}$ is the restriction of $\mathbf{t}$ to $J \setminus I$, and k is the length of $\mathbf{i}$).

We now fix k and $(\mathbf{i}, \mathbf{j})$, and try to simplify the inner sum $\sum_{\substack{\mathbf{t} \in J^n : \\ \operatorname{std}(\mathbf{t}) = w^{-1}; \\ \mathbf{t} \in \mathbf{i} \sqcup\!\sqcup \widetilde{\mathbf{j}}}} (\mathbf{X}, \mathbf{Y})_{\mathbf{t}}$ on the right hand side of (8.1.9). First we notice that this sum is nonempty if and only if there exists some $\mathbf{t} \in \mathbf{i} \sqcup\!\sqcup \widetilde{\mathbf{j}}$ satisfying $\operatorname{std}(\mathbf{t}) = w^{-1}$. This existence is easily seen to be equivalent to $w^{-1} \in \operatorname{std}(\mathbf{i}) \sqcup\!\sqcup \operatorname{std}(\mathbf{j})[k]$ (since the standardization of any shuffle in $\mathbf{i} \sqcup\!\sqcup \widetilde{\mathbf{j}}$ is the corresponding shuffle in $\operatorname{std}(\mathbf{i}) \sqcup\!\sqcup \operatorname{std}(\mathbf{j})[k]$). This, in turn, is equivalent to $\operatorname{std}(\mathbf{i}) = (\operatorname{std}(w_1, \ldots, w_k))^{-1}$ and $\operatorname{std}(\mathbf{j}) = (\operatorname{std}(w_{k+1}, \ldots, w_n))^{-1}$ (according to the equivalence (i) $\Longleftrightarrow$ (ii) in Lemma 8.1.6). Hence, the inner sum on the right hand side of (8.1.9) is nonempty if and only if $\operatorname{std}(\mathbf{i}) = (\operatorname{std}(w_1, \ldots, w_k))^{-1}$ and $\operatorname{std}(\mathbf{j}) = (\operatorname{std}(w_{k+1}, \ldots, w_n))^{-1}$. When it is nonempty, it has only one addend[368], and this addend is $(\mathbf{X}, \mathbf{Y})_{\mathbf{t}} = \mathbf{X}_{\mathbf{i}} \mathbf{Y}_{\mathbf{j}}$ (since $\mathbf{t} \in \mathbf{i} \sqcup\!\sqcup \widetilde{\mathbf{j}}$). Summarizing, we see that the inner sum on the right hand side of (8.1.9) equals $\mathbf{X}_{\mathbf{i}} \mathbf{Y}_{\mathbf{j}}$ when $\operatorname{std}(\mathbf{i}) = (\operatorname{std}(w_1, \ldots, w_k))^{-1}$ and $\operatorname{std}(\mathbf{j}) = (\operatorname{std}(w_{k+1}, \ldots, w_n))^{-1}$, and is empty otherwise. Thus, (8.1.9) simplifies to

$$\begin{aligned} \Delta F_w &= \sum_{k=0}^{n} \sum_{\substack{(\mathbf{i}, \mathbf{j}) \in I^k \times I^{n-k} : \\ \operatorname{std}(\mathbf{i}) = (\operatorname{std}(w_1, \ldots, w_k))^{-1} \\ \operatorname{std}(\mathbf{j}) = (\operatorname{std}(w_{k+1}, \ldots, w_n))^{-1}}} \mathbf{X}_{\mathbf{i}} \mathbf{Y}_{\mathbf{j}} \\ &= \sum_{k=0}^{n} F_{\operatorname{std}(w_1, \ldots, w_k)}(\mathbf{X}) F_{\operatorname{std}(w_{k+1}, \ldots, w_n)}(\mathbf{Y}) \\ &= \sum_{k=0}^{n} F_{\operatorname{std}(w_1, \ldots, w_k)} \otimes F_{\operatorname{std}(w_{k+1}, \ldots, w_n)} \in \operatorname{FQSym} \otimes \operatorname{FQSym}. \end{aligned}$$

This proves (8.1.2), and thus the first assertion of Proposition 8.1.9.

From this, it is easy to derive that Δ satisfies coassociativity (i.e., the diagram (1.2.1) holds for $C = \operatorname{FQSym}$). (Alternatively, one can obtain this from the associativity of multiplication using Corollary 8.1.11.) We have already verified the rule (8.1.2). The connected graded structure on FQSym gives a counit and an antipode for free. □

Exercise 8.1.10. We say that a permutation $w \in \mathfrak{S}_n$ is *connected* if n is a positive integer and if there exists no $i \in \{1, 2, \ldots, n-1\}$ satisfying $f(\{1, 2, \ldots, i\}) = \{1, 2, \ldots, i\}$. Let $\mathfrak{CS}$ denote the set of all connected permutations of all $n \in \mathbb{N}$. Show that FQSym is a free (noncommutative) $\mathbf{k}$-algebra with generators $(F_w)_{w \in \mathfrak{CS}}$. (This statement means that $(F_{w_1} F_{w_2} \cdots F_{w_k})_{k \in \mathbb{N};\ (w_1, w_2, \ldots, w_k) \in \mathfrak{CS}^k}$ is a basis of the $\mathbf{k}$-module FQSym.)

[**Hint:** This is a result of Poirier and Reutenauer [172, Theorem 2.1]; it is much easier than the similar Theorem 6.4.3.]

[368] In fact, the elements $\operatorname{std}(\mathbf{t})$ for $\mathbf{t} \in \mathbf{i} \sqcup\!\sqcup \widetilde{\mathbf{j}}$ are distinct, and thus only one of them can equal w^{-1}.

Corollary 8.1.11. *The Hopf algebra* FQSym *is self-dual: Let* $\{G_w\}$ *be the dual* $\mathbf{k}$*-basis to the* $\mathbf{k}$*-basis* $\{F_w\}$ *for* FQSym. *Then, the* $\mathbf{k}$*-linear map sending* $G_w \longmapsto F_{w^{-1}}$ *is a Hopf algebra isomorphism* $\operatorname{FQSym}^o \longrightarrow \operatorname{FQSym}$.

Proof. For any $0 \le k \le n$, any $u \in \mathfrak{S}_k$ and any $v \in \mathfrak{S}_{n-k}$, one has

$$F_{u^{-1}} F_{v^{-1}} = \sum_{w^{-1} \in u^{-1} \sqcup v^{-1}[k]} F_{w^{-1}} = \sum_{\substack{w \in \mathfrak{S}_n: \\ \operatorname{std}(w_1,\ldots,w_k)=u \\ \operatorname{std}(w_{k+1},\ldots,w_n)=v}} F_{w^{-1}}$$

via the equivalence of (i) and (ii) in Lemma 8.1.6. On the other hand, in FQSym^o, the dual $\mathbf{k}$-basis $\{G_w\}$ to the $\mathbf{k}$-basis $\{F_w\}$ for FQSym should have product formula

$$G_u G_v = \sum_{\substack{w \in \mathfrak{S}_n: \\ \operatorname{std}(w_1,\ldots,w_k)=u \\ \operatorname{std}(w_{k+1},\ldots,w_n)=v}} G_w$$

coming from the coproduct formula (8.1.2) for FQSym in the $\{F_w\}$-basis. Comparing these equalities, we see that the $\mathbf{k}$-linear map τ sending $G_w \longmapsto F_{w^{-1}}$ is an isomorphism $\operatorname{FQSym}^o \longrightarrow \operatorname{FQSym}$ of $\mathbf{k}$-algebras. Hence, the adjoint $\tau^* : \operatorname{FQSym}^o \to (\operatorname{FQSym}^o)^o$ of this map is an isomorphism of $\mathbf{k}$-coalgebras. But identifying $(\operatorname{FQSym}^o)^o$ with FQSym in the natural way (since FQSym is of finite type), we easily see that $\tau^* = \tau$, whence τ itself is an isomorphism of both $\mathbf{k}$-algebras and $\mathbf{k}$-coalgebras, hence of $\mathbf{k}$-bialgebras, hence of Hopf algebras. □

We can now be a bit more precise about the relations between the various algebras

$$\Lambda, \operatorname{QSym}, \operatorname{NSym}, \operatorname{FQSym}, R\langle \mathbf{X} \rangle, R(\mathbf{x}).$$

Not only does FQSym allow one to *lift* the Hopf structure of QSym, it dually allows one to *extend* the Hopf structure of NSym. To set up this duality, note that Corollary 8.1.11 motivates the choice of an inner product on FQSym in which

$$(F_u, F_v) := \delta_{u^{-1},v}.$$

We wish to identify the images of the ribbon basis $\{R_\alpha\}$ of NSym when included in FQSym.

Definition 8.1.12. For any composition α, define an element $\mathbf{R}_\alpha$ of FQSym by

$$\mathbf{R}_\alpha := \sum_{\substack{w \in \mathfrak{S}_{|\alpha|}: \\ \operatorname{Des}(w)=D(\alpha)}} F_{w^{-1}} = \sum_{\substack{(w,\mathbf{i}): \\ w \in \mathfrak{S}_{|\alpha|}; \\ \operatorname{Des}(w)=D(\alpha); \\ \operatorname{std}(\mathbf{i})=w}} \mathbf{X}_\mathbf{i} = \sum_{\mathbf{i}:\operatorname{Des}(\mathbf{i})=D(\alpha)} \mathbf{X}_\mathbf{i},$$

where the *descent set* of a sequence $\mathbf{i} = (i_1, \ldots, i_n)$ is defined by

$$\operatorname{Des}(\mathbf{i}) := \{j \in \{1, 2, \ldots, n-1\} : i_j > i_{j+1}\} = \operatorname{Des}(\operatorname{std}(\mathbf{i})).$$

Alternatively,

$$\mathbf{R}_\alpha = \sum_T \mathbf{X}_T \tag{8.1.10}$$

in which the sum is over column-strict tableaux of the ribbon skew shape $\operatorname{Rib}(\alpha)$, and $\mathbf{X}_T = \mathbf{X}_\mathbf{i}$ in which $\mathbf{i}$ is the sequence of entries of T read in order from the southwest toward the northeast.

Example 8.1.13. Taking $\alpha = (1, 3, 2)$, with ribbon shape and column-strict fillings T as shown:

$$\operatorname{Rib}(\alpha) = \begin{array}{cccc} & & \square & \square \\ \square & \square & \square & \\ \square & & & \end{array} \qquad \text{and} \qquad T = \begin{array}{ccccccc} & & & & i_5 & \le & i_6 \\ & & & & \wedge & & \\ i_2 & \le & i_3 & \le & i_4 & & \\ \wedge & & & & & & \\ i_1 & & & & & & \end{array}$$

one has that

$$\mathbf{R}_{(1,3,2)} = \sum_{\substack{\mathbf{i}=(i_1,i_2,i_3,i_4,i_5,i_6): \\ \operatorname{Des}(\mathbf{i})=D(\alpha)=\{1,4\}}} \mathbf{X}_\mathbf{i} = \sum_{i_1 > i_2 \le i_3 \le i_4 > i_5 \le i_6} X_{i_1} X_{i_2} X_{i_3} X_{i_4} X_{i_5} X_{i_6} = \sum_T \mathbf{X}_T.$$

Corollary 8.1.14. *For every $n \in \mathbb{N}$ and $w \in \mathfrak{S}_n$, we let $\gamma(w)$ denote the unique composition α of n satisfying $D(\alpha) = \operatorname{Des}(w)$.*

(a) *The $\mathbf{k}$-linear map*

$$\begin{array}{ccc} \operatorname{FQSym} & \overset{\pi}{\twoheadrightarrow} & \operatorname{QSym}, \\ F_w & \longmapsto & L_{\gamma(w)} \end{array}$$

is a surjective Hopf algebra homomorphism.

(b) *The $\mathbf{k}$-linear map*

$$\begin{array}{ccc} \operatorname{NSym} & \overset{\iota}{\hookrightarrow} & \operatorname{FQSym}, \\ R_\alpha & \longmapsto & \mathbf{R}_\alpha \end{array}$$

is an injective Hopf algebra homomorphism.

(c) *The linear maps π and ι are adjoint maps with respect to the above choice of inner product on* FQSym *and the usual dual pairing between* NSym *and* QSym.

Now, consider the abelianization map $\operatorname{ab} : R\langle \mathbf{X} \rangle \twoheadrightarrow R(\mathbf{x})$ *defined as the continuous $\mathbf{k}$-algebra homomorphism sending the noncommutative variable X_i to the commutative x_i.*

(d) *The map π is a restriction of* ab.

(e) *The map ι lets one factor the surjection* $\operatorname{NSym} \twoheadrightarrow \Lambda$ *as follows:*

$$\begin{array}{ccccc} \operatorname{NSym} & \rightarrow & \operatorname{FQSym} \hookrightarrow R\langle \mathbf{X} \rangle & \overset{\operatorname{ab}}{\rightarrow} & R(\mathbf{x}), \\ R_\alpha & \longmapsto & \mathbf{R}_\alpha & \longmapsto & s_{\operatorname{Rib}(\alpha)}(\mathbf{x}). \end{array}$$

Proof. Given $n \in \mathbb{N}$, each composition α of n can be written in the form $\gamma(w)$ for some $w \in \mathfrak{S}_n$.[369] Hence, each fundamental quasisymmetric function L_α lies in the image of π. Thus, π is surjective.

Also, for each $n \in \mathbb{N}$ and $\alpha \in \operatorname{Comp}_n$, the element $\mathbf{R}_\alpha$ is a nonempty sum of noncommutative monomials (nonempty because α can be written in the form $\gamma(w)$ for some $w \in \mathfrak{S}_n$). Moreover, the elements $\mathbf{R}_\alpha$ for varying n and α are supported on disjoint monomials. Thus, these elements are linearly independent. Hence, the map ι is injective.

(d) Let $\mathfrak{A}$ denote the totally ordered set $\{1 < 2 < 3 < \cdots\}$ of positive integers. For each word $w = (w_1, w_2, \ldots, w_n) \in \mathfrak{A}^n$, we define a monomial $\mathbf{x}_w$ in $\mathbf{k}[[\mathbf{x}]]$ by $\mathbf{x}_w = x_{w_1} x_{w_2} \cdots x_{w_n}$.

Let $n \in \mathbb{N}$ and $\sigma \in \mathfrak{S}_n$. Then,

$$L_{\gamma(\sigma)} = \sum_{\substack{w \in \mathfrak{A}^n; \\ \operatorname{std} w = \sigma^{-1}}} \mathbf{x}_w$$

(by Lemma 5.3.6). But (8.1.3) (applied to $w = \sigma$) yields

$$F_\sigma = \sum_{\substack{\mathbf{i} = (i_1, \ldots, i_n): \\ \operatorname{std}(\mathbf{i}) = \sigma^{-1}}} \mathbf{X}_{\mathbf{i}} = \sum_{\substack{w \in \mathfrak{A}^n; \\ \operatorname{std} w = \sigma^{-1}}} \mathbf{X}_w$$

and thus

$$\operatorname{ab}(F_\sigma) = \operatorname{ab}\left(\sum_{\substack{w \in \mathfrak{A}^n; \\ \operatorname{std} w = \sigma^{-1}}} \mathbf{X}_w \right) = \sum_{\substack{w \in \mathfrak{A}^n; \\ \operatorname{std} w = \sigma^{-1}}} \underbrace{\operatorname{ab}(\mathbf{X}_w)}_{= \mathbf{x}_w} = \sum_{\substack{w \in \mathfrak{A}^n; \\ \operatorname{std} w = \sigma^{-1}}} \mathbf{x}_w = L_{\gamma(\sigma)} = \pi(F_\sigma).$$

We have shown this for all $n \in \mathbb{N}$ and $\sigma \in \mathfrak{S}_n$. Thus, π is a restriction of ab. This proves Corollary 8.1.14(d).

[369] Indeed, write our composition α as $(\alpha_1, \alpha_2, \ldots, \alpha_k)$. Then, we can pick w to be the permutation whose first α_1 entries are the largest α_1 elements of $\{1, 2, \ldots, n\}$ in increasing order; whose next α_2 entries are the next-largest α_2 elements of $\{1, 2, \ldots, n\}$ in increasing order; and so on. This permutation w will satisfy $\operatorname{Des}(w) = \{\alpha_1, \alpha_1 + \alpha_2, \ldots, \alpha_1 + \alpha_2 + \cdots + \alpha_{k-1}\} = D(\alpha)$ and thus $\gamma(w) = \alpha$.

(a) Let $n \in \mathbb{N}$ and $w = (w_1, w_2, \ldots, w_n) \in \mathfrak{S}_n$. Let α be the composition $\gamma(w)$ of n. Thus, the definition of π yields $\pi(F_w) = L_\alpha$. But applying the map $\pi \otimes \pi$ to the equality (8.1.2), we obtain

$$\begin{aligned} (\pi \otimes \pi)(\Delta F_w) &= (\pi \otimes \pi)\left(\sum_{k=0}^{n} F_{\operatorname{std}(w_1,w_2,\ldots,w_k)} \otimes F_{\operatorname{std}(w_{k+1},w_{k+2},\ldots,w_n)}\right) \\ &= \sum_{k=0}^{n} \pi\left(F_{\operatorname{std}(w_1,w_2,\ldots,w_k)}\right) \otimes \pi\left(F_{\operatorname{std}(w_{k+1},w_{k+2},\ldots,w_n)}\right) \\ &= \sum_{k=0}^{n} L_{\gamma(\operatorname{std}(w_1,w_2,\ldots,w_k))} \otimes L_{\gamma(\operatorname{std}(w_{k+1},w_{k+2},\ldots,w_n))} \qquad (8.1.11) \end{aligned}$$

(by the definition of π). Now, for each $k \in \{0, 1, \ldots, n\}$, the two compositions $\gamma(\operatorname{std}(w_1, w_2, \ldots, w_k))$ $\gamma(\operatorname{std}(w_{k+1}, w_{k+2}, \ldots, w_n))$ form a pair (β, γ) of compositions satisfying[370] either $\beta \cdot \gamma = \alpha$ or $\beta \odot \gamma = \alpha$, and in fact they form the only such pair satisfying $|\beta| = k$ and $|\gamma| = n - k$. Thus, the right hand side of (8.1.11) can be rewritten as

$$\sum_{\substack{(\beta,\gamma):\\ \beta\cdot\gamma=\alpha \text{ or } \beta\odot\gamma=\alpha}} L_\beta \otimes L_\gamma.$$

But this sum is ΔL_α, as we know from (5.2.5). Hence, (8.1.11) becomes

$$(\pi \otimes \pi)(\Delta F_w) = \Delta L_\alpha = \Delta(\pi(F_w)) \qquad (\text{since } L_\alpha = \pi(F_w)).$$

We have proven this for each $n \in \mathbb{N}$ and $w \in \mathfrak{S}_n$. Thus, we have proven that $(\pi \otimes \pi) \circ \Delta_{\mathrm{FQSym}} = \Delta_{\mathrm{QSym}} \circ \pi$. Combined with $\epsilon_{\mathrm{FQSym}} = \epsilon_{\mathrm{QSym}} \circ \pi$ (which is easy to check), this shows that π is a coalgebra homomorphism.

We can similarly see that π is an algebra homomorphism by checking that it respects the product (compare (5.2.6) and (8.1.1)). However, this also follows trivially from Corollary 8.1.14(d).

Thus, π is a bialgebra morphism, and therefore a Hopf algebra morphism (by Corollary 1.4.27). This proves Corollary 8.1.14(a).

(c) For any composition α and any $w \in \mathfrak{S}$, we have

$$\begin{aligned} (\iota(R_\alpha), F_w) = (\mathbf{R}_\alpha, F_w) = \sum_{u:\operatorname{Des}(u)=D(\alpha)} (F_{u^{-1}}, F_w) = \begin{cases} 1, & \text{if } \operatorname{Des}(w) = D(\alpha); \\ 0, & \text{otherwise} \end{cases} = \begin{cases} 1, & \text{if } \gamma(w) = \alpha; \\ 0, & \text{otherwise} \end{cases} \\ = (R_\alpha, L_{\gamma(w)}) = (R_\alpha, \pi(F_w)). \end{aligned}$$

Thus, the maps π and ι are adjoint. This proves Corollary 8.1.14(c).

(b) Again, there are several ways to prove this. Here is one:

First, note that $\iota(1) = 1$ (because $R_\varnothing = 1$ and $\mathbf{R}_\varnothing = 1$). Next, let α and β be two nonempty compositions. Let $m = |\alpha|$ and $n = |\beta|$. Then, $R_\alpha R_\beta = R_{\alpha\cdot\beta} + R_{\alpha\odot\beta}$ (by (5.4.11)) and thus

$$\begin{aligned} \iota(R_\alpha R_\beta) = \iota(R_{\alpha\cdot\beta} + R_{\alpha\odot\beta}) &= \underbrace{\iota(R_{\alpha\cdot\beta})}_{=\mathbf{R}_{\alpha\cdot\beta}=\sum_{\mathbf{i}:\operatorname{Des}(\mathbf{i})=D(\alpha\cdot\beta)} \mathbf{X}_\mathbf{i}} + \underbrace{\iota(R_{\alpha\odot\beta})}_{=\mathbf{R}_{\alpha\odot\beta}=\sum_{\mathbf{i}:\operatorname{Des}(\mathbf{i})=D(\alpha\odot\beta)} \mathbf{X}_\mathbf{i}} \\ &= \sum_{\mathbf{i}:\operatorname{Des}(\mathbf{i})=D(\alpha\cdot\beta)} \mathbf{X}_\mathbf{i} + \sum_{\mathbf{i}:\operatorname{Des}(\mathbf{i})=D(\alpha\odot\beta)} \mathbf{X}_\mathbf{i} = \sum_{\mathbf{i}:\operatorname{Des}(\mathbf{i})=D(\alpha\cdot\beta) \text{ or } \operatorname{Des}(\mathbf{i})=D(\alpha\odot\beta)} \mathbf{X}_\mathbf{i} \\ &= \sum_{\substack{\mathbf{i}=(i_1,i_2,\ldots,i_{m+n}):\\ \operatorname{Des}(i_1,i_2,\ldots,i_m)=D(\alpha) \text{ and}\\ \operatorname{Des}(i_{m+1},i_{m+2},\ldots,i_{m+n})=D(\beta)}} \mathbf{X}_\mathbf{i} \qquad (8.1.12) \end{aligned}$$

(since the words $\mathbf{i}$ of length $m + n$ satisfying $\operatorname{Des}(\mathbf{i}) = D(\alpha \cdot \beta)$ or $\operatorname{Des}(\mathbf{i}) = D(\alpha \odot \beta)$ are precisely the words $\mathbf{i} = (i_1, i_2, \ldots, i_{m+n})$ satisfying $\operatorname{Des}(i_1, i_2, \ldots, i_m) = D(\alpha)$ and $\operatorname{Des}(i_{m+1}, i_{m+2}, \ldots, i_{m+n}) = D(\beta)$). But choosing a word $\mathbf{i} = (i_1, i_2, \ldots, i_{m+n})$ satisfying $\operatorname{Des}(i_1, i_2, \ldots, i_m) = D(\alpha)$ and $\operatorname{Des}(i_{m+1}, i_{m+2}, \ldots, i_{m+n}) =$

[370] See Definition 5.2.14 for the notation we are using.

$D(\beta)$ is tantamount to choosing a pair $(\mathbf{u}, \mathbf{v})$ of a word $\mathbf{u} = (i_1, i_2, \ldots, i_m)$ satisfying $\operatorname{Des}\mathbf{u} = D(\alpha)$ and a word $\mathbf{v} = (i_{m+1}, i_{m+2}, \ldots, i_{m+n})$ satisfying $\operatorname{Des}\mathbf{v} = D(\beta)$. Thus, (8.1.12) becomes

$$\begin{aligned}\iota(R_\alpha R_\beta) &= \sum_{\substack{\mathbf{i}=(i_1,i_2,\ldots,i_{m+n}):\\ \operatorname{Des}(i_1,i_2,\ldots,i_m)=D(\alpha) \text{ and}\\ \operatorname{Des}(i_{m+1},i_{m+2},\ldots,i_{m+n})=D(\beta)}} \mathbf{X_i} = \sum_{\mathbf{u}:\operatorname{Des}\mathbf{u}=D(\alpha)} \ \sum_{\mathbf{v}:\operatorname{Des}\mathbf{v}=D(\beta)} \mathbf{X_u X_v} \\ &= \underbrace{\left(\sum_{\mathbf{u}:\operatorname{Des}\mathbf{u}=D(\alpha)} \mathbf{X_u} \right)}_{=\mathbf{R}_\alpha=\iota(R_\alpha)} \underbrace{\left(\sum_{\mathbf{v}:\operatorname{Des}\mathbf{v}=D(\beta)} \mathbf{X_v} \right)}_{=\mathbf{R}_\beta=\iota(R_\beta)} = \iota(R_\alpha)\,\iota(R_\beta).\end{aligned}$$

Thus, we have proven the equality $\iota(R_\alpha R_\beta) = \iota(R_\alpha)\,\iota(R_\beta)$ whenever α and β are two nonempty compositions. It also holds if we drop the "nonempty" requirement (since $R_\varnothing = 1$ and $\iota(1) = 1$). Thus, the $\mathbf{k}$-linear map ι respects the multiplication. Since $\iota(1) = 1$, this shows that ι is a $\mathbf{k}$-algebra homomorphism.

For each $n \in \mathbb{N}$, we let id_n be the identity permutation in $\mathfrak{S}_n$. Next, we observe that each $n \in \mathbb{N}$ satisfies $H_n = R_{(n)}$ (this follows, e.g., from (5.4.9), because the composition (n) is coarsened only by itself). Hence, each $n \in \mathbb{N}$ satisfies

$$\begin{aligned}\iota(H_n) = \iota\left(R_{(n)}\right) = \mathbf{R}_{(n)} &= \sum_{\substack{w\in\mathfrak{S}_n:\\ \operatorname{Des}(w)=D((n))}} F_{w^{-1}} \\ &= F_{\operatorname{id}_n^{-1}} \qquad \text{(since the only } w \in \mathfrak{S}_n \text{ satisfying } \operatorname{Des}(w) = D((n)) \text{ is } \operatorname{id}_n) \\ &= F_{\operatorname{id}_n}.\end{aligned} \tag{8.1.13}$$

In order to show that ι is a $\mathbf{k}$-coalgebra homomorphism, it suffices to check the equalities $(\iota \otimes \iota) \circ \Delta_{\mathrm{NSym}} = \Delta_{\mathrm{FQSym}} \circ \iota$ and $\epsilon_{\mathrm{NSym}} = \epsilon_{\mathrm{FQSym}} \circ \iota$. We shall only prove the first one, since the second is easy. Since ι, Δ_{NSym} and Δ_{FQSym} are $\mathbf{k}$-algebra homomorphisms, it suffices to check it on the generators $H_1, H_2, H_3, \ldots$ of NSym. But on these generators, it follows from comparing

$$\begin{aligned}((\iota \otimes \iota) \circ \Delta_{\mathrm{NSym}})(H_n) &= (\iota \otimes \iota)(\Delta_{\mathrm{NSym}} H_n) = (\iota \otimes \iota)\left(\sum_{i+j=n} H_i \otimes H_j\right) \qquad \text{(by (5.4.2))} \\ &= \sum_{i+j=n} \underbrace{\iota(H_i)}_{\substack{=F_{\operatorname{id}_i} \\ \text{(by (8.1.13))}}} \otimes \underbrace{\iota(H_j)}_{\substack{=F_{\operatorname{id}_j} \\ \text{(by (8.1.13))}}} = \sum_{i+j=n} F_{\operatorname{id}_i} \otimes F_{\operatorname{id}_j} = \sum_{k=0}^{n} F_{\operatorname{id}_k} \otimes F_{\operatorname{id}_{n-k}}\end{aligned}$$

with

$$\begin{aligned}(\Delta_{\mathrm{FQSym}} \circ \iota)(H_n) &= \Delta_{\mathrm{FQSym}}(\iota(H_n)) = \Delta_{\mathrm{FQSym}}(F_{\operatorname{id}_n}) \qquad \text{(by (8.1.13))} \\ &= \sum_{k=0}^{n} F_{\operatorname{id}_k} \otimes F_{\operatorname{id}_{n-k}} \qquad \text{(by (8.1.2))}.\end{aligned}$$

Thus, we know that ι is a $\mathbf{k}$-algebra homomorphism and a $\mathbf{k}$-coalgebra homomorphism. Hence, ι is a bialgebra morphism, and therefore a Hopf algebra morphism (by Corollary 1.4.27). This proves Corollary 8.1.14(b).

An alternative proof of Corollary 8.1.14(b) can be obtained by adjointness from Corollary 8.1.14(a). Both the inner product on FQSym and the dual pairing $(\cdot,\cdot) : \mathrm{NSym} \otimes \mathrm{QSym} \to \mathbf{k}$ respect the Hopf structures (i.e., the maps Δ_{NSym} and m_{QSym} are mutually adjoint with respect to these forms, and so are the maps m_{NSym} and Δ_{QSym}, and the maps Δ_{FQSym} and m_{FQSym}, and so on). Corollary 8.1.14(c) shows that the map ι is adjoint to the map π with respect to these two bilinear forms. Hence, we have a commutative diagram

$$\begin{tikzcd}
\mathrm{NSym} \arrow[r, hook, "\iota"] \arrow[d, "\cong"] & \mathrm{FQSym} \arrow[d, "\cong"] \\
\mathrm{QSym}^o \arrow[r, "\pi^*"'] & \mathrm{FQSym}^o
\end{tikzcd}$$

of Hopf algebras (where the two vertical arrows are the isomorphisms induced by the two bilinear forms). Thus, Corollary 8.1.14(b) follows from Corollary 8.1.14(a) by duality.

(e) For each composition α, the abelianization map ab sends the noncommutative tableau monomial $\mathbf{X}_T$ to the commutative tableau monomial $\mathbf{x}_T$ whenever T is a tableau of ribbon shape $\operatorname{Rib}(\alpha)$. Thus, ab sends $\mathbf{R}_\alpha$ to $s_{\operatorname{Rib}(\alpha)}(\mathbf{x})$ (because of the formula (8.1.10)). Hence, the composition $\mathrm{NSym} \to \mathrm{FQSym} \hookrightarrow R\langle\mathbf{X}\rangle \overset{\mathrm{ab}}{\to} R(\mathbf{x})$ does indeed send R_α to $s_{\operatorname{Rib}(\alpha)}(\mathbf{x})$. But so does the projection $\pi : \mathrm{NSym} \to \Lambda$, according to Theorem 5.4.10(b). Hence, the composition factors the projection. This proves Corollary 8.1.14(e). □

We summarize some of this picture as follows:

$$\begin{tikzcd}
\mathrm{FQSym} \arrow[r, dash, dotted, "\text{dual}"] & \mathrm{FQSym} \arrow[d, "\pi"] \\
\mathrm{NSym} \arrow[u, hook, "\iota"'] \arrow[r, dash, dotted, "\text{dual}"] \arrow[d, "\pi"] & \mathrm{QSym} \\
\Lambda \arrow[r, dash, dotted, "\text{dual}"] & \Lambda \arrow[u, hook]
\end{tikzcd}$$

Furthermore, if we denote by ι the canonical inclusion $\Lambda \to \mathrm{QSym}$ as well, then the diagram

$$\begin{tikzcd}
& \mathrm{FQSym} \arrow[dr, "\pi"] & \\
\mathrm{NSym} \arrow[ur, hook, "\iota"'] \arrow[dr, "\pi"] & & \mathrm{QSym} \\
& \Lambda \arrow[ur, hook, "\iota"'] &
\end{tikzcd}$$

is commutative (according to Corollary 8.1.14(e)).

Remark 8.1.15. Different notations for FQSym appear in the literature. In the book [24] (which presents an unusual approach to the character theory of the symmetric group using FQSym), the Hopf algebra FQSym is called $\mathcal{P}$, and its basis that we call $\{G_w\}_{w\in\mathfrak{S}_n}$ is denoted $\{w\}_{w\in\mathfrak{S}_n}$. In [93, Chapter 7], the Hopf algebra FQSym and its basis $\{F_w\}_{w\in\mathfrak{S}_n}$ are denoted MPR and $\{w\}_{w\in\mathfrak{S}_n}$, respectively.

9. Further topics

The following is a list of topics that were, at one point, planned to be touched in class, but did not make the cut. They might get elaborated upon in a future version of these notes.

9.0.1. 0-*Hecke algebras.*

- **Review of representation theory of finite-dimensional algebras.**
  Review the notions of indecomposables, simples, projectives, along with the theorems of Krull-Remak-Schmidt, of Jordan-Hölder, and the two kinds of Grothendieck groups dual to each other.
- **0-Hecke algebra representation theory.**
  Describe the simples and projectives, following Denton, Hivert, Schilling, Thiery [49] on $\mathcal{J}$-trivial monoids.
- **Nsym and Qsym as Grothendieck groups.**
  Give Krob and Thibon's interpretation (see [216, §5] for a brief summary) of
  - QSym and the Grothendieck group of composition series, and
  - NSym and the Grothendieck group of projectives.

  Remark 9.0.1. Mention P. McNamara's interpretation, in the case of *supersolvable lattices*, of the Ehrenborg quasisymmetric function as the composition series enumerator for an $H_n(0)$-action on the maximal chains

9.0.2. *Aguiar-Bergeron-Sottile character theory Part II: Odd and even characters, subalgebras.*

9.0.3. *Face enumeration, Eulerian posets, and cd-indices.* Borrowing from Billera's ICM notes [19].

- f-vectors, h-vectors
- flag f-vectors, flag h-vectors
- ab-indices and cd-indices

9.0.4. *Other topics.*

- Loday-Ronco Hopf algebra of planar binary trees [137]
- Poirier-Reutenauer Hopf algebra of tableaux
- Reading Hopf algebra of Baxter permutations
- Hopf monoids, e.g. of Hopf algebra of generalized permutohedra, of matroids, of graphs, Stanley chromatic symmetric functions and Tutte polynomials
- Lam-Pylyavskyy Hopf algebra of set-valued tableaux
- Connes-Kreimer Hopf algebra and renormalization
- Noncommutative symmetric functions and $\Omega\Sigma\mathbb{C}P^{\infty}$
- Maschke's theorem and "integrals" for Hopf algebras
- Nichols-Zoeller structure theorem and group-like elements
- Cartier-Milnor-Moore structure theorem and primitive elements
- Quasi-triangular Hopf algebras and quantum groups
- The Steenrod algebra, its dual, and tree Hopf algebras
- Ringel-Hall algebras of quivers
- Ellis-Khovanov odd symmetric function Hopf algebras [67] (see also Lauda-Russell [123])

Student talks given in class were:

(1) Al Garver, on Maschke's theorem for finite-dimensional Hopf algebras
(2) Jonathan Hahn, on the paper by Humpert and Martin.
(3) Emily Gunawan, on the paper by Lam, Lauve and Sottile.
(4) Jonas Karlsson, on the paper by Connes and Kreimer
(5) Thomas McConville, on Butcher's group and generalized Runge-Kutta methods.
(6) Cihan Bahran, on universal enveloping algebras and the Poincaré-Birkhoff-Witt theorem.
(7) Theodosios Douvropolos, on the Cartier-Milnor-Moore theorem.
(8) Alex Csar, on the Loday-Ronco Hopf algebra of binary trees
(9) Kevin Dilks, on Reading's Hopf algebra of (twisted) Baxter permutations
(10) Becky Patrias, on the paper by Lam and Pylyavskyy
(11) Meng Wu, on multiple zeta values and Hoffman's homomorphism from QSym

10. Some open problems and conjectures

- Is there a proof of the Assaf-McNamara skew Pieri rule that gives a resolution of Specht or Schur/Weyl modules whose character corresponds to $s_{\lambda/\mu}h_n$, whose terms model their alternating sum?
- Explicit antipodes in the Lam-Pylyavskyy Hopf algebras? (Answered by Patrias in [170].)
- P. McNamara's question [152, Question 7.1]: are P-partition enumerators irreducible for connected posets P?
- Stanley's question: are the only P-partition enumerators which are symmetric (not just quasisymmetric) those for which P is a skew shape with a column-strict labelling?
- Does Stanley's chromatic symmetric function distinguish trees?
- Hoffman's stuffle conjecture
- Billera-Brenti's nonnegativity conjecture for the total cd-index of Bruhat intervals ([20, Conjecture 6.1])

11. Appendix: Some basics

In this appendix, we briefly discuss some basic notions from linear algebra and elementary combinatorics that are used in these notes.

11.1. Linear expansions and triangularity.

In this Section, we shall recall some fundamental results from linear algebra (most importantly, the notions of a change-of-basis matrix and of a unitriangular matrix), but in greater generality than how it is usually done in textbooks. We shall use these results later when studying bases of combinatorial Hopf algebras; but per se, this section has nothing to do with Hopf algebras.

11.1.1. *Matrices.*

Let us first define the notion of a matrix whose rows and columns are indexed by arbitrary objects (as opposed to numbers):[371]

Definition 11.1.1. Let S and T be two sets. An *$S \times T$-matrix over* $\mathbf{k}$ shall mean a family $\left(a_{s,t}\right)_{(s,t)\in S\times T} \in \mathbf{k}^{S\times T}$ of elements of $\mathbf{k}$ indexed by elements of $S \times T$. Thus, the set of all $S\times T$-matrices over $\mathbf{k}$ is $\mathbf{k}^{S\times T}$.

We shall abbreviate "$S \times T$-matrix over $\mathbf{k}$" by "$S\times T$-matrix" when the value of $\mathbf{k}$ is clear from the context.

This definition of $S \times T$-matrices generalizes the usual notion of matrices (i.e., the notion of $n \times m$-matrices): Namely, if $n \in \mathbb{N}$ and $m \in \mathbb{N}$, then the $\{1, 2, \ldots, n\} \times \{1, 2, \ldots, m\}$-matrices are precisely the $n \times m$-matrices (in the usual meaning of this word). We shall often use the word "*matrix*" for both the usual notion of matrices and for the more general notion of $S \times T$-matrices.

Various concepts defined for $n \times m$-matrices (such as addition and multiplication of matrices, or the notion of a row) can be generalized to $S \times T$-matrices in a straightforward way. The following four definitions are examples of such generalizations:

Definition 11.1.2. Let S and T be two sets.

(a) The sum of two $S \times T$-matrices is defined by $\left(a_{s,t}\right)_{(s,t)\in S\times T} + \left(b_{s,t}\right)_{(s,t)\in S\times T} = \left(a_{s,t} + b_{s,t}\right)_{(s,t)\in S\times T}$.

(b) If $u \in \mathbf{k}$ and if $\left(a_{s,t}\right)_{(s,t)\in S\times T} \in \mathbf{k}^{S\times T}$, then we define $u\left(a_{s,t}\right)_{(s,t)\in S\times T}$ to be the $S \times T$-matrix $\left(ua_{s,t}\right)_{(s,t)\in S\times T}$.

(c) Let $A = \left(a_{s,t}\right)_{(s,t)\in S\times T}$ be an $S \times T$-matrix. For every $s \in S$, we define the *s-th row of* A to be the $\{1\} \times T$-matrix $\left(a_{s,t}\right)_{(i,t)\in\{1\}\times T}$. (Notice that $\{1\} \times T$-matrices are a generalization of row vectors.) Similarly, for every $t \in T$, we define the *t-th column of* A to be the $S \times \{1\}$-matrix $\left(a_{s,t}\right)_{(s,i)\in S\times\{1\}}$.

Definition 11.1.3. Let S be a set.

(a) The *$S \times S$ identity matrix* is defined to be the $S \times S$-matrix $\left(\delta_{s,t}\right)_{(s,t)\in S\times S}$. This $S \times S$-matrix is denoted by I_S. (Notice that the $n \times n$ identity matrix I_n is $I_{\{1,2,\ldots,n\}}$ for each $n \in \mathbb{N}$.)

(b) An $S \times S$-matrix $\left(a_{s,t}\right)_{(s,t)\in S\times S}$ is said to be *diagonal* if every $(s,t) \in S \times T$ satisfying $s \neq t$ satisfies $a_{s,t} = 0$.

(c) Let $A = \left(a_{s,t}\right)_{(s,t)\in S\times S}$ be an $S \times S$-matrix. The *diagonal* of A means the family $\left(a_{s,s}\right)_{s\in S}$. The *diagonal entries* of A are the entries of this diagonal $\left(a_{s,s}\right)_{s\in S}$.

Definition 11.1.4. Let S, T and U be three sets. Let $A = \left(a_{s,t}\right)_{(s,t)\in S\times T}$ be an $S \times T$-matrix, and let $B = \left(b_{t,u}\right)_{(t,u)\in T\times U}$ be a $T \times U$-matrix. Assume that the sum $\sum_{t\in T} a_{s,t}b_{t,u}$ is well-defined for every $(s,u) \in S \times U$. (For example, this is guaranteed to hold if the set T is finite. For infinite T, it may and may not hold.) Then, the $S \times U$-matrix AB is defined by

$$AB = \left(\sum_{t\in T} a_{s,t}b_{t,u}\right)_{(s,u)\in S\times U}.$$

Definition 11.1.5. Let S and T be two finite sets. We say that an $S\times T$-matrix A is *invertible* if and only if there exists a $T \times S$-matrix B satisfying $AB = I_S$ and $BA = I_T$. In this case, this matrix B is unique; it is denoted by A^{-1} and is called the *inverse* of A.

[371] As before, $\mathbf{k}$ denotes a commutative ring.

The definitions that we have just given are straightforward generalizations of the analogous definitions for $n \times m$-matrices; thus, unsurprisingly, many properties of $n \times m$-matrices still hold for $S \times T$-matrices. For example:

Proposition 11.1.6. (a) *Let S and T be two sets. Let A be an $S \times T$-matrix. Then, $I_S A = A$ and $A I_T = A$.*

(b) *Let S, T and U be three sets such that T is finite. Let A and B be two $S \times T$-matrices. Let C be a $T \times U$-matrix. Then, $(A + B)\, C = AC + BC$.*

(c) *Let S, T, U and V be four sets such that T and U are finite. Let A be an $S \times T$-matrix. Let B be a $T \times U$-matrix. Let C be a $U \times V$-matrix. Then, $(AB)\, C = A\,(BC)$.*

The proof of Proposition 11.1.6 (and of similar properties that will be left unstated) is analogous to the proofs of the corresponding properties of $n \times m$-matrices.[372] As a consequence of these properties, it is easy to see that if S is any finite set, then $\mathbf{k}^{S \times S}$ is a $\mathbf{k}$-algebra.

In general, $S \times T$-matrices (unlike $n \times m$-matrices) do not have a predefined order on their rows and their columns. Thus, the classical notion of a triangular $n \times n$-matrix cannot be generalized to a notion of a "triangular $S \times S$-matrix" when S is just a set with no additional structure. However, when S is a poset, such a generalization can be made:

Definition 11.1.7. Let S be a poset. Let $A = (a_{s,t})_{(s,t) \in S \times S}$ be an $S \times S$-matrix.

(a) The matrix A is said to be *triangular* if and only if every $(s, t) \in S \times S$ which does not satisfy $t \leq s$ must satisfy $a_{s,t} = 0$. (Here, $\leq$ denotes the smaller-or-equal relation of the poset S.)

(b) The matrix A is said to be *unitriangular* if and only if A is triangular and has the further property that, for every $s \in S$, we have $a_{s,s} = 1$.

(c) The matrix A is said to be *invertibly triangular* if and only if A is triangular and has the further property that, for every $s \in S$, the element $a_{s,s}$ of $\mathbf{k}$ is invertible.

Of course, all three notions of "triangular", "unitriangular" and "invertibly triangular" depend on the partial order on S.

Clearly, every invertibly triangular $S \times S$-matrix is triangular. Also, every unitriangular $S \times S$-matrix is invertibly triangular (because the element 1 of $\mathbf{k}$ is invertible).

We can restate the definition of "invertibly triangular" as follows: The matrix A is said to be *invertibly triangular* if and only if it is triangular and its diagonal entries are invertible. Similarly, we can restate the definition of "unitriangular" as follows: The matrix A is said to be *unitriangular* if and only if it is triangular and all its diagonal entries equal 1.

Definition 11.1.7(a) generalizes both the notion of upper-triangular matrices and the notion of lower-triangular matrices. To wit:

Example 11.1.8. Let $n \in \mathbb{N}$. Let N_1 be the poset whose ground set is $\{1, 2, \ldots, n\}$ and whose smaller-or-equal relation $\leq_1$ is given by

$$s \leq_1 t \iff s \leq t \text{ (as integers).}$$

(This is the usual order relation on this set.) Let N_2 be the poset whose ground set is $\{1, 2, \ldots, n\}$ and whose order relation $\leq_2$ is given by

$$s \leq_2 t \iff s \geq t \text{ (as integers).}$$

Let $A \in \mathbf{k}^{n \times n}$.

(a) The matrix A is upper-triangular if and only if A is triangular when regarded as an $N_1 \times N_1$-matrix.

[372] A little **warning**: In Proposition 11.1.6(c), the condition that T and U be finite can be loosened (we leave this to the interested reader), but cannot be completely disposed of. It can happen that both $(AB)\, C$ and $A\,(BC)$ are defined, but $(AB)\, C = A\,(BC)$ does not hold (if we remove this condition). For example, this happens if $S = \mathbb{Z}$, $T = \mathbb{Z}$, $U = \mathbb{Z}$, $V = \mathbb{Z}$, $A = \left(\begin{cases} 1, & \text{if } i \geq j; \\ 0, & \text{if } i < j \end{cases} \right)_{(i,j) \in \mathbb{Z} \times \mathbb{Z}}$, $B = (\delta_{i,j} - \delta_{i,j+1})_{(i,j) \in \mathbb{Z} \times \mathbb{Z}}$ and $C = \left(\begin{cases} 0, & \text{if } i \geq j; \\ 1, & \text{if } i < j \end{cases} \right)_{(i,j) \in \mathbb{Z} \times \mathbb{Z}}$. (Indeed, in this example, it is easy to check that $AB = I_{\mathbb{Z}}$ and $BC = -I_{\mathbb{Z}}$ and thus $\underbrace{(AB)}_{= I_{\mathbb{Z}}} C = I_{\mathbb{Z}} C = C \neq -A = A \underbrace{(-I_{\mathbb{Z}})}_{= BC} = A\,(BC)$.)

This seeming paradox is due to the subtleties of rearranging infinite sums (similarly to how a conditionally convergent series of real numbers can change its value when its entries are rearranged).

(b) The matrix A is lower-triangular if and only if A is triangular when regarded as an $N_2 \times N_2$-matrix.

More interesting examples of triangular matrices are obtained when the order on S is not a total order:

Example 11.1.9. Let S be the poset whose ground set is $\{1, 2, 3\}$ and whose smaller relation $<_S$ is given by $1 <_S 2$ and $3 <_S 2$. Then, the triangular $S \times S$-matrices are precisely the 3×3-matrices of the form $\begin{pmatrix} a_{1,1} & 0 & 0 \\ a_{2,1} & a_{2,2} & a_{2,3} \\ 0 & 0 & a_{3,3} \end{pmatrix}$ with $a_{1,1}, a_{2,1}, a_{2,2}, a_{2,3}, a_{3,3} \in \mathbf{k}$.

We shall now state some basic properties of triangular matrices:

Proposition 11.1.10. *Let S be a finite poset.*

(a) *The triangular $S \times S$-matrices form a subalgebra of the $\mathbf{k}$-algebra $\mathbf{k}^{S\times S}$.*
(b) *The invertibly triangular $S \times S$-matrices form a group with respect to multiplication.*
(c) *The unitriangular $S \times S$-matrices form a group with respect to multiplication.*
(d) *Any invertibly triangular $S \times S$-matrix is invertible, and its inverse is again invertibly triangular.*
(e) *Any unitriangular $S \times S$-matrix is invertible, and its inverse is again unitriangular.*

Exercise 11.1.11. Prove Proposition 11.1.10.

11.1.2. *Expansion of a family in another.* We will often study situations where two families $(e_s)_{s\in S}$ and $(f_t)_{t\in T}$ of vectors in a $\mathbf{k}$-module M are given, and the vectors e_s can be written as linear combinations of the vectors f_t. In such situations, we can form an $S \times T$-matrix out of the coefficients of these linear combinations; this is one of the ways how matrices arise in the theory of modules. Let us define the notations we are going to use in such situations:

Definition 11.1.12. Let M be a $\mathbf{k}$-module. Let $(e_s)_{s\in S}$ and $(f_t)_{t\in T}$ be two families of elements of M. (The sets S and T may and may not be finite.)

Let $A = (a_{s,t})_{(s,t)\in S\times T}$ be an $S \times T$-matrix. Assume that, for every $s \in S$, all but finitely many $t \in T$ satisfy $a_{s,t} = 0$. (This assumption is automatically satisfied if T is finite.)

We say that the family $(e_s)_{s\in S}$ *expands in the family* $(f_t)_{t\in T}$ *through the matrix* A if

$$\text{every } s \in S \text{ satisfies } e_s = \sum_{t\in T} a_{s,t} f_t. \tag{11.1.1}$$

In this case, we furthermore say that the matrix A is a *change-of-basis matrix* (or *transition matrix*) from the family $(e_s)_{s\in S}$ to the family $(f_t)_{t\in T}$.

Remark 11.1.13. The notation in Definition 11.1.12 is not really standard; even we ourselves will occasionally deviate in its use. In the formulation "the family $(e_s)_{s\in S}$ expands in the family $(f_t)_{t\in T}$ through the matrix A", the word "in" can be replaced by "with respect to", and the word "through" can be replaced by "using".

The notion of a "change-of-basis matrix" is slightly misleading, because neither of the families $(e_s)_{s\in S}$ and $(f_t)_{t\in T}$ has to be a basis. Our use of the words "transition matrix" should not be confused with the different meaning that these words have in the theory of Markov chains. The indefinite article in "a change-of-basis matrix" is due to the fact that, for given families $(e_s)_{s\in S}$ and $(f_t)_{t\in T}$, there might be more than one change-of-basis matrix from $(e_s)_{s\in S}$ to $(f_t)_{t\in T}$. (There also might be no such matrix.) When $(e_s)_{s\in S}$ and $(f_t)_{t\in T}$ are bases of the $\mathbf{k}$-module M, there exists precisely one change-of-basis matrix from $(e_s)_{s\in S}$ to $(f_t)_{t\in T}$.

So a change-of-basis matrix $A = (a_{s,t})_{(s,t)\in S\times T}$ from one family $(e_s)_{s\in S}$ to another family $(f_t)_{t\in T}$ allows us to write the elements of the former family as linear combinations of the elements of the latter (using (11.1.1)). When such a matrix A is invertible (and the sets S and T are finite[373]), it also (indirectly) allows us to do the opposite: i.e., to write the elements of the latter family as linear combinations of the elements of the former. This is because if A is an invertible change-of-basis matrix from $(e_s)_{s\in S}$ to $(f_t)_{t\in T}$, then A^{-1} is a change-of-basis matrix from $(f_t)_{t\in T}$ to $(e_s)_{s\in S}$. This is part (a) of the following theorem:

[373] We are requiring the finiteness of S and T mainly for the sake of simplicity. We could allow S and T to be infinite, but then we would have to make some finiteness requirements on A and A^{-1}.

Theorem 11.1.14. *Let M be a $\mathbf{k}$-module. Let S and T be two finite sets. Let $(e_s)_{s\in S}$ and $(f_t)_{t\in T}$ be two families of elements of M.*

Let A be an invertible $S \times T$-matrix. Thus, A^{-1} is a $T \times S$-matrix.

Assume that the family $(e_s)_{s\in S}$ expands in the family $(f_t)_{t\in T}$ through the matrix A. Then:

(a) *The family $(f_t)_{t\in T}$ expands in the family $(e_s)_{s\in S}$ through the matrix A^{-1}.*
(b) *The $\mathbf{k}$-submodule of M spanned by the family $(e_s)_{s\in S}$ is the $\mathbf{k}$-submodule of M spanned by the family $(f_t)_{t\in T}$.*
(c) *The family $(e_s)_{s\in S}$ spans the $\mathbf{k}$-module M if and only if the family $(f_t)_{t\in T}$ spans the $\mathbf{k}$-module M.*
(d) *The family $(e_s)_{s\in S}$ is $\mathbf{k}$-linearly independent if and only if the family $(f_t)_{t\in T}$ is $\mathbf{k}$-linearly independent.*
(e) *The family $(e_s)_{s\in S}$ is a basis of the $\mathbf{k}$-module M if and only if the family $(f_t)_{t\in T}$ is a basis of the $\mathbf{k}$-module M.*

Exercise 11.1.15. Prove Theorem 11.1.14.

Definition 11.1.16. Let M be a $\mathbf{k}$-module. Let S be a finite poset. Let $(e_s)_{s\in S}$ and $(f_s)_{s\in S}$ be two families of elements of M.

(a) We say that the family $(e_s)_{s\in S}$ *expands triangularly in the family* $(f_s)_{s\in S}$ if and only if there exists a triangular $S \times S$-matrix A such that the family $(e_s)_{s\in S}$ expands in the family $(f_s)_{s\in S}$ through the matrix A.
(b) We say that the family $(e_s)_{s\in S}$ *expands invertibly triangularly in the family* $(f_s)_{s\in S}$ if and only if there exists an invertibly triangular $S \times S$-matrix A such that the family $(e_s)_{s\in S}$ expands in the family $(f_s)_{s\in S}$ through the matrix A.
(c) We say that the family $(e_s)_{s\in S}$ *expands unitriangularly in the family* $(f_s)_{s\in S}$ if and only if there exists a unitriangular $S \times S$-matrix A such that the family $(e_s)_{s\in S}$ expands in the family $(f_s)_{s\in S}$ through the matrix A.

Clearly, if the family $(e_s)_{s\in S}$ expands unitriangularly in the family $(f_s)_{s\in S}$, then it also expands invertibly triangularly in the family $(f_s)_{s\in S}$ (because any unitriangular matrix is an invertibly triangular matrix).

We notice that in Definition 11.1.16, the two families $(e_s)_{s\in S}$ and $(f_s)_{s\in S}$ must be indexed by one and the same set S.

The concepts of "expanding triangularly", "expanding invertibly triangularly" and "expanding unitriangularly" can also be characterized without referring to matrices, as follows:

Remark 11.1.17. Let M be a $\mathbf{k}$-module. Let S be a finite poset. Let $(e_s)_{s\in S}$ and $(f_s)_{s\in S}$ be two families of elements of M. Let $<$ denote the smaller relation of the poset S, and let $\leq$ denote the smaller-or-equal relation of the poset S. Then:

(a) The family $(e_s)_{s\in S}$ expands triangularly in the family $(f_s)_{s\in S}$ if and only if every $s \in S$ satisfies
$$e_s = (\text{a } \mathbf{k}\text{-linear combination of the elements } f_t \text{ for } t \in S \text{ satisfying } t \leq s).$$
(b) The family $(e_s)_{s\in S}$ expands invertibly triangularly in the family $(f_s)_{s\in S}$ if and only if every $s \in S$ satisfies
$$e_s = \alpha_s f_s + (\text{a } \mathbf{k}\text{-linear combination of the elements } f_t \text{ for } t \in S \text{ satisfying } t < s)$$
for some invertible $\alpha_s \in \mathbf{k}$.
(c) The family $(e_s)_{s\in S}$ expands unitriangularly in the family $(f_s)_{s\in S}$ if and only if every $s \in S$ satisfies
$$e_s = f_s + (\text{a } \mathbf{k}\text{-linear combination of the elements } f_t \text{ for } t \in S \text{ satisfying } t < s).$$

All three parts of Remark 11.1.17 follow easily from the definitions.

Example 11.1.18. Let $n \in \mathbb{N}$. For this example, let S be the poset $\{1, 2, \ldots, n\}$ (with its usual order). Let M be a $\mathbf{k}$-module, and let $(e_s)_{s\in S}$ and $(f_s)_{s\in S}$ be two families of elements of M. We shall identify these families $(e_s)_{s\in S}$ and $(f_s)_{s\in S}$ with the n-tuples $(e_1, e_2, \ldots, e_n)$ and $(f_1, f_2, \ldots, f_n)$. Then, the family $(e_s)_{s\in S} = (e_1, e_2, \ldots, e_n)$ expands triangularly in the family $(f_s)_{s\in S} = (f_1, f_2, \ldots, f_n)$ if and only if, for every $s \in \{1, 2, \ldots, n\}$, the vector e_s is a $\mathbf{k}$-linear combination of $f_1, f_2, \ldots, f_s$. Moreover, the family $(e_s)_{s\in S} = (e_1, e_2, \ldots, e_n)$ expands unitriangularly in the family $(f_s)_{s\in S} = (f_1, f_2, \ldots, f_n)$ if and only if, for every $s \in \{1, 2, \ldots, n\}$, the vector e_s is a sum of f_s with a $\mathbf{k}$-linear combination of $f_1, f_2, \ldots, f_{s-1}$.

Corollary 11.1.19. *Let M be a $\mathbf{k}$-module. Let S be a finite poset. Let $(e_s)_{s\in S}$ and $(f_s)_{s\in S}$ be two families of elements of M. Assume that the family $(e_s)_{s\in S}$ expands invertibly triangularly in the family $(f_s)_{s\in S}$. Then:*

(a) *The family $(f_s)_{s\in S}$ expands invertibly triangularly in the family $(e_s)_{s\in S}$.*
(b) *The $\mathbf{k}$-submodule of M spanned by the family $(e_s)_{s\in S}$ is the $\mathbf{k}$-submodule of M spanned by the family $(f_s)_{s\in S}$.*
(c) *The family $(e_s)_{s\in S}$ spans the $\mathbf{k}$-module M if and only if the family $(f_s)_{s\in S}$ spans the $\mathbf{k}$-module M.*
(d) *The family $(e_s)_{s\in S}$ is $\mathbf{k}$-linearly independent if and only if the family $(f_s)_{s\in S}$ is $\mathbf{k}$-linearly independent.*
(e) *The family $(e_s)_{s\in S}$ is a basis of the $\mathbf{k}$-module M if and only if the family $(f_s)_{s\in S}$ is a basis of the $\mathbf{k}$-module M.*

Exercise 11.1.20. Prove Remark 11.1.17 and Corollary 11.1.19.

An analogue of Corollary 11.1.19 can be stated for unitriangular expansions, but we leave this to the reader.

12. Further hints to the exercises (work in progress)

The following pages contain hints to (some of[374]) the exercises in the text (beyond the hints occasionally included in the exercises themselves). Some of the hints rise to the level of outlined solutions.

Detailed solutions (sometimes different from the solutions hinted at) can be found in Chapter 13.

Warning: The hints below are new and have never been proofread. Typos (or worse) are likely. In case of doubt, consult the detailed solutions.

12.1. Hints for Chapter 1.

Hint to Exercise 1.2.3. The claim of the exercise is dual to the classical fact that if A is a $\mathbf{k}$-module and $m : A \otimes A \to A$ is a $\mathbf{k}$-linear map, then there exists *at most one* $\mathbf{k}$-linear map $u : \mathbf{k} \to A$ such that the diagram (1.1.2) commutes[375]. Take any proof of this latter fact, rewrite it in an "element-free" fashion[376], and "reverse all arrows". This will yield a solution to Exercise 1.2.3.

For an alternative solution, use Sweedler notation (as in (1.2.3)) as follows: The commutativity of the diagram (1.2.2) says that

$$c = \sum_{(c)} \epsilon(c_1) c_2 = \sum_{(c)} \epsilon(c_2) c_1 \qquad \text{for each } c \in C.$$

Thus, if ϵ_1 and ϵ_2 are two $\mathbf{k}$-linear maps $\epsilon : C \to \mathbf{k}$ such that the diagram (1.2.2) commutes, then each $c \in C$ satisfies

$$c = \sum_{(c)} \epsilon_1(c_1) c_2 = \sum_{(c)} \epsilon_1(c_2) c_1$$

and

$$c = \sum_{(c)} \epsilon_2(c_1) c_2 = \sum_{(c)} \epsilon_2(c_2) c_1.$$

Apply ϵ_2 to both sides of the equality $c = \sum_{(c)} \epsilon_1(c_2) c_1$, and apply ϵ_1 to both sides of the equality $c = \sum_{(c)} \epsilon_2(c_1) c_2$. Compare the results, and conclude that $\epsilon_1 = \epsilon_2$.

Hint to Exercise 1.3.4. Part (a) is well-known, and part (b) is dual to part (a). So the trick is (again) to rewrite the classical proof of part (a) in an "element-free" way, and then "reversing all arrows". Alternatively, part (b) can be solved using Sweedler notation.

Hint to Exercise 1.3.6. Same method as for Exercise 1.3.4 above.

Hint to Exercise 1.3.13. (a) Use the following fact from linear algebra: If U, V, U' and V' are four $\mathbf{k}$-modules, and $\phi : U \to U'$ and $\psi : V \to V'$ are two surjective $\mathbf{k}$-linear maps, then the kernel of $\phi \otimes \psi : U \otimes V \to U' \otimes V'$ is

$$\ker(\phi \otimes \psi) = (\ker \phi) \otimes V + U \otimes (\ker \psi).$$

(b) The fact just mentioned also holds if we no longer require ϕ and ψ to be surjective, but instead require $\mathbf{k}$ to be a field.

Hint to Exercise 1.3.18. Let $f : V \to W$ be an invertible graded $\mathbf{k}$-linear map. Let $n \in \mathbb{N}$ and $w \in W_n$. Show that the n-th homogeneous component of $f^{-1}(w)$ is also a preimage of w under f, and thus must equal $f^{-1}(w)$. Therefore, $f^{-1}(w) \in W_n$.

Hint to Exercise 1.3.19. (a) Define the $\mathbf{k}$-linear map $\widetilde{\Delta} : A \to A \otimes A$ by $\widetilde{\Delta}(x) = \Delta(x) - (x \otimes 1 + 1 \otimes x)$. Argue that $\widetilde{\Delta}$ is graded, so its kernel $\ker \widetilde{\Delta}$ is a graded $\mathbf{k}$-submodule of A. But this kernel is precisely $\mathfrak{p}$.

(b) The hard part is to show that $\epsilon(\mathfrak{p}) = 0$. To do so, consider any $x \in \mathfrak{p}$, and apply the map $\epsilon \otimes \mathrm{id}$ to both sides of the equality $\Delta(x) = x \otimes 1 + 1 \otimes x$. The result simplifies to $x = \epsilon(x) \cdot 1_A + x$. Thus, $\epsilon(x) \cdot 1_A = 0$. Now apply ϵ to this, thus obtaining $\epsilon(x) = 0$.

[374] Currently only the ones from Chapter 1.

[375] This fact is just the linearization of the known fact that any binary operation has at most one neutral element.

[376] This means rewriting it completely in terms of linear maps rather than elements. For example, instead of talking about $m(m(a \otimes b) \otimes c)$ for three elements $a, b, c \in A$, you should talk about the map $m \circ (m \otimes \mathrm{id}_A) : A \otimes A \otimes A \to A$ (which is, of course, the map that sends each $a \otimes b \otimes c$ to $m(m(a \otimes b) \otimes c)$). Instead of computing with elements, you should compute with maps (and commutative diagrams).

Hint to Exercise 1.3.20. (a) This follows from $1_A \in A_0$, which is part of what it means for A to be a graded $\mathbf{k}$-algebra.

(b) Let $\epsilon' : A_0 \to \mathbf{k}$ be the restriction of the map ϵ to A_0. We know that ϵ' is surjective (since $\epsilon'(1_A) = 1_{\mathbf{k}}$), and that both A_0 and $\mathbf{k}$ are free $\mathbf{k}$-modules of rank 1 (since connectedness of A means $A_0 \cong \mathbf{k}$ as $\mathbf{k}$-modules). It is an an easy exercise in linear algebra to conclude from these facts that ϵ' is an isomorphism. Since $\epsilon' \circ u = \mathrm{id}_{\mathbf{k}}$, we thus conclude that $u : \mathbf{k} \to A_0$ is an isomorphism as well (from $\mathbf{k}$ to A_0).

(c) This follows from part (b).

(e) This follows from how we solved part (b).

(d) Since the bialgebra A is graded, the map ϵ must be graded. Thus, for each positive integer n, we have $\epsilon(A_n) \subset \mathbf{k}_n = 0$. This quickly yields $\epsilon(I) = 0$ (where $I = \bigoplus_{n>0} A_n$), hence $I \subset \ker \epsilon$. On the other hand, $\ker \epsilon \subset I$ can be shown as follows: Let $a \in \ker \epsilon$; write a in the form $a = a' + a''$ for some $a' \in A_0$ and some $a'' \in I$, and then argue that $0 = \epsilon(a) = \epsilon(a' + a'') = \epsilon(a') + \underbrace{\epsilon(a'')}_{\substack{=0 \\ (\text{since } a'' \in I \subset \ker \epsilon)}} = \epsilon(a')$, so that $a' = 0$ by part (e) and therefore $a \in I$.

(f) This is most intuitive with Sweedler notation: Let $x \in A$. Then, $\Delta(x) = \sum_{(x)} x_1 \otimes x_2$. Applying $\mathrm{id} \otimes \epsilon$ and recalling the commutativity of (1.2.2), we thus get $x = \sum_{(x)} \epsilon(x_2) x_1$. Thus,

$$\underbrace{\Delta(x)}_{=\sum_{(x)} x_1 \otimes x_2} - \underbrace{x}_{=\sum_{(x)} \epsilon(x_2) x_1} \otimes 1 = \sum_{(x)} x_1 \otimes x_2 - \sum_{(x)} \epsilon(x_2) x_1 \otimes 1$$
$$= \sum_{(x)} \underbrace{x_1}_{\in A} \otimes \underbrace{(x_2 - \epsilon(x_2) \cdot 1)}_{\substack{\in \ker \epsilon = I \\ (\text{by part (d)})}} \in A \otimes I.$$

(g) Let $x \in I$. Proceeding similarly to part (f), show that

$$\Delta(x) - 1 \otimes x - x \otimes 1 + \epsilon(x) 1 \otimes 1 = \sum_{(x)} \underbrace{(x_1 - \epsilon(x_1) \cdot 1)}_{\substack{\in \ker \epsilon = I \\ (\text{by part (d)})}} \otimes \underbrace{(x_2 - \epsilon(x_2) \cdot 1)}_{\substack{\in \ker \epsilon = I \\ (\text{by part (d)})}} \in I \otimes I.$$

Since $x \in I = \ker \epsilon$, the $\epsilon(x) 1 \otimes 1$ term on the left hand side vanishes.

(h) This follows from part (g), since a simple homogeneity argument shows that $(I \otimes I)_n = \sum_{k=1}^{n-1} A_k \otimes A_{n-k}$.

Hint to Exercise 1.3.24. We need to check the four equalities $D_q \circ m = m \circ (D_q \otimes D_q)$ and $D_q \circ u = u$ and $(D_q \otimes D_q) \circ \Delta = \Delta \circ D_q$ and $\epsilon \circ D_q = \epsilon$. This can easily be done by hand (just check everything on homogeneous elements); a more erudite proof proceeds as follows: Generalize the map D_q to a map $D_{q,V} : V \to V$ defined (in the same way as D_q) for every graded $\mathbf{k}$-module V, and show that these maps $D_{q,V}$ are functorial (i.e., if $f : V \to W$ is a graded $\mathbf{k}$-linear map between two graded $\mathbf{k}$-modules V and W, then $D_{q,W} \circ f = f \circ D_{q,V}$) and "respect tensor products" (i.e., we have $D_{q,V \otimes W} = D_{q,V} \otimes D_{q,W}$ for any two graded $\mathbf{k}$-modules V and W). The four equalities are then easily obtained from these two facts, without having to introduce elements.

Hint to Exercise 1.3.26. (a) Our definition of the $\mathbf{k}$-coalgebra $A \otimes B$ yields

$$\Delta_{A \otimes B} = (\mathrm{id}_A \otimes T_{A,B} \otimes \mathrm{id}_B) \circ (\Delta_A \otimes \Delta_B) \qquad \text{and} \qquad \epsilon_{A \otimes B} = \theta \circ (\epsilon_A \otimes \epsilon_B),$$

where θ is the canonical $\mathbf{k}$-module isomorphism $\mathbf{k} \otimes \mathbf{k} \to \mathbf{k}$. All maps on the right hand sides are $\mathbf{k}$-algebra homomorphisms (see Exercise 1.3.6(a)); thus, so are $\Delta_{A \otimes B}$ and $\epsilon_{A \otimes B}$.

(b) Straightforward.

Hint to Exercise 1.4.2. Simple computation (either element-free or with Sweedler notation).

Hint to Exercise 1.4.4. Simple computation (either element-free or with Sweedler notation).

Hint to Exercise 1.4.5. Straightforward computation, best done using Sweedler notation.

Hint to Exercise 1.4.15. Use Exercise 1.4.2.

Hint to Exercise 1.4.19. The following is more context than hint (see the last paragraph for an actual hint).

It is easiest to prove this by calculating with elements. To wit, in order to prove that two **k**-linear maps from $A^{\otimes(k+1)}$ are identical, it suffices to show that they agree on all pure tensors $a_1 \otimes a_2 \otimes \cdots \otimes a_{k+1} \in A^{\otimes(k+1)}$. But the recursive definition of $m^{(k)}$ shows that

$$m^{(k)}\left(a_1 \otimes a_2 \otimes \cdots \otimes a_{k+1}\right) = a_1\left(a_2\left(a_3\left(\cdots\left(a_k a_{k+1}\right)\cdots\right)\right)\right) \tag{12.1.1}$$

for all $a_1, a_2, \ldots, a_{k+1} \in A$. Now, the "general associativity" law (a fundamental result in abstract algebra, commonly used without mention) says that, because the multiplication of A is associative, the parentheses in the product $a_1\left(a_2\left(a_3\left(\cdots\left(a_k a_{k+1}\right)\cdots\right)\right)\right)$ can be omitted without making it ambiguous – i.e., any two ways of parenthesizing the product $a_1 a_2 \cdots a_{k+1}$ evaluate to the same result. (For example, for $k = 4$, this says that

$$a_1\left(a_2\left(a_3 a_4\right)\right) = a_1\left(\left(a_2 a_3\right) a_4\right) = \left(a_1 a_2\right)\left(a_3 a_4\right) = \left(a_1\left(a_2 a_3\right)\right) a_4 = \left(\left(a_1 a_2\right) a_3\right) a_4$$

for all $a_1, a_2, a_3, a_4 \in A$.) Thus, we can rewrite (12.1.1) as

$$m^{(k)}\left(a_1 \otimes a_2 \otimes \cdots \otimes a_{k+1}\right) = a_1 a_2 \cdots a_{k+1}.$$

Using this formula, all four parts of the exercise become trivial: For example, part (a) simply says that

$$a_1 a_2 \cdots a_{k+1} = \left(a_1 a_2 \cdots a_{i+1}\right)\left(a_{i+2} a_{i+3} \cdots a_{k+1}\right)$$

for all $a_1, a_2, \ldots, a_{k+1} \in A$, because we have

$$\left(m \circ \left(m^{(i)} \otimes m^{(k-1-i)}\right)\right)\left(a_1 \otimes a_2 \otimes \cdots \otimes a_{k+1}\right) = \left(a_1 a_2 \cdots a_{i+1}\right)\left(a_{i+2} a_{i+3} \cdots a_{k+1}\right).$$

Likewise, part (c) simply says that

$$a_1 a_2 \cdots a_{k+1} = a_1 a_2 \cdots a_i \left(a_{i+1} a_{i+2}\right) a_{i+3} a_{i+4} \cdots a_{k+1}$$

for all $a_1, a_2, \ldots, a_{k+1} \in A$. Parts (b) and (d) are particular cases of parts (a) and (c), respectively.

Of course, in order for this to be a complete solution, you have to prove the "general associativity" law used above. It turns out that doing so is not much easier than solving the exercise from scratch (in fact, part (a) of the exercise is an equivalent form of the "general associativity" law). So we can just as well start from scratch and solve part (a) directly by induction on k, then derive part (b) as its particular case, then solve part (c) by induction on k using the result of part (b), then derive part (d) as a particular case of (c).

Hint to Exercise 1.4.20. If you have solved Exercise 1.4.19 in an "element-free" way, then you can reverse all arrows in said solution and thus obtain a solution to Exercise 1.4.20.

Hint to Exercise 1.4.22. (a) Induction on k, using Exercise 1.3.6(b).

(b) This is dual to (a).

(d) For every **k**-coalgebra C, consider the map $\Delta_C^{(k)} : C \to C^{\otimes(k+1)}$ (this is the map $\Delta^{(k)}$ defined in Exercise 1.4.20). This map $\Delta_C^{(k)}$ is clearly functorial in C. By this we mean that if C and D are any two **k**-coalgebras, and $f : C \to D$ is any **k**-coalgebra homomorphism, then the diagram

$$\begin{tikzcd}
C \arrow[r, "f"] \arrow[d, "\Delta_C^{(k)}"] & D \arrow[d, "\Delta_D^{(k)}"] \\
C^{\otimes(k+1)} \arrow[r, "f^{\otimes(k+1)}"'] & D^{\otimes(k+1)}
\end{tikzcd}$$

commutes. Now, apply this to $C = H^{\otimes(\ell+1)}$, $D = H$ and $f = m_H^{(\ell)}$ (using part (a)).

(c) This is dual to (d).

Hint to Exercise 1.4.23. Induction on k.

Hint to Exercise 1.4.28. This is dual to Proposition 1.4.10, so the usual strategy (viz., rewriting element-free and reversing all arrows) applies.

Hint to Exercise 1.4.29. (a) A straightforward generalization of the proof of Proposition 1.4.10 (which corresponds to the particular case when $C = A$ and $r = \mathrm{id}$) does the trick.

(b) This is dual to (a).

(c) Easy.

(d) Apply Exercise 1.4.29(a) to $C = A$ and $r = \mathrm{id}_A$; then, apply Proposition 1.4.26(a) to $H = A$ and $\alpha = S$.

(e) Let $s : C \to A$ be the $\mathbf{k}$-linear map that sends every homogeneous element $c \in C_n$ (for every $n \in \mathbb{N}$) to the n-th homogeneous component of $r^{\star(-1)}(c)$. Then, s is graded, and (this takes some work) is also a $\star$-inverse to r. But r has only one $\star$-inverse.

Hint to Exercise 1.4.30. (a) Rewrite the assumption as $m \circ (P \otimes \mathrm{id}) \circ T \circ \Delta = u \circ \epsilon$, where T is the twist map $T_{A,A}$. Proposition 1.4.10 leads to $m \circ (S \otimes S) = S \circ m \circ T$ and $u = S \circ u$. Exercise 1.4.28 leads to $(S \otimes S) \circ \Delta = T \circ \Delta \circ S$ and $\epsilon \circ S = \epsilon$. Use these to show that $(P \circ S) \star S = u \circ \epsilon$, so that $P \circ S = \mathrm{id}$. Also, show that $S \star (S \circ P) = u \circ \epsilon$, so that $S \circ P = \mathrm{id}$.

(b) Similar to (a).

(c) Let A be a connected graded Hopf algebra. Just as a left $\star$-inverse S to id_A has been constructed in the proof of Proposition 1.4.16, we could construct a $\mathbf{k}$-linear map $P : A \to A$ such that every $a \in A$ satisfies $\sum_{(a)} P(a_2) \cdot a_1 = u(\epsilon(a))$. Now apply part (a).

Hint to Exercise 1.4.32. Since D is a direct summand of C, we can identify the tensor products $D \otimes C$, $C \otimes D$ and $D \otimes D$ with their canonical images inside $C \otimes C$. Now, we can show that $\Delta(D) \subset D \otimes D$ as follows: Let $p : C \to D$ be the canonical projection from C onto its direct summand D; then, $\Delta(D) \subset D \otimes C$ shows that $(p \otimes \mathrm{id}) \circ \Delta = \Delta$, and $\Delta(D) \subset C \otimes D$ shows that $(\mathrm{id} \otimes p) \circ \Delta = \Delta$. Hence,

$$\underbrace{(p \otimes p)}_{=(p\otimes \mathrm{id})\circ(\mathrm{id}\otimes p)} \circ \Delta = (p \otimes \mathrm{id}) \circ \underbrace{(\mathrm{id} \otimes p) \circ \Delta}_{=\Delta} = (p \otimes \mathrm{id}) \circ \Delta = \Delta.$$

This yields $\Delta(D) \subset D \otimes D$. Hence, we get a map $\Delta_D : D \to D \otimes D$ by restricting Δ. Obviously, the map $\epsilon : C \to \mathbf{k}$ restricts to a map $\epsilon_D : D \to \mathbf{k}$ as well. It remains to check the commutativity of the diagrams (1.2.1) and (1.2.2) for D instead of C; but this is inherited from C.

Hint to Exercise 1.4.33. (a) Let $\widetilde{f} = (\mathrm{id}_C \otimes f \otimes \mathrm{id}_C) \circ \Delta^{(2)} : C \to C \otimes U \otimes C$; then, $K = \ker \widetilde{f}$. Show (by manipulation of maps, using Exercise 1.4.20(b)) that $(\mathrm{id}_C \otimes \mathrm{id}_U \otimes \Delta) \circ \widetilde{f} = \left(\widetilde{f} \otimes \mathrm{id}_C\right) \circ \Delta$. Now,

$$K = \ker \widetilde{f} \subset \ker \left(\underbrace{(\mathrm{id}_C \otimes \mathrm{id}_U \otimes \Delta) \circ \widetilde{f}}_{=\left(\widetilde{f}\otimes \mathrm{id}_C\right)\circ\Delta} \right) = \ker \left(\left(\widetilde{f} \otimes \mathrm{id}_C \right) \circ \Delta \right) = \Delta^{-1} \left(\ker \left(\widetilde{f} \otimes \mathrm{id}_C \right) \right)$$

and therefore

$$\begin{aligned} \Delta(K) \subset \ker \left(\widetilde{f} \otimes \mathrm{id}_C \right) &= \underbrace{\left(\ker \widetilde{f}\right)}_{=K} \otimes C \qquad \text{(since tensoring over a field is left-exact)} \\ &= K \otimes C. \end{aligned}$$

Similarly, $\Delta(K) \subset C \otimes K$. Now, apply Exercise 1.4.32 to $D = K$.

(b) Let E be a $\mathbf{k}$-subcoalgebra of C which is a subset of $\ker f$. Then, $\Delta^{(2)}(E) \subset E \otimes E \otimes E$ (since E is a subcoalgebra) and $f(E) = 0$ (since $E \subset \ker f$). Now,

$$\begin{aligned} \left((\mathrm{id}_C \otimes f \otimes \mathrm{id}_C) \circ \Delta^{(2)} \right)(E) &= (\mathrm{id}_C \otimes f \otimes \mathrm{id}_C) \left(\underbrace{\Delta^{(2)}(E)}_{\subset E \otimes E \otimes E} \right) \\ &\subset (\mathrm{id}_C \otimes f \otimes \mathrm{id}_C)(E \otimes E \otimes E) \\ &= \mathrm{id}_C(E) \otimes \underbrace{f(E)}_{=0} \otimes \mathrm{id}_C(E) = 0. \end{aligned}$$

Hence, $E \subset \ker \left((\mathrm{id}_C \otimes f \otimes \mathrm{id}_C) \circ \Delta^{(2)} \right) = K$.

[*Remark:* Exercise 1.4.33(a) would not hold if we allowed $\mathbf{k}$ to be an arbitrary commutative ring rather than a field.]

Hint to Exercise 1.4.34. (a) Here is Takeuchi's argument: We know that the map $h\mid_{C_0}\in \operatorname{Hom}(C_0, A)$ is $\star$-invertible; let $\widetilde{g}$ be its $\star$-inverse. Extend $\widetilde{g}$ to a $\mathbf{k}$-linear map $g : C \to A$ by defining it as 0 on every C_n for $n > 0$. It is then easy to see that $(h \star g)\mid_{C_0} = (g \star h)\mid_{C_0} = (u\epsilon)\mid_{C_0}$. This allows us to assume WLOG that $h\mid_{C_0} = (u\epsilon)\mid_{C_0}$ (because once we know that $h \star g$ and $g \star h$ are $\star$-invertible, it follows that so is h). Assuming this, we conclude that $h - u\epsilon$ annihilates C_0. Define f as $h - u\epsilon$. Now, we can proceed as in the proof of Proposition 1.4.24 to show that $\sum_{k\geq 0} (-1)^k f^{\star k}$ is a well-defined linear map $C \to A$ and a two-sided $\star$-inverse for h. Thus, h is $\star$-invertible, and part (a) of the exercise is proven. (An alternative proof proceeds by mimicking the proof of Proposition 1.4.16, again by first assuming WLOG that $h\mid_{C_0} = (u\epsilon)\mid_{C_0}$.)

(b) Apply part (a) to $C = A$ and the map $\operatorname{id}_A : A \to A$.

(c) Applying part (b), we see that A is a Hopf algebra (since $A_0 = \mathbf{k}$ is a Hopf algebra) in the setting of Proposition 1.4.16. This yields the existence of the antipode. Its uniqueness is trivial, and its gradedness follows from Exercise 1.4.29(e).

Hint to Exercise 1.4.35. (a) Let I be a two-sided coideal of A such that $I \cap \mathfrak{p} = 0$ and such that $I = \bigoplus_{n\geq 0} (I \cap A_n)$. Let $I_n = I \cap A_n$ for every $n \in \mathbb{N}$. Then, $I = \bigoplus_{n\geq 0} I_n$. Since I is a two-sided coideal, we have $\epsilon(I) = 0$.

We want to prove that $I = 0$. It clearly suffices to show that every $n \in \mathbb{N}$ satisfies $I_n = 0$ (since $I = \bigoplus_{n\geq 0} I_n$). We shall show this by strong induction: We fix an $N \in \mathbb{N}$, and we assume (as induction hypothesis) that $I_n = 0$ for all $n < N$. We must prove that $I_N = 0$.

Fix $i \in I_N$; we aim to show that $i = 0$. We have $i \in I_N \subset A_N$ and thus $\Delta(i) \in (A \otimes A)_N$ (since Δ is a graded map). On the other hand, from $i \in I_N \subset I$, we obtain

$$\begin{aligned}\Delta(i) \in \Delta(I) &\subset \underbrace{I}_{=\bigoplus_{n\geq 0} I_n} \otimes \underbrace{A}_{=\bigoplus_{m\geq 0} A_m} + \underbrace{A}_{=\bigoplus_{m\geq 0} A_m} \otimes \underbrace{I}_{=\bigoplus_{n\geq 0} I_n} \qquad \text{(since I is a two-sided coideal)}\\ &= \sum_{(m,n)\in\mathbb{N}^2} I_n \otimes A_m + \sum_{(m,n)\in\mathbb{N}^2} A_m \otimes I_n.\end{aligned}$$

Combining this with $\Delta(i) \in (A \otimes A)_N$, we obtain

$$\begin{aligned}\Delta(i) &\in \sum_{\substack{(m,n)\in\mathbb{N}^2;\\ m+n=N}} I_n \otimes A_m + \sum_{\substack{(m,n)\in\mathbb{N}^2;\\ m+n=N}} A_m \otimes I_n \qquad \left(\text{since $I_n \otimes A_m$ and $A_m \otimes I_n$ are subsets of } (A \otimes A)_{n+m}\right)\\ &= \sum_{n=0}^{N} I_n \otimes A_{N-n} + \sum_{n=0}^{N} A_{N-n} \otimes I_n\\ &= I_N \otimes \underbrace{A_0}_{=\mathbf{k}\cdot 1_A} + \sum_{n=0}^{N-1} \underbrace{I_n}_{\substack{=0\\ \text{(by the induction}\\ \text{hypothesis)}}} \otimes A_{N-n} + \underbrace{A_0}_{=\mathbf{k}\cdot 1_A} \otimes I_N + \sum_{n=0}^{N-1} A_{N-n} \otimes \underbrace{I_n}_{\substack{=0\\ \text{(by the induction}\\ \text{hypothesis)}}}\\ &= I_N \otimes (\mathbf{k} \cdot 1_A) + (\mathbf{k} \cdot 1_A) \otimes I_N.\end{aligned}$$

In other words,

$$\Delta(i) = j \otimes 1_A + 1_A \otimes k \tag{12.1.2}$$

for some $j, k \in I_N$. By applying $\epsilon \otimes \operatorname{id}$ to both sides of this equality, and recalling the commutativity of (1.2.2), we obtain $i = \epsilon(j) 1_A + k$. But $\epsilon(j) = 0$ (since $j \in I_N \subset I$, so $\epsilon(j) \in \epsilon(I) = 0$), so this simplifies to $i = k$. Similarly, $i = j$. Hence, (12.1.2) rewrites as $\Delta(i) = i \otimes 1_A + 1_A \otimes i$, which shows that $i \in \mathfrak{p}$, hence $i \in I \cap \mathfrak{p} = 0$ and thus $i = 0$. This was for proved for each $i \in I_N$, so we obtain $I_N = 0$. This completes the induction step, and so part (a) is solved.

(b) Exercise 1.3.13(a) shows that $\ker f$ is a two-sided coideal of C. If $f\mid_{\mathfrak{p}}$ is injective, then $(\ker f) \cap \mathfrak{p} = 0$. Now, apply part (a) of the current exercise to $I = \ker f$.

(c) Proceed as in part (b), but use Exercise 1.3.13(b) instead of Exercise 1.3.13(a).

Hint to Exercise 1.5.4. (a) Straightforward (if slightly laborious) computations.

(b) Direct verification (the hard part of which has been done in (1.3.7) already).

(c) For every subset S of a $\mathbf{k}$-module U, we let $\langle S \rangle$ denote the $\mathbf{k}$-submodule of U spanned by S. Our definition of J thus becomes

$$J = T(\mathfrak{p}) \cdot C \cdot T(\mathfrak{p}), \tag{12.1.3}$$

where $C = \langle xy - yx - [x, y] \mid x, y \in \mathfrak{p} \rangle$. A simple computation shows that each element of C is primitive. Hence,

$$\Delta(C) \subset C \otimes T(\mathfrak{p}) + T(\mathfrak{p}) \otimes C.$$

Applying Δ to both sides of (12.1.3), and recalling that Δ is a $\mathbf{k}$-algebra homomorphism, we find

$$\begin{aligned} \Delta(J) &= \underbrace{\Delta(T(\mathfrak{p}))}_{\subset T(\mathfrak{p}) \otimes T(\mathfrak{p})} \cdot \underbrace{\Delta(C)}_{\subset C \otimes T(\mathfrak{p}) + T(\mathfrak{p}) \otimes C} \cdot \underbrace{\Delta(T(\mathfrak{p}))}_{\subset T(\mathfrak{p}) \otimes T(\mathfrak{p})} \\ &\subset (T(\mathfrak{p}) \otimes T(\mathfrak{p})) \cdot (C \otimes T(\mathfrak{p}) + T(\mathfrak{p}) \otimes C) \cdot (T(\mathfrak{p}) \otimes T(\mathfrak{p})) \\ &= J \otimes T(\mathfrak{p}) + T(\mathfrak{p}) \otimes J. \end{aligned}$$

A similar (but simpler) argument shows $\epsilon(J) = 0$. Thus, J is a two-sided coideal of $T(\mathfrak{p})$. This yields that $T(\mathfrak{p})/J$ is a $\mathbf{k}$-bialgebra.

(d) We need to show that $S(J) \subset J$. This can be done in a similar way as we proved $\Delta(J) \subset J \otimes T(\mathfrak{p}) + T(\mathfrak{p}) \otimes J$ in part (c), once you know (from Proposition 1.4.10) that the antipode S of $T(\mathfrak{p})$ is a $\mathbf{k}$-algebra anti-homomorphism.

Hint to Exercise 1.5.5. Straightforward and easy verification.

Hint to Exercise 1.5.6. Straightforward and easy verification. Parts (a) and (b) are dual, of course.

Hint to Exercise 1.5.8. (a) Straightforward and easy verification.

(b) The dual says the following: Let A and B be two $\mathbf{k}$-coalgebras, at least one of which is cocommutative. Prove that the $\mathbf{k}$-coalgebra anti-homomorphisms from A to B are the same as the $\mathbf{k}$-coalgebra homomorphisms from A to B.

Hint to Exercise 1.5.9. For every $1 \leq i < j \leq k$, let $t_{i,j}$ be the transposition in $\mathfrak{S}_k$ which transposes i with j. It is well-known that the symmetric group $\mathfrak{S}_k$ is generated by the transpositions $t_{i,i+1}$ with i ranging over $\{1, 2, \ldots, k-1\}$. However, we have $(\rho(\pi)) \circ (\rho(\psi)) = \rho(\pi\psi)$ for any two elements π and ψ of $\mathfrak{S}_k$. Thus, it suffices to check that

$$m^{(k-1)} \circ (\rho(t_{i,i+1})) = m^{(k-1)} \qquad \text{for all } i \in \{1, 2, \ldots, k-1\}.$$

But this is not hard to check using $m^{(k-1)} = m^{(k-2)} \circ (\mathrm{id}_{A^{\otimes(i-1)}} \otimes m \otimes \mathrm{id}_{A^{\otimes(k-1-i)}})$ (a consequence of Exercise 1.4.19(c)) and $m \circ T = m$.

Hint to Exercise 1.5.10. Here is the dual statement: Let C be a cocommutative $\mathbf{k}$-coalgebra, and let $k \in \mathbb{N}$. The symmetric group $\mathfrak{S}_k$ acts on the k-fold tensor power $C^{\otimes k}$ by permuting the tensor factors: $\sigma(v_1 \otimes v_2 \otimes \cdots \otimes v_k) = v_{\sigma^{-1}(1)} \otimes v_{\sigma^{-1}(2)} \otimes \cdots \otimes v_{\sigma^{-1}(k)}$ for all $v_1, v_2, \ldots, v_k \in C$ and $\sigma \in \mathfrak{S}_k$. For every $\pi \in \mathfrak{S}_k$, denote by $\rho(\pi)$ the action of π on $C^{\otimes k}$ (this is an endomorphism of $C^{\otimes k}$). Show that every $\pi \in \mathfrak{S}_k$ satisfies $(\rho(\pi)) \circ \Delta^{(k-1)} = \Delta^{(k-1)}$. (Recall that $\Delta^{(k-1)} : C \to C^{\otimes k}$ is defined as in Exercise 1.4.20 for $k \geq 1$, and by $\Delta^{(-1)} = \epsilon : C \to \mathbf{k}$ for $k = 0$.)

Hint to Exercise 1.5.11. (a) Use Exercise 1.5.6(b) and Exercise 1.3.6(a) to represent $f \star g$ as a composition of three $\mathbf{k}$-algebra homomorphisms.

(b) Induction on k, using part (a).

(c) Use Proposition 1.4.10, Proposition 1.4.26(a) and the easy fact that a composition of a $\mathbf{k}$-algebra homomorphism with a $\mathbf{k}$-algebra anti-homomorphism (in either order) always is a $\mathbf{k}$-algebra anti-homomorphism.

(d) Use Exercise 1.5.6(b). Then, proceed by induction on k as in the solution of Exercise 1.4.22(a).

(e) Use Proposition 1.4.3.

(f) Let H be a commutative $\mathbf{k}$-bialgebra. Let k and ℓ be two nonnegative integers. Then, Exercise 1.5.11(b) (applied to $A = H$ and $f_i = \mathrm{id}_H$) yields that $\mathrm{id}_H^{\star k}$ is a $\mathbf{k}$-algebra homomorphism $H \to H$. Now, apply Exercise 1.5.11(e) to H, H, H, H, ℓ, id_H, $\mathrm{id}_H^{\star k}$ and id_H instead of C, C', A, A', k, f_i, α and γ.

(g) This is an exercise in bootstrapping. First, let $k \in \mathbb{N}$. Then, part (b) of this exercise shows that $\mathrm{id}_H^{\star k}$ is a $\mathbf{k}$-algebra homomorphism. Use this together with part (c) to conclude that $\mathrm{id}_H^{\star k} \circ S$ is again a $\mathbf{k}$-algebra homomorphism and a $\star$-inverse to $\mathrm{id}_H^{\star k}$; thus, $\mathrm{id}_H^{\star k} \circ S = \left(\mathrm{id}_H^{\star k}\right)^{\star(-1)} = \mathrm{id}_H^{\star(-k)}$, and this map $\mathrm{id}_H^{\star(-k)}$ is a $\mathbf{k}$-algebra homomorphism.

Now forget that we fixed k. We thus have shown that $\mathrm{id}_H^{\star k}$ and $\mathrm{id}_H^{\star(-k)}$ are $\mathbf{k}$-algebra homomorphisms for each $k \in \mathbb{N}$. In other words,

$$\mathrm{id}_H^{\star k} \text{ is a } \mathbf{k}\text{-algebra homomorphism} \qquad \text{for every } k \in \mathbb{Z}. \tag{12.1.4}$$

Furthermore, we have proved the equality $\mathrm{id}_H^{\star k} \circ S = \mathrm{id}_H^{\star(-k)}$ for each $k \in \mathbb{N}$. Repeating the proof of this, but now taking $k \in \mathbb{Z}$ instead of $k \in \mathbb{N}$, we conclude that it also holds for each $k \in \mathbb{Z}$ (since we already have proved (12.1.4)). In other words,

$$\mathrm{id}_H^{\star(-k)} = \mathrm{id}_H^{\star k} \circ S \qquad \text{for every } k \in \mathbb{Z}. \tag{12.1.5}$$

Now, fix two integers k and ℓ. From (12.1.4), we know that $\mathrm{id}_H^{\star k}$ is a $\mathbf{k}$-algebra homomorphism. Hence, if ℓ is nonnegative, then we can prove $\mathrm{id}_H^{\star k} \circ \mathrm{id}_H^{\star \ell} = \mathrm{id}_H^{\star(k\ell)}$ just as we did in the solution to Exercise 1.5.11(f). But the case when ℓ is negative can be reduced to the previous case by applying (12.1.5) (once to $-\ell$ instead of k, and once again to $-k\ell$ instead of k). Thus, in each case, we obtain $\mathrm{id}_H^{\star k} \circ \mathrm{id}_H^{\star \ell} = \mathrm{id}_H^{\star(k\ell)}$.

(h) The dual of Exercise 1.5.11(a) is the following exercise:

> If H is a $\mathbf{k}$-bialgebra and C is a cocommutative $\mathbf{k}$-coalgebra, and if f and g are two $\mathbf{k}$-coalgebra homomorphisms $C \to H$, then prove that $f \star g$ also is a $\mathbf{k}$-coalgebra homomorphism $C \to H$.

The dual of Exercise 1.5.11(b) is the following exercise:

> If H is a $\mathbf{k}$-bialgebra and C is a cocommutative $\mathbf{k}$-coalgebra, and if $f_1, f_2, \ldots, f_k$ are several $\mathbf{k}$-coalgebra homomorphisms $C \to H$, then prove that $f_1 \star f_2 \star \cdots \star f_k$ also is a $\mathbf{k}$-coalgebra homomorphism $C \to H$.

The dual of Exercise 1.5.11(c) is the following exercise:

> If H is a Hopf algebra and C is a cocommutative $\mathbf{k}$-coalgebra, and if $f : C \to H$ is a $\mathbf{k}$-coalgebra homomorphism, then prove that $S \circ f : C \to H$ (where S is the antipode of H) is again a $\mathbf{k}$-coalgebra homomorphism, and is a $\star$-inverse to f.

The dual of Exercise 1.5.11(d) is the following exercise:

> If C is a cocommutative $\mathbf{k}$-coalgebra, then show that $\Delta^{(k)}$ is a $\mathbf{k}$-coalgebra homomorphism for every $k \in \mathbb{N}$. (The map $\Delta^{(k)} : C \to C^{\otimes(k+1)}$ is defined as in Exercise 1.4.20.)

The dual of Exercise 1.5.11(e) is Exercise 1.5.11(e) itself (up to renaming objects and maps).

The dual of Exercise 1.5.11(f) is the following exercise:

> If H is a cocommutative $\mathbf{k}$-bialgebra, and k and ℓ are two nonnegative integers, then prove that $\mathrm{id}_H^{\star \ell} \circ \mathrm{id}_H^{\star k} = \mathrm{id}_H^{\star(\ell k)}$.

The dual of Exercise 1.5.11(g) is the following exercise:

> If H is a cocommutative $\mathbf{k}$-Hopf algebra, and k and ℓ are two integers, then prove that $\mathrm{id}_H^{\star \ell} \circ \mathrm{id}_H^{\star k} = \mathrm{id}_H^{\star(\ell k)}$.

Hint to Exercise 1.5.13. This is dual to Corollary 1.4.12 (but can also easily be shown using Exercise 1.4.29(b), Exercise 1.5.8(b) and Proposition 1.4.26(b)).

Hint to Exercise 1.5.14. (a) This can be proved computationally (using Sweedler notation), but there is a nicer argument as well:

A *coderivation* of a $\mathbf{k}$-coalgebra (C, Δ, ϵ) is defined as a $\mathbf{k}$-linear map $F : C \to C$ such that $\Delta \circ F = (F \otimes \mathrm{id} + \mathrm{id} \otimes F) \circ \Delta$. (The reader can check that this axiom is the result of writing the axiom for a derivation in element-free terms and reversing all arrows. Nothing less should be expected.) It is easy to see that E is a coderivation. Hence, it will be enough to check that $(S \star f)(a)$ and $(f \star S)(a)$ are primitive whenever $f : A \to A$ is a coderivation and $a \in A$. So fix a coderivation $f : A \to A$. Notice that the antipode S of A is a coalgebra anti-endomorphism (by Exercise 1.4.28), thus a coalgebra endomorphism (by Exercise 1.5.8(b)).

Thus, $\Delta \circ S = (S \otimes S) \circ \Delta$. Moreover, $\Delta : A \to A \otimes A$ is a coalgebra homomorphism (by Exercise 1.5.6(a)) and an algebra homomorphism (since A is a bialgebra). Applying (1.4.2) to $A \otimes A$, A, A, Δ, id_A, S and f instead of A', C, C', α, γ, f and g, we obtain

$$\begin{aligned}
\Delta \circ (S \star f) &= \underbrace{(\Delta \circ S)}_{=(S \otimes S)\circ\Delta} \star \underbrace{(\Delta \circ f)}_{\substack{=(f\otimes \mathrm{id} + \mathrm{id} \otimes f)\circ\Delta \\ (\text{since } f \text{ is a coderivation})}} \\
&= ((S \otimes S) \circ \Delta) \star ((f \otimes \mathrm{id} + \mathrm{id} \otimes f) \circ \Delta) = ((S \otimes S) \star (f \otimes \mathrm{id} + \mathrm{id} \otimes f)) \circ \Delta \\
&= \underbrace{((S \otimes S) \star (f \otimes \mathrm{id}))}_{\substack{=(S\star f)\otimes(S\star\mathrm{id}) \\ (\text{by Exercise 1.4.4(a)})}} \circ \Delta + \underbrace{((S \otimes S) \star (\mathrm{id} \otimes f))}_{\substack{=(S\star\mathrm{id})\otimes(S\star f) \\ (\text{by Exercise 1.4.4(a)})}} \circ \Delta \\
&= \left((S \star f) \otimes \underbrace{(S \star \mathrm{id})}_{=u\epsilon} \right) \circ \Delta + \left(\underbrace{(S \star \mathrm{id})}_{=u\epsilon} \otimes (S \star f) \right) \circ \Delta \\
&= ((S \star f) \otimes u\epsilon) \circ \Delta + (u\epsilon \otimes (S \star f)) \circ \Delta.
\end{aligned}$$

Hence, every $a \in A$ satisfies

$$\begin{aligned}
(\Delta \circ (S \star f))(a) &= (((S \star f) \otimes u\epsilon) \circ \Delta + (u\epsilon \otimes (S \star f)) \circ \Delta)(a) \\
&= (S \star f)(a) \otimes 1 + 1 \otimes (S \star f)(a)
\end{aligned}$$

(after some brief computations using (1.2.2)). In other words, for every $a \in A$, the element $(S \star f)(a)$ is primitive. Similarly the same can be shown for $(f \star S)(a)$, and so we are done.

(b) is a very simple computation. (Alternatively, the $(S \star E)(p) = E(p)$ part follows from applying part (c) to $a = 1$, and similarly one can show $(E \star S)(p) = E(p)$.)

(c) This is another computation, using Proposition 1.4.17 and the (easy) observation that E is a derivation of the algebra A.

(d) Assume that the graded algebra $A = \bigoplus_{n \geq 0} A_n$ is connected and that $\mathbb{Q}$ is a subring of $\mathbf{k}$. Let B be the $\mathbf{k}$-subalgebra of A generated by $\mathfrak{p}$. In order to prove part (d), we need to show that $A \subset B$. Clearly, it suffices to show that $A_n \subset B$ for every $n \in \mathbb{N}$. We prove this by strong induction on n; thus, we fix some $n \in \mathbb{N}$, and assume as induction hypothesis that $A_m \subset B$ for every $m < n$. Our goal is then to show that $A_n \subset B$. This being trivial for $n = 0$ (since A is connected), we WLOG assume that $n > 0$. Let $a \in A_n$. Part (a) of this exercise yields $(S \star E)(a) \in \mathfrak{p} \subset B$. On the other hand, Exercise 1.3.20(h) (applied to $x = a$) yields

$$\Delta(a) \in 1 \otimes a + a \otimes 1 + \sum_{k=1}^{n-1} A_k \otimes A_{n-k}.$$

Hence, from the definition of convolution, we obtain

$$\begin{aligned}
(S \star E)(a) &\in \underbrace{S(1)}_{=1} E(a) + S(a) \underbrace{E(1)}_{=0} + \underbrace{(m \circ (S \otimes E)) \left(\sum_{k=1}^{n-1} A_k \otimes A_{n-k} \right)}_{=\sum_{k=1}^{n-1} S(A_k) E(A_{n-k})} \\
&= E(a) + \sum_{k=1}^{n-1} \underbrace{S(A_k)}_{\substack{\subset A_k \\ (\text{since } S \text{ is graded})}} \underbrace{E(A_{n-k})}_{\substack{\subset A_{n-k} \subset B \\ (\text{by the induction} \\ \text{hypothesis})}} \subset E(a) + \sum_{k=1}^{n-1} \underbrace{A_k}_{\substack{\subset B \\ (\text{by the induction} \\ \text{hypothesis})}} B \subset E(a) + B
\end{aligned}$$

(since B is a subalgebra). Hence, $E(a) \in (S \star E)(a) + B = B$ (since $(S \star E)(a) \in B$). Since $E(a) = na$, this becomes $na \in B$, thus $a \in B$ (since $\mathbb{Q}$ is a subring of $\mathbf{k}$). Since we have shown this for each $a \in A_n$, we thus obtain $A_n \subset B$, and our induction is complete.

This solution of part (d) is not the most generalizable one – for instance, (d) also holds if A is connected filtered instead of connected graded, and then a different argument is necessary. This is a part of the Cartier-Milnor-Moore theorem, and appears e.g. in [60, §3.2].

(e) If $a \in T(V)$ is homogeneous of positive degree and $p \in V$, then part (c) quickly yields $(S \star E)(ap) = [(S \star E)(a), p]$. This allows proving (e) by induction over n, with the induction base $n = 1$ being a consequence of part (b).

Hint to Exercise 1.6.1. (a) This can be done by diagram chasing. For example, if $\mathfrak{m}$ denotes the map $\Delta_C^* \circ \rho_{C,C} : C^* \otimes C^* \to C^*$, then the diagram

$$\begin{tikzcd}
 & & & C^* \otimes C^* \otimes C^* \arrow[dl, "\rho_{C,C} \otimes \mathrm{id}"'] \arrow[dr, "\mathrm{id} \otimes \rho_{C,C}"] \arrow[ddlll, bend right, "\mathfrak{m} \otimes \mathrm{id}"'] \arrow[ddrrr, bend left, "\mathrm{id} \otimes \mathfrak{m}"] & & & \\
 & & (C \otimes C)^* \otimes C^* \arrow[dll, "\Delta_C^* \otimes \mathrm{id}"'] \arrow[dr, "\rho_{C \otimes C, C}"] & & C^* \otimes (C \otimes C)^* \arrow[dl, "\rho_{C, C \otimes C}"'] \arrow[drr, "\mathrm{id} \otimes \Delta_C^*"] & & \\
C^* \otimes C^* \arrow[drr, "\rho_{C,C}"] \arrow[ddrrr, bend right, "\mathfrak{m}"'] & & & (C \otimes C \otimes C)^* \arrow[dl, "(\Delta_C \otimes \mathrm{id})^*"'] \arrow[dr, "(\mathrm{id} \otimes \Delta_C)^*"] & & & C^* \otimes C^* \arrow[dll, "\rho_{C,C}"'] \arrow[ddlll, bend left, "\mathfrak{m}"] \\
 & & (C \otimes C)^* \arrow[dr, "\Delta_C^*"] & & (C \otimes C)^* \arrow[dl, "\Delta_C^*"'] & & \\
 & & & C^* & & &
\end{tikzcd}$$

is commutative (since each of its little triangles and squares is); thus, $\mathfrak{m} \circ (\mathfrak{m} \otimes \mathrm{id}) = \mathfrak{m} \circ (\mathrm{id} \otimes \mathfrak{m})$ for $\mathfrak{m}$. This proves that the diagram (1.1.1) commutes for our algebra C^*. The commutativity of (1.1.2) is obtained similarly.

Alternatively, we could also solve part (a) trivially by first solving part (b) and then recalling Exercise 1.4.2.

(b) Straightforward verification on pure tensors.

(c) Let $C = \bigoplus_{n \geq 0} C_n$ be a graded $\mathbf{k}$-coalgebra. For every $n \in \mathbb{N}$, we identify $(C_n)^*$ with a $\mathbf{k}$-submodule of C^*, namely with the $\mathbf{k}$-submodule $\{f \in C^* \mid f(C_p) = 0 \text{ for all } p \in \mathbb{N} \text{ satisfying } p \neq n\}$. By the definition of C^o, we have $C^o = \bigoplus_{n \geq 0} (C_n)^*$. Hence, it remains to show that $(C_a)^* (C_b)^* \subset (C_{a+b})^*$ for all $a, b \in \mathbb{N}$, and that $1_{C^*} \in (C_0)^*$. But this is straightforward using the gradedness of Δ and ϵ.

(d) Diagram chasing or simple element-wise verification.

(e) Simple linear algebra (no Hopf algebras involved here).

(f) The "only if" direction is proved in the same way as part (d) (or as a corollary of part (d), since D° and C° are subalgebras of D^* and C^*). It remains to prove the "if" direction.

Assume that $f^* : D^o \to C^o$ is a $\mathbf{k}$-algebra morphism. We want to show that $f : C \to D$ is a $\mathbf{k}$-coalgebra morphism. In other words, we want to show that the two diagrams

(12.1.6)
$$\begin{tikzcd}
C \arrow[r, "f"] \arrow[d, "\Delta_C"'] & D \arrow[d, "\Delta_D"] \\
C \otimes C \arrow[r, "f \otimes f"] & D \otimes D
\end{tikzcd}
\qquad \text{and} \qquad
\begin{tikzcd}
C \arrow[rr, "f"] \arrow[dr, "\epsilon_C"'] & & D \arrow[dl, "\epsilon_D"] \\
 & \mathbf{k} &
\end{tikzcd}$$

commute. Let us start with the left one of these diagrams. The graded $\mathbf{k}$-module D is of finite type, and therefore the map $\rho_{D,D} : D^o \otimes D^o \to (D \otimes D)^o$ (a restriction of the map $\rho_{D,D} : D^* \otimes D^* \to (D \otimes D)^*$) is an isomorphism. Its inverse $\rho_{D,D}^{-1} : (D \otimes D)^o \to D^o \otimes D^o$ is therefore well-defined[377]. We can thus form the

[377] Beware: we don't have an inverse of the non-restricted map $\rho_{D,D} : D^* \otimes D^* \to (D \otimes D)^*$.

(asymmetric!) diagram

$$\begin{tikzcd}
D^o \arrow[rrr, "f^*"] & & & C^o & . \\
& D^o \otimes D^o \arrow[ul, "m_{D^*}"'] \arrow[r, "f^* \otimes f^*"] & C^o \otimes C^o \arrow[ur, "m_{C^*}"] \arrow[dr, "\rho_{C,C}"] & & \\
(D \otimes D)^o \arrow[uu, "\Delta_D^*"] \arrow[ur, "\rho_{D,D}^{-1}"] \arrow[rrr, "(f \otimes f)^*"'] & & & (C \otimes C)^o \arrow[uu, "\Delta_C^*"'] &
\end{tikzcd} \tag{12.1.7}$$

(The arrows labelled m_{C^*} and m_{D^*} could just as well have been labelled m_{C^o} and m_{D^o}, since the multiplication maps m_{C^o} and m_{D^o} are restrictions of m_{C^*} and m_{D^*}.) Argue that the diagram (12.1.7) commutes. Thus, $f^* \circ \Delta_D^* = \Delta_C^* \circ (f \otimes f)^*$ as maps from $(D \otimes D)^o$ to C^o. In other words, $(\Delta_D \circ f)^* = ((f \otimes f) \circ \Delta_C)^*$ as maps from $(D \otimes D)^o$ to C^o. But a general linear-algebraic fact states that if U and V are two graded **k**-modules such that V is of finite type, and if α and β are two graded **k**-linear maps $U \to V$ such that $\alpha^* = \beta^*$ as maps from V^o to U^o, then $\alpha = \beta$ [378]. Hence, $(\Delta_D \circ f)^* = ((f \otimes f) \circ \Delta_C)^*$ leads to $\Delta_D \circ f = (f \otimes f) \circ \Delta_C$. In other words, the first diagram in (12.1.6) commutes. The second is similar but easier. Thus, f is a **k**-coalgebra morphism, and the "if" direction is proved.

Hint to Exercise 1.6.4. Straightforward computations. For part (d), first show (independently of whether **k** is a field and its characteristic) that $\left(f^{(1)}\right)^m = m! f^{(m)}$ for every $m \in \mathbb{N}$.

Hint to Exercise 1.6.5. It is best to solve parts (c) and (d) before approaching (b).

(a) Both maps $\Delta_{\operatorname{Sym} V}$ and

$$\begin{array}{ccc} \mathbf{k}[\mathbf{x}] & \xrightarrow{\Delta} & \mathbf{k}[\mathbf{x}, \mathbf{y}], \\ f(x_1, \ldots, x_n) & \longmapsto & f(x_1 + y_1, \ldots, x_n + y_n) \end{array}$$

are **k**-algebra homomorphisms. Thus, in order to check that they are equal, it suffices to verify that they agree on V (since V generates $\operatorname{Sym} V$).

(c) This is a straightforward computation unless you get confused with the topologist's sign convention. The latter convention affects the twist map $T = T_{T(V),T(V)} : T(V) \otimes T(V) \to T(V) \otimes T(V)$ (in particular, we now have $T(x \otimes x) = -x \otimes x$ instead of $T(x \otimes x) = x \otimes x$), and thus also affects the multiplication in the **k**-algebra $T(V) \otimes T(V)$, because this multiplication is given by

$$m_{T(V) \otimes T(V)} = \left(m_{T(V)} \otimes m_{T(V)}\right) \circ (\operatorname{id} \otimes T \otimes \operatorname{id}).$$

Make sure you understand why this leads to $(1 \otimes x) \cdot (x \otimes 1) = -x \otimes x$ (whereas $(x \otimes 1) \cdot (1 \otimes x) = x \otimes x$).

(d) The trickiest part is showing that J is a graded **k**-submodule of $T(V)$. It suffices to check that J is generated (as a two-sided ideal) by homogeneous elements[379]; however, this is not completely trivial, as the designated generators x^2 for $x \in V$ need not be homogeneous. However, it helps to observe that J is also the two-sided ideal generated by the set

$$\{x \otimes x\}_{x \in V \text{ is homogeneous}} \cup \{x \otimes y + y \otimes x\}_{x,y \in V \text{ are homogeneous}}$$

(why?), which set does consist of homogeneous elements. Thus, J is a graded **k**-submodule of $T(V)$. From part (c), it is easy to observe that J is a two-sided coideal of $T(V)$ as well. Hence, $T(V)/J$ inherits a graded **k**-bialgebra structure from $T(V)$. The rest is easy.

(b) is now a consequence of what has been done in (d).

Hint to Exercise 1.6.6. Easy and straightforward.

Hint to Exercise 1.6.8. The hint after the exercise shows the way; here are a few more pointers. The solution proceeds in two steps:

- *Step 1:* Show that Proposition 1.6.7 holds when V is a finite free **k**-module.
- *Step 2:* Use this to conclude that Proposition 1.6.7 always holds.

[378] This follows immediately from Exercise 1.6.1 (e).

[379] Make sure you understand why.

The trick to Step 1 is to reduce the proof to Example 1.6.3. In a bit more detail: If V is a finite free $\mathbf{k}$-module with basis $(v_1, v_2, \ldots, v_n)$, then we know from Example 1.6.3 that the graded dual A^o of its tensor algebra $A := T(V)$ is a Hopf algebra whose basis $\left\{y_{(i_1,i_2,\ldots,i_\ell)}\right\}$ is indexed by words in the alphabet $I := \{1, 2, \ldots, n\}$. This allows us to define a $\mathbf{k}$-linear map $\phi : A^o \to T(V)$ by setting

$$\phi\left(y_{(i_1,i_2,\ldots,i_\ell)}\right) = v_{i_1} v_{i_2} \cdots v_{i_\ell} \qquad \text{for every } \ell \in \mathbb{N} \text{ and } (i_1, i_2, \ldots, i_\ell) \in I^\ell.$$

This $\mathbf{k}$-linear map ϕ then is an isomorphism from the Hopf algebra A^o to the putative Hopf algebra $\left(\operatorname{Sh}(V), \sqcup\!\sqcup, 1_{T(V)}, \Delta_{\sqcup\!\sqcup}, \epsilon, S\right)$, in the sense that it is invertible (since it sends a basis to a basis) and satisfies the five equalities

$$\begin{aligned}
\phi \circ m_{A^o} &= m_{\sqcup\!\sqcup} \circ (\phi \otimes \phi), \\
\phi \circ u_{A^o} &= u, \\
(\phi \otimes \phi) \circ \Delta_{A^o} &= \Delta_{\sqcup\!\sqcup} \circ \phi, \\
\epsilon_{A^o} &= \epsilon \circ \phi, \\
\phi \circ S_{A^o} &= S \circ \phi
\end{aligned}$$

(check all these – for instance, the first of these equalities follows by comparing (1.6.4) with the definition of $\sqcup\!\sqcup$). Thus, the latter putative Hopf algebra is an actual Hopf algebra (since the former is). This proves Proposition 1.6.7 for our finite free V, and thus completes Step 1.

Step 2 demonstrates the power of functoriality. We want to prove Proposition 1.6.7 in the general case, knowing that it holds when V is finite free. So let V be an arbitrary $\mathbf{k}$-module. For the sake of brevity, we shall write $\mathbf{V}$ for $T(V)$. Let $m_{\sqcup\!\sqcup}$ denote the $\mathbf{k}$-linear map $\mathbf{V} \otimes \mathbf{V} \to \mathbf{V}$ which sends every $a \otimes b$ to $a \sqcup\!\sqcup b$. One of the things that need to be shown is the commutativity of the diagram

(12.1.8)
$$\begin{tikzcd}
 & \mathbf{V} \otimes \mathbf{V} \arrow[dl, "\Delta_{\sqcup\!\sqcup} \otimes \Delta_{\sqcup\!\sqcup}"'] \arrow[dr, "m_{\sqcup\!\sqcup}"] & \\
\mathbf{V} \otimes \mathbf{V} \otimes \mathbf{V} \otimes \mathbf{V} \arrow[d, "\mathrm{id} \otimes T \otimes \mathrm{id}"'] & & \mathbf{V} \arrow[ddl, "\Delta_{\sqcup\!\sqcup}"] \\
\mathbf{V} \otimes \mathbf{V} \otimes \mathbf{V} \otimes \mathbf{V} \arrow[dr, "m_{\sqcup\!\sqcup} \otimes m_{\sqcup\!\sqcup}"'] & & \\
 & \mathbf{V} \otimes \mathbf{V} &
\end{tikzcd},$$

where T is the twist map $T_{\mathbf{V},\mathbf{V}}$. By linearity, it is clearly enough to verify this only on the pure tensors; that is, it is enough to check that every $a \in \mathbf{V}$ and $b \in \mathbf{V}$ satisfy

$$((m_{\sqcup\!\sqcup} \otimes m_{\sqcup\!\sqcup}) \circ (\mathrm{id} \otimes T \otimes \mathrm{id}) \circ (\Delta_{\sqcup\!\sqcup} \otimes \Delta_{\sqcup\!\sqcup}))(a \otimes b) = (\Delta_{\sqcup\!\sqcup} \circ m_{\sqcup\!\sqcup})(a \otimes b). \tag{12.1.9}$$

So let $a, b \in \mathbf{V}$ be arbitrary. WLOG assume that $a = v_1 v_2 \cdots v_p$ and $b = v_{p+1} v_{p+2} \cdots v_{p+q}$ for some $p, q \in \mathbb{N}$ and $v_1, v_2, \ldots, v_{p+q} \in V$. Define W to be the free $\mathbf{k}$-module with basis $(x_1, x_2, \ldots, x_{p+q})$, and let $\mathbf{W}$ be its tensor algebra $T(W)$. Then, W is a finite free $\mathbf{k}$-module, and so we know from Step 1 that Proposition 1.6.7 holds for W instead of V. But we can define a $\mathbf{k}$-linear map $f : W \to V$ that sends $x_1, x_2, \ldots, x_{p+q}$ to $v_1, v_2, \ldots, v_{p+q}$, respectively. This map $f : W \to V$ clearly induces a $\mathbf{k}$-algebra homomorphism $\mathbf{f} := T(f) : \mathbf{W} \to \mathbf{V}$ that respects all relevant shuffle-algebraic structure (i.e., it satisfies $\mathbf{f} \circ m_{\sqcup\!\sqcup} = m_{\sqcup\!\sqcup} \circ (\mathbf{f} \otimes \mathbf{f})$ and $(\mathbf{f} \otimes \mathbf{f}) \circ \Delta_{\sqcup\!\sqcup} = \Delta_{\sqcup\!\sqcup} \circ \mathbf{f}$ and so on), simply because this structure has been defined canonically in terms of each of V and W. Thus, in the diagram

$$\begin{tikzcd}
 & & \mathbf{W}\otimes\mathbf{W} \arrow[d, "\mathbf{f}\otimes\mathbf{f}"'] \arrow[ddll, "\Delta_{\sqcup\!\sqcup}\otimes\Delta_{\sqcup\!\sqcup}"'] \arrow[dddrr, "m_{\sqcup\!\sqcup}"] & & & , \\
 & & \mathbf{V}\otimes\mathbf{V} \arrow[dl, "\Delta_{\sqcup\!\sqcup}\otimes\Delta_{\sqcup\!\sqcup}"'] \arrow[ddr, "m_{\sqcup\!\sqcup}"'] & & & \\
\mathbf{W}\otimes\mathbf{W}\otimes\mathbf{W}\otimes\mathbf{W} \arrow[r, "\mathbf{f}\otimes\mathbf{f}\otimes\mathbf{f}\otimes\mathbf{f}"'] \arrow[dd, "\mathrm{id}\otimes T\otimes\mathrm{id}"'] & \mathbf{V}\otimes\mathbf{V}\otimes\mathbf{V}\otimes\mathbf{V} \arrow[dd, "\mathrm{id}\otimes T\otimes\mathrm{id}"'] & & & & \\
 & & & \mathbf{V} \arrow[ddl, "\Delta_{\sqcup\!\sqcup}"'] & \mathbf{W} \arrow[l, "\mathbf{f}"] \arrow[dddll, "\Delta_{\sqcup\!\sqcup}"] & \\
\mathbf{W}\otimes\mathbf{W}\otimes\mathbf{W}\otimes\mathbf{W} \arrow[r, "\mathbf{f}\otimes\mathbf{f}\otimes\mathbf{f}\otimes\mathbf{f}"] \arrow[ddrr, "m_{\sqcup\!\sqcup}\otimes m_{\sqcup\!\sqcup}"'] & \mathbf{V}\otimes\mathbf{V}\otimes\mathbf{V}\otimes\mathbf{V} \arrow[dr, "m_{\sqcup\!\sqcup}\otimes m_{\sqcup\!\sqcup}"'] & & & & \\
 & & \mathbf{V}\otimes\mathbf{V} & & & \\
 & & \mathbf{W}\otimes\mathbf{W} \arrow[u, "\mathbf{f}\otimes\mathbf{f}"] & & &
\end{tikzcd}$$

all the little quadrilaterals commute. The outer pentagon also commutes, since Proposition 1.6.7 holds for W instead of V. If $\mathbf{f}$ was surjective, then we would be able to conclude that the inner pentagon also commutes, so we would immediately get the commutativity of (12.1.8). But even if $\mathbf{f}$ is not surjective, we are almost there: The inner pentagon commutes on the image of the map $\mathbf{f}\otimes\mathbf{f}:\mathbf{W}\otimes\mathbf{W}\to\mathbf{V}\otimes\mathbf{V}$ (because when we start at $\mathbf{W}\otimes\mathbf{W}$, we can walk around the outer pentagon instead, which is known to commute), but this image contains $a\otimes b$ (since $a=v_1v_2\cdots v_p=\mathbf{f}(x_1x_2\cdots x_p)$ and similarly $b=\mathbf{f}(x_{p+1}x_{p+2}\cdots x_{p+q})$), so we conclude that (12.1.9) holds, as we wanted to show.

This is only one of the diagrams we need to prove in order to prove Proposition 1.6.7, but the other diagrams are done in the exact same way.

Hint to Exercise 1.7.9. Straightforward reasoning using facts like "a union of finitely many finite sets is finite" and "a tensor is a sum of finitely many pure tensors".

Hint to Exercise 1.7.13. Parts (a), (b), (d) and (e) of Proposition 1.7.11 are easy. (In proving (1.7.3) and later, it helps to first establish an extension of (1.7.2) to infinite sums[380].) For part (c), recall that the binomial formula $(a+b)^n=\sum_{k=0}^{n}\binom{n}{k}a^kb^{n-k}$ holds for any two commuting elements a and b of any ring (such as f and g in the convolution algebra $\operatorname{Hom}(C,A)$). Part (f) follows from (e) using (1.7.3). Part (g) is best proved in two steps: First, use induction to prove part (g) in the case when $u=T^k$ for some $k\in\mathbb{N}$ (this relies on (1.7.3)); then, notice that both sides of (1.7.7) depend $\mathbf{k}$-linearly on u, whence the general case follows (up to some mudfighting with infinite sums). Part (h) is an instance of the "local $\star$-nilpotence" already observed in the proof of Proposition 1.4.7. Part (j) follows from (h). Part (i) follows from Proposition 1.4.3 (applied to $C'=C$, $A'=B$, $\gamma=\mathrm{id}_C$ and $\alpha=s$) in a similar way as part (g) followed from (1.7.3).

Hint to Exercise 1.7.20. Proposition 1.7.15 is a classical result, often proved by a lazy reference to the mythical complex analysis class the reader has surely seen it in. Here is a do-it-yourself purely algebraic proof:

[380] Namely: Let $(r_q)_{q\in Q}\in(\mathbf{k}[[T]])^Q$ be a family of power series such that the (possibly infinite) sum $\sum_{q\in Q}r_q$ converges in $\mathbf{k}[[T]]$. Let $f\in\mathfrak{n}(C,A)$. Then, the family $((r_q)^{\star}(f))_{q\in Q}\in(\operatorname{Hom}(C,A))^Q$ is pointwise finitely supported and satisfies $\left(\sum_{q\in Q}r_q\right)^{\star}(f)=\sum_{q\in Q}(r_q)^{\star}(f)$.

- *Step 1:* If $u, v \in \mathbf{k}[[T]]$ are two power series having the same constant term and satisfying $\frac{d}{dT}u = \frac{d}{dT}v$, then $u = v$. This simple lemma (whose analogue for differentiable functions is a fundamental fact of real analysis) is easily proved by comparing coefficients in $\frac{d}{dT}u = \frac{d}{dT}v$ and recalling that $\mathbf{k}$ is a $\mathbb{Q}$-algebra (so $1, 2, 3, \ldots$ are invertible in $\mathbf{k}$).
- *Step 2:* If $u, v \in \mathbf{k}[[T]]$ are two power series having constant term 1 and satisfying $\left(\frac{d}{dT}u\right) \cdot v = \left(\frac{d}{dT}v\right) \cdot u$, then $u = v$. This can be proved by applying Step 1 to uv^{-1} and 1 instead of u and v.
- *Step 3:* The power series $\overline{\log}\left[\overline{\exp}\right]$ and $\overline{\exp}\left[\overline{\log}\right]$ are well-defined and have constant term 0. (Easy.)
- *Step 4:* If $w \in \mathbf{k}[[T]]$ is a power series having constant term 0, then

$$\frac{d}{dT}\left(\overline{\exp}[w]\right) = \left(\frac{d}{dT}w\right) \cdot \exp[w] \qquad \text{and}$$
$$\frac{d}{dT}\left(\overline{\log}[w]\right) = \left(\frac{d}{dT}w\right) \cdot \frac{1}{1+w}.$$

These formulas can be derived from the chain rule, or more directly from $\overline{\exp}[w] = \sum_{n \geq 1} \frac{1}{n!}w^n$ and $\overline{\log}[w] = \sum_{n \geq 1} \frac{(-1)^{n-1}}{n}w^n$.
- *Step 5:* Show $\overline{\exp}\left[\overline{\log}\right] = T$ by applying Step 2 to $u = \exp\left[\overline{\log}\right]$ and $v = 1 + T$.
- *Step 6:* Show $\overline{\log}\left[\overline{\exp}\right] = T$ by applying Step 1 to $u = \overline{\log}\left[\overline{\exp}\right]$ and $v = T$.

Lemma 1.7.16 easily follows from Proposition 1.7.11(f).

Remains to prove Proposition 1.7.18. It is easy to see that $\log^\star(\exp^\star f) = \overline{\log}^\star(\overline{\exp}^\star f)$ for each $f \in \mathfrak{n}(C, A)$; thus, Proposition 1.7.18(a) follows from (1.7.7) using Proposition 1.7.15 and Proposition 1.7.11(f) (since $T^\star(f) = f$). A similar argument yields Proposition 1.7.18(b) (this time, we need to observe that $\exp^\star(\log^\star g) = \overline{\exp}^\star\left(\overline{\log}^\star(g - u_A\epsilon_C)\right) + u_A\epsilon_C$ first). To prove Proposition 1.7.18(c), first use Proposition 1.7.11(c) to show that $\exp^\star(f + g)$ is well-defined; then, apply the well-known fact that $\exp(x + y) = \exp x \cdot \exp y$ for any two commuting elements x and y of a ring (provided the exponentials are well-defined; some yak-shaving is required here to convince oneself that the infinite sums behave well)[381]. Part (d) is trivial. Part (e) is an induction on n. Part (f) is a rehash of the definition of $\log^\star(f + u_A\epsilon_C) = \overline{\log}^\star f$.

Hint to Exercise 1.7.28. Proposition 1.7.21(a) is easily proved by unpacking the definition of convolution (just like Proposition 1.4.3). Part (b) follows from (a) by induction.

The trick to Proposition 1.7.22 is to realize that if $f \in \operatorname{Hom}(C, A)$ is as in Proposition 1.7.22, then every $x, y \in C$ satisfy

$$f(xy) = \epsilon(y) f(x) + \epsilon(x) f(y), \tag{12.1.10}$$

because $xy - \epsilon(x) y - \epsilon(y) x = \epsilon(x)\epsilon(y) \cdot 1 + \underbrace{(x - \epsilon(x))}_{\in \ker \epsilon} \underbrace{(y - \epsilon(y))}_{\in \ker \epsilon}$ is annihilated by f. Once this equality is known, it is not hard to prove Proposition 1.7.22 "by hand" by induction on n (using Sweedler notation). Alternatively, for a cleaner proof, the equality (12.1.10) can be restated in an element-free way as

$$f \circ m_C = m_A \circ (f \otimes \mathfrak{i} + \mathfrak{i} \otimes f),$$

[381] If you have not seen this well-known fact, prove it by a quick computation using the binomial formula.

where $\mathfrak{i} = u_A \circ \epsilon_C$ is the unity of the $\mathbf{k}$-algebra $(\operatorname{Hom}(C, A), \star)$; then, an application of Proposition 1.7.21(b) shows that every $n \in \mathbb{N}$ satisfies

$$f^{\star n} \circ m_C = m_A \circ \underbrace{(f \otimes \mathfrak{i} + \mathfrak{i} \otimes f)^{\star n}}_{\substack{=\sum_{i=0}^{n} \binom{n}{i} (f\otimes\mathfrak{i})^{\star i} \star (\mathfrak{i}\otimes f)^{\star(n-i)} \\ \text{(by the binomial formula,} \\ \text{since } f\otimes\mathfrak{i} \text{ and } \mathfrak{i}\otimes f \text{ commute in} \\ \text{the convolution algebra } \operatorname{Hom}(C\otimes C, A\otimes A))}} = m_A \circ \left(\sum_{i=0}^{n} \binom{n}{i} \underbrace{(f \otimes \mathfrak{i})^{\star i} \star (\mathfrak{i} \otimes f)^{\star(n-i)}}_{\substack{=f^{\star i}\otimes f^{\star(n-i)} \\ \text{(by repeated application of Exercise 1.4.4(a))}}} \right)$$

$$= m_A \circ \left(\sum_{i=0}^{n} \binom{n}{i} f^{\star i} \otimes f^{\star(n-i)} \right),$$

which is precisely Proposition 1.7.22 (restated in an element-free way).

Proposition 1.7.23 is an easy consequence of Proposition 1.7.22, since $(\exp^\star f)(xy) = \sum_{n\in\mathbb{N}} \dfrac{1}{n!} f^{\star n}(xy)$. (Again, fighting infinite sums is probably the most laborious part of the proof.)

Lemma 1.7.24 can be reduced to the fact that the matrix $\left(i^{N+1-j}\right)_{i,j=1,2,\ldots,N+1} \in \mathbb{Q}^{(N+1)\times(N+1)}$ is invertible (since its determinant is the Vandermonde determinant $\prod_{1\leq i<j\leq N+1} \underbrace{(i-j)}_{\neq 0} \neq 0$) and thus has trivial kernel (not just over $\mathbb{Q}$, but on any torsionfree abelian group).

Lemma 1.7.25 follows from Lemma 1.7.24, because a finitely supported family indexed by nonnegative integers must become all zeroes from some point on.

The proof of Proposition 1.7.26 is rather surprising: It suffices to show that $f(xy) = 0$ for all $x, y \in \ker \epsilon$. So let us fix $x, y \in \ker \epsilon$. Proposition 1.7.11(h) yields $f \in \mathfrak{n}(C, A)$. Let $t \in \mathbb{N}$ be arbitrary. Then, Proposition 1.7.18(e) (applied to $n = t$) shows that $tf \in \mathfrak{n}(C, A)$ and $\exp^\star(tf) = (\exp^\star f)^{\star t}$. But Exercise 1.5.11(b) shows that $(\exp^\star f)^{\star t}$ is a $\mathbf{k}$-algebra homomorphism $C \to A$. Hence, $(\exp^\star f)^{\star t}(xy) = (\exp^\star f)^{\star t}(x) \cdot (\exp^\star f)^{\star t}(y)$. Rewriting $(\exp^\star f)^{\star t}$ as $\exp^\star(tf) = \sum_{n\in\mathbb{N}} \dfrac{1}{n!} f^{\star n} t^n$ on both sides, and multiplying out the right hand side, we can rewrite this as

$$\sum_{k\in\mathbb{N}} \frac{1}{k!} f^{\star k}(xy)\, t^k = \sum_{k\in\mathbb{N}} \left(\sum_{i=0}^{k} \frac{f^{\star i}(x)}{i!} \cdot \frac{f^{\star(k-i)}(y)}{(k-i)!} \right) t^k.$$

In other words,

$$\sum_{k\in\mathbb{N}} w_k t^k = 0, \qquad \text{where we set } w_k = \frac{1}{k!} f^{\star k}(xy) - \sum_{i=0}^{k} \frac{f^{\star i}(x)}{i!} \cdot \frac{f^{\star(k-i)}(y)}{(k-i)!}.$$

But we have proved this for all $t \in \mathbb{N}$. Thus, Lemma 1.7.25 shows that

$$w_k = 0 \qquad \text{for every } k \in \mathbb{N}.$$

Applying this to $k = 1$ and simplifying, we obtain $f(xy) - \epsilon(x) f(y) - f(x) \epsilon(y) = 0$. Since $x, y \in \ker \epsilon$, this simplifies even further to $f(xy) = 0$, which proves Proposition 1.7.26.

Finally, we need to prove Proposition 1.7.27. Set $F = \exp^\star f$ and $\widetilde{F} = F - u_A \epsilon_C$, so that $\widetilde{F} \in \mathfrak{n}(C, A)$. Then, Proposition 1.7.23 shows that $F : C \to A$ is a $\mathbf{k}$-algebra homomorphism, so it remains to show that F is surjective. But it is easy to see using Proposition 1.7.18(a) that $f = \overline{\log}^\star \widetilde{F}$.

Define $\widetilde{\operatorname{id}} \in \mathfrak{n}(C, C)$ by $\widetilde{\operatorname{id}} = \operatorname{id}_C - u_C \epsilon_C$. Then, it is not hard to see that $F \circ \widetilde{\operatorname{id}} = \widetilde{F}$. Hence, $f = \overline{\log}^\star \underbrace{\widetilde{F}}_{=F\circ\widetilde{\operatorname{id}}} = \overline{\log}^\star \left(F \circ \widetilde{\operatorname{id}} \right) = F \circ \left(\overline{\log}^\star \left(\widetilde{\operatorname{id}} \right) \right)$ (by Proposition 1.7.11(i), since F is a $\mathbf{k}$-algebra homomorphism). Therefore, $f(C) \subset F(C)$. Since F is a $\mathbf{k}$-algebra homomorphism, this entails that $F(C)$ is a $\mathbf{k}$-subalgebra of A that contains $f(C)$ as a subset. But this causes $F(C)$ to be the whole A (since $f(C)$ generates A). Thus, F is surjective, so Proposition 1.7.27 is proven.

Hint to Exercise 1.7.33. We must prove Theorem 1.7.29. Part (a) is easy. For the remainder of the proof, we set $\widetilde{\operatorname{id}} = \operatorname{id}_A - u_A \epsilon_A \in \operatorname{End} A$, and equip ourselves with some simple lemmas:

- The kernel $\ker \epsilon$ is an ideal of A.
- We have $\widetilde{\operatorname{id}} \in \mathfrak{n}(A, A)$ and $\ker \widetilde{\operatorname{id}} = \mathbf{k} \cdot 1_A$ and $\widetilde{\operatorname{id}}(A) = \ker \epsilon$.
- We have $A / \left(\mathbf{k} \cdot 1_A + (\ker \epsilon)^2\right) \cong (\ker \epsilon) / (\ker \epsilon)^2$ as $\mathbf{k}$-modules.

Now, to the proof of Theorem 1.7.29(b). Using $\mathfrak{e} = \log^\star(\operatorname{id}_A) = \overline{\log}^\star \widetilde{\operatorname{id}}$ and $\widetilde{\operatorname{id}}(1_A) = 0$, it is easy to see that $\mathfrak{e}(1_A) = 0$. Hence, $\mathfrak{e}(A_0) = 0$ since A is connected. Thus, Proposition 1.7.26 shows that $\mathfrak{e}\left((\ker \epsilon)^2\right) = 0$ (since $\exp^\star \mathfrak{e} = \operatorname{id}_A$ is a $\mathbf{k}$-algebra homomorphism). Combined with $\mathfrak{e}(1_A) = 0$, this yields $\mathbf{k} \cdot 1_A + (\ker \epsilon)^2 \subset \ker \mathfrak{e}$. But this inclusion is actually an equality, as we can show by the following computation: We have $\mathfrak{e} = \overline{\log}^\star \widetilde{\operatorname{id}} = \sum_{n \geq 1} \dfrac{(-1)^{n-1}}{n} \widetilde{\operatorname{id}}^{\star n}$, and therefore each $x \in A$ satisfies

$$\mathfrak{e}(x) = \sum_{n \geq 1} \frac{(-1)^{n-1}}{n} \widetilde{\operatorname{id}}^{\star n}(x) = \underbrace{\widetilde{\operatorname{id}}(x)}_{\substack{= x - \epsilon(x) 1_A \\ \text{(by the definition of } \widetilde{\operatorname{id}})}} + \sum_{n \geq 2} \frac{(-1)^{n-1}}{n} \underbrace{\widetilde{\operatorname{id}}^{\star n}(x)}_{\substack{\in \left(\widetilde{\operatorname{id}}(A)\right)^n \\ \text{(by induction on } n, \\ \text{using the definition} \\ \text{of convolution)}}}$$

$$\in x - \epsilon(x) 1_A + \sum_{n \geq 2} \frac{(-1)^{n-1}}{n} \Bigg(\underbrace{\widetilde{\operatorname{id}}(A)}_{= \ker \epsilon}\Bigg)^n = x - \underbrace{\epsilon(x)}_{\in \mathbf{k}} 1_A + \underbrace{\sum_{n \geq 2} \frac{(-1)^{n-1}}{n} (\ker \epsilon)^n}_{\subset (\ker \epsilon)^2} \subset x - \mathbf{k} \cdot 1_A + (\ker \epsilon)^2,$$

so that

$$x - \mathfrak{e}(x) \in \mathbf{k} \cdot 1_A + (\ker \epsilon)^2. \tag{12.1.11}$$

If $x \in \ker \mathfrak{e}$, then this simplifies to $x \in \mathbf{k} \cdot 1_A + (\ker \epsilon)^2$. Thus, $\ker \mathfrak{e} \subset \mathbf{k} \cdot 1_A + (\ker \epsilon)^2$. Combining this with $\mathbf{k} \cdot 1_A + (\ker \epsilon)^2 \subset \ker \mathfrak{e}$, we obtain $\ker \mathfrak{e} = \mathbf{k} \cdot 1_A + (\ker \epsilon)^2$. But the homomorphism theorem yields

$$\mathfrak{e}(A) \cong A / \underbrace{\ker \mathfrak{e}}_{= \mathbf{k} \cdot 1_A + (\ker \epsilon)^2} = A / \left(\mathbf{k} \cdot 1_A + (\ker \epsilon)^2\right) \cong (\ker \epsilon) / (\ker \epsilon)^2 \qquad \text{(as seen above)}$$

as $\mathbf{k}$-modules. This completes the proof of Theorem 1.7.29(b).

Theorem 1.7.29(c) just requires showing that $\mathfrak{q}(A_0) = 0$, which is a consequence of $\mathfrak{e}(A_0) = 0$.

Next, we shall prove Theorem 1.7.29(d). We have $\mathfrak{q} \in \mathfrak{n}(A, \operatorname{Sym}(\mathfrak{e}(A)))$. Furthermore, $\mathfrak{q}(A)$ generates the $\mathbf{k}$-algebra $\operatorname{Sym}(\mathfrak{e}(A))$ (since $\mathfrak{q}(A) = \operatorname{Sym}^1(\mathfrak{e}(A))$). From Theorem 1.7.29(b), we get $\ker \mathfrak{e} = \mathbf{k} \cdot 1_A + (\ker \epsilon)^2$, from which we easily obtain $\mathfrak{q}(1_A) = 0$ and $\mathfrak{q}\left((\ker \epsilon)^2\right) = 0$. Thus, Proposition 1.7.27 (applied to A, $\operatorname{Sym}(\mathfrak{e}(A))$ and $\mathfrak{q}$ instead of C, A and f) shows that $\exp^\star \mathfrak{q} : A \to \operatorname{Sym}(\mathfrak{e}(A))$ is a surjective $\mathbf{k}$-algebra homomorphism. But $\mathfrak{s}$ is a $\mathbf{k}$-algebra homomorphism $\operatorname{Sym}(\mathfrak{e}(A)) \to A$ and satisfies $\mathbf{i} = \mathfrak{s} \circ \iota_{\mathfrak{e}(A)}$ (by its definition). Thus, Proposition 1.7.11(i) (applied to A, $\operatorname{Sym}(\mathfrak{e}(A))$, A, $\mathfrak{s}$, $\exp$ and $\mathfrak{q}$ instead of C, A, B, s, u and f) shows that $\mathfrak{s} \circ \mathfrak{q} \in \mathfrak{n}(A, A)$ and $\exp^\star(\mathfrak{s} \circ \mathfrak{q}) = \mathfrak{s} \circ (\exp^\star \mathfrak{q})$. However, it is easy to see that $\mathfrak{s} \circ \mathfrak{q} = \mathfrak{e}$ (since $\mathbf{i} = \mathfrak{s} \circ \iota_{\mathfrak{e}(A)}$); this lets us rewrite the equality $\exp^\star(\mathfrak{s} \circ \mathfrak{q}) = \mathfrak{s} \circ (\exp^\star \mathfrak{q})$ as $\exp^\star \mathfrak{e} = \mathfrak{s} \circ (\exp^\star \mathfrak{q})$. Comparing this with $\exp^\star \mathfrak{e} = \operatorname{id}_A$, we obtain $\mathfrak{s} \circ (\exp^\star \mathfrak{q}) = \operatorname{id}_A$. Since $\exp^\star \mathfrak{q}$ is surjective, this entails that the maps $\exp^\star \mathfrak{q}$ and $\mathfrak{s}$ are mutually inverse. This proves Theorem 1.7.29(d).

Theorem 1.7.29(d) shows that $A \cong \operatorname{Sym}(\mathfrak{e}(A))$ as $\mathbf{k}$-algebras, but Theorem 1.7.29(b) shows that $\mathfrak{e}(A) \cong (\ker \epsilon) / (\ker \epsilon)^2$ as $\mathbf{k}$-modules. Combining these, we obtain Theorem 1.7.29(e).

Finally, to prove Theorem 1.7.29(f), we notice that each $x \in A$ satisfies

$$\begin{aligned} x - \mathfrak{e}(x) &\in \mathbf{k} \cdot 1_A + (\ker \epsilon)^2 && \text{(by (12.1.11))} \\ &= \ker \mathfrak{e} && \text{(by Theorem 1.7.29(b))} \end{aligned}$$

and thus $0 = \mathfrak{e}(x - \mathfrak{e}(x)) = \mathfrak{e}(x) - (\mathfrak{e} \circ \mathfrak{e})(x)$.

Acknowledgements

The authors thank the following for helpful comments and/or teaching them about Hopf algebras: Marcelo Aguiar, Federico Ardila, Lou Billera, Richard Ehrenborg, Mark Haiman, Florent Hivert, Christophe Hohlweg, Jia Huang, Jang Soo Kim, Aaron Lauve, Dominique Manchon, John Palmieri, Alexander Postnikov, Margie Readdy, Nathan Reading, Christophe Reutenauer, Hans-Jürgen Schneider, Richard Stanley, Josh Swanson, Muge Taskin, Jean-Yves Thibon.

Parts of this text have been written during stays at the Mathematisches Forschungsinstitut Oberwolfach (2019 and 2020)[382] and at the Institut Mittag–Leffler Djursholm (Spring 2020, supported by the Swedish Research Council under grant no. 2016-06596); DG thanks both for their hospitality.

References

[1] Eiichi Abe. Hopf algebras. *Cambridge Tracts in Mathematics* **74**. Cambridge University Press, Cambridge-New York, 1980.

[2] Marcelo Aguiar, et al. (28 authors). Supercharacters, symmetric functions in noncommuting variables, and related Hopf algebras. *Adv. Math.* **229** (2012), 2310–2337. https://doi.org/10.1016/j.aim.2011.12.024 . Also available as arXiv:1009.4134v2.

[3] Marcelo Aguiar and Federico Ardila. Hopf monoids and generalized permutahedra. arXiv:1709.07504v1.

[4] Marcelo Aguiar, Nantel Bergeron, and Frank Sottile. Combinatorial Hopf algebras and generalized Dehn-Sommerville relations. *Compos. Math.* **142** (2006), pp. 1–30. A newer version of this paper appears at http://pi.math.cornell.edu/~maguiar/CHalgebra.pdf.

[5] Marcelo Aguiar, Aaron Lauve. The characteristic polynomial of the Adams operators on graded connected Hopf algebras. *Algebra & Number Theory* **9-3** (2015), 547–583. Also available at http://pi.math.cornell.edu/~maguiar/adams.pdf and as arXiv:1403.7584v2.

[6] Marcelo Aguiar and Swapneel Mahajan. Monoidal functors, species and Hopf algebras. *CRM Monograph Series* **29**. American Mathematical Society, Providence, RI, 2010. Available at http://pi.math.cornell.edu/~maguiar/a.pdf

[7] Marcelo Aguiar and Frank Sottile. Structure of the Malvenuto-Reutenauer Hopf algebra of permutations. *Adv. Math.* **191** (2005), 225–275. https://doi.org/10.1016/j.aim.2004.03.007.
A preprint is available at http://pi.math.cornell.edu/~maguiar/MR.pdf

[8] Nicolas Andruskiewitsch, Walter Ferrer Santos. The beginnings of the theory of Hopf algebras. *Acta Appl Math* **108** (2009), 3–17. See also a corrected postprint published on arXiv as arXiv:0901.2460v3.

[9] Sami H. Assaf and Peter R.W. McNamara. A Pieri rule for skew shapes. *J. Combin. Theory, Ser. A* **118** (2011), 277–290. https://doi.org/10.1016/j.jcta.2010.03.010

[10] Olga Azenhas. Littlewood-Richardson fillings and their symmetries. *Matrices and group representations* (Coimbra, 1998), 81–92, Textos Mat. Ser. B, 19, Univ. Coimbra, Coimbra, 1999. http://www.mat.uc.pt/~oazenhas/graciano+.pdf.

[11] Olga Azenhas, Ronald C. King, Itaru Terada. The involutive nature of the Littlewood-Richardson commutativity bijection. arXiv:1603.05037v1.

[12] Andrew Baker, and Birgit Richter. Quasisymmetric functions from a topological point of view. *Math. Scand.* **103** (2008), 208–242. http://dx.doi.org/10.7146/math.scand.a-15078

[13] Farzin Barekat, Victor Reiner, Stephanie van Willigenburg. Corrigendum to "Coincidences among skew Schur functions" [*Adv. Math.* **216** (2007), 118–152]. *Adv. Math.* **220** (2009), 1655–1656. See also a corrected version of this paper on arXiv:math/0602634v4.

[14] Carolina Benedetti, Joshua Hallam, John Machacek. Combinatorial Hopf Algebras of Simplicial Complexes. arXiv:1505.04458v2. (Published in: *SIAM J. Discrete Math.* **30** (3), 1737–1757.)

[15] Carolina Benedetti, Bruce Sagan. Antipodes and involutions. arXiv:1410.5023v4. (Published in: *Journal of Combinatorial Theory, Series A* **148** (2017), 275–315.)

[16] Georgia Benkart, Frank Sottile, Jeffrey Stroomer. Tableau Switching: Algorithms and Applications. *Journal of Combinatorial Theory, Series A* **76**, 1, October 1996, 11–43. https://doi.org/10.1006/jcta.1996.0086
Preprint available at http://www.math.tamu.edu/~sottile/research/pdf/switching.pdf

[17] Chris Berg, Nantel Bergeron, Franco Saliola, Luis Serrano, Mike Zabrocki. A lift of the Schur and Hall-Littlewood bases to non-commutative symmetric functions. *Canad. J. Math.* **66** (2014), 525–565. http://dx.doi.org/10.4153/CJM-2013-013-0.
A preprint is arXiv:1208.5191v3.

[18] Nantel Bergeron, Mike Zabrocki. The Hopf algebras of symmetric functions and quasi-symmetric functions in non-commutative variables are free and co-free. *Journal of Algebra and Its Applications* **08**, Issue 04, August 2009, 581–600. A preprint also appears at arXiv:math/0509265v3.

[19] Louis J. Billera. Flag enumeration in polytopes, Eulerian partially ordered sets and Coxeter groups. *Proceedings of the International Congress of Mathematicians* **IV**, 2389–2415, Hindustan Book Agency, New Delhi, 2010. http://pi.math.cornell.edu/~billera/papers/eulericm.pdf

[20] Louis J. Billera, Francesco Brenti. Quasisymmetric functions and Kazhdan-Lusztig polynomials. arXiv:0710.3965v2. Published in: Israel Journal of Mathematics, August 2011, 184, pp. 317–348. https://doi.org/10.1007/s11856-011-0070-0

[382] This research was supported through the programme "Oberwolfach Leibniz Fellows" by the Mathematisches Forschungsinstitut Oberwolfach in 2019 and 2020.

[21] Louis J. Billera, Ning Jia, and Victor Reiner. A quasisymmetric function for matroids. *European J. Combin.* **30** (2009), pp. 1727–1757. https://doi.org/10.1016/j.ejc.2008.12.007 . A preprint also appears at arXiv:math/0606646v3.

[22] Anders Björner. Some combinatorial and algebraic properties of Coxeter complexes and Tits buildings. *Adv. in Math.* **52** (1984), 173–212. https://doi.org/10.1016/0001-8708(84)90021-5

[23] Jonah Blasiak. Kronecker coefficients for one hook shape. *Seminaire Lotharingien de Combinatoire* **77** (2017), B77c. https://www.emis.de/journals/SLC/wpapers/s77blasiak.html

[24] D. Blessenohl, H. Laue. Algebraic combinatorics related to the free Lie algebra. *Seminaire Lotharingien de Combinatoire* **29** (1992), B29e. https://www.emis.de/journals/SLC/opapers/s29laue.html

[25] Dieter Blessenohl, Manfred Schocker. Noncommutative character theory of the symmetric group. Imperial College Press 2005. https://www.worldscientific.com/worldscibooks/10.1142/p369

[26] Ben Blum-Smith, Samuel Coskey. The Fundamental Theorem on Symmetric Polynomials: History's First Whiff of Galois Theory. arXiv:1301.7116v4. An updated version was published in: The College Mathematics Journal Vol. 48, No. 1 (January 2017), pp. 18–29. https://doi.org/10.4169/college.math.j.48.1.18

[27] N. Bourbaki. Éléments de Mathématique: Groupes et algèbres de Lie, Chapitres 2 et 3. Springer, Heidelberg 2006.

[28] Thomas Britz, Sergey Fomin. Finite posets and Ferrers shapes. *Adv. in Math.* **158**, Issue 1, 1 March 2001, 86–127. Better version to be found on arXiv as arXiv:math/9912126v1.

[29] N.G. de Bruijn, D.A. Klarner. Multisets of aperiodic cycles. *SIAM J. Alg. Disc. Math.* **3** (1982), no. 3, 359–368. https://pure.tue.nl/ws/files/1674487/597568.pdf

[30] Daniel Bump. Notes on representations of $GL(r)$ over a finite field. Available at http://math.stanford.edu/~bump/.

[31] Emily Burgunder. Eulerian idempotent and Kashiwara-Vergne conjecture. *Annales de l'institut Fourier* **58** (2008), Issue 4, 1153–1184. https://eudml.org/doc/10345.

[32] Lynne M. Butler, Alfred W. Hales. Nonnegative Hall polynomials. *Journal of Algebraic Combinatorics* **2** (1993), Issue 2, 125–135. https://www.emis.de/journals/JACO/Volume2_2/l42886q158156k2u.html

[33] Stefaan Caenepeel, J. Vercruysse. Hopf algebras. *Lecture notes, Vrije Universiteit Brussel* **2013**. http://homepages.ulb.ac.be/~scaenepe/Hopfalgebra.pdf

[34] Peter J. Cameron. Notes on matroids and codes. *Lecture notes*, 2000. http://www.maths.qmul.ac.uk/~pjc/comb/matroid.pdf

[35] Pierre F. Cartier. A primer of Hopf algebras. Frontiers in number theory, physics, and geometry. II, 537–615, Springer, Berlin, 2007.
A preprint is available at http://preprints.ihes.fr/2006/M/M-06-40.pdf

[36] Vyjayanthi Chari, and Andrew N. Pressley. A guide to quantum groups. Cambridge University Press, Cambridge, 1994.

[37] Sunil K. Chebolu, Jan Minac. Counting irreducible polynomials over finite fields using the inclusion-exclusion principle. *Math. Mag.* **84** (2011), 369–371. A preprint is arXiv:1001.0409v6.

[38] K.T. Chen, R.H. Fox, R.C. Lyndon. Free Differential Calculus, IV: The Quotient Groups of the Lower Central Series. *Annals of Mathematics* **68** (1), 81–95. https://doi.org/10.2307/1970044

[39] Sergei Chmutov, Sergei V. Duzhin, Jacob Mostovoy. Introduction to Vassiliev Knot Invariants. CUP 2012.
Various preprint versions can be found at https://people.math.osu.edu/chmutov.1/preprints/ , at http://www.pdmi.ras.ru/~duzhin/papers/cdbook/ and at arXiv:1103.5628v3.

[40] Keith Conrad. Expository papers ("Blurbs"), specifically *Tensor Products I*, *Tensor Products II*, *Exterior Powers*. http://www.math.uconn.edu/~kconrad/blurbs

[41] Henry Crapo and William Schmitt. Primitive elements in the matroid-minor Hopf algebra. *J. Algebraic Combin.* **28** (2008), 43–64. https://doi.org/10.1007/s10801-007-0066-3.
A preprint is arXiv:math/0511033v1.

[42] ———. A unique factorization theorem for matroids. *J. Combin. Theory Ser. A* **112** (2005), 222–249. https://doi.org/10.1016/j.jcta.2005.02.004

[43] ———. A free subalgebra of the algebra of matroids. *European J. Combin.* **26** (2005), 1066–1085. https://doi.org/10.1016/j.ejc.2004.05.006

[44] William Crawley-Boevey. *Lectures on representation theory and invariant theory*, Bielefeld 1989/90. Available from https://www.math.uni-bielefeld.de/~wcrawley/.

[45] Maxime Crochemore, Jacques Désarménien, Dominique Perrin. A note on the Burrows–Wheeler transformation. *Theoretical Computer Science* **332** (2005), pp. 567–572. https://doi.org/10.1016/j.tcs.2004.11.014
A preprint is arXiv:cs/0502073.

[46] Geir Dahl. Network flows and combinatorial matrix theory. Lecture notes, 4 September 2013. http://www.uio.no/studier/emner/matnat/math/MAT-INF4110/h13/lecturenotes/combmatrix.pdf

[47] Sorin Dascalescu, Constantin Nastasescu, Serban Raianu. Hopf algebras. An introduction. *Monographs and Textbooks in Pure and Applied Mathematics* **235**. Marcel Dekker, Inc., New York, 2001.

[48] Barry Dayton. Witt vectors, the Grothendieck Burnside ring, and Necklaces. http://orion.neiu.edu/~bhdayton/necksum.htm

[49] Tom Denton, Florent Hivert, Anne Schilling, and Nicolas M. Thiéry. On the representation theory of finite J-trivial monoids. *Sém. Lothar. Combin.* **64** (2010/11), Art. B64d, 44 pp. https://www.emis.de/journals/SLC/wpapers/s64dehiscth.html

[50] Jacques Désarménien, Michelle L. Wachs. Descent classes of permutations with a given number of fixed points. *Journal of Combinatorial Theory, Series A* **64**, Issue 2, pp. 311–328. https://doi.org/10.1016/0097-3165(93)90100-M

[51] Persi Diaconis, Michael Mc Grath, Jim Pitman. Riffle shuffles, cycles, and descents. *Combinatorica* **15**(1), 1995, pp. 11–29. https://doi.org/10.1007/bf01294457

[52] Persi Diaconis, C.Y. Amy Pang and Arun Ram. Hopf algebras and Markov chains: Two examples and a theory. *J. Algebraic Combin.* **39**, Issue 3, May 2014, 527–585. A newer version is available at https://amypang.github.io/papers/hpmc.pdf

[53] Francesco Dolce, Antonio Restivo, Christophe Reutenauer. On generalized Lyndon words. Theoretical Computer Science **777** (2019), 232–242. Also available at arXiv:1812.04515v1.

[54] Francesco Dolce, Antonio Restivo, Christophe Reutenauer. Some variations on Lyndon words. arXiv:1904.00954v1.

[55] William F. Doran IV. A Proof of Reutenauer's $-q_{(n)}$ Conjecture. *J. Combin. Theory, Ser. A* **74** (1996), 342–344. https://doi.org/10.1006/jcta.1996.0056

[56] Andreas W. M. Dress, and Christian Siebeneicher. On the number of solutions of certain linear diophantine equations. *Hokkaido Math. J.* **19** (1990), pp. 385–401. http://www.math.sci.hokudai.ac.jp/hmj/page/19-3/pdf/HMJ_19_3_1990_385-401.pdf

[57] Andreas W. M. Dress, and Christian Siebeneicher. The Burnside Ring of the Infinite Cyclic Group and Its Relations to the Necklace Algebra, λ-Rings, and the Universal Ring of Witt Vectors. *Advances in Mathematics* **78** (1989), 1–41. https://doi.org/10.1016/0001-8708(89)90027-3

[58] Gérard Duchamp, Florent Hivert, and Jean-Yves Thibon. Noncommutative symmetric functions VI. Free quasi-symmetric functions and related algebras. *Internat. J. Algebra Comput.* **12** (2002), 671–717. A preprint is available at http://monge.univ-mlv.fr/~hivert/PAPER/NCSF6.ps.

[59] Gérard H. E. Duchamp, Nguyen Hoang-Nghia, Thomas Krajewski, Adrian Tanasa. Recipe theorem for the Tutte polynomial for matroids, renormalization group-like approach. *Advances in Applied Mathematics* **51**(3), 345—358. https://doi.org/10.1016/j.aam.2013.04.006

[60] Gérard Henry Edmond Duchamp, Vincel Hoang Ngoc Minh, Christophe Tollu, Bùi Chiên, Nguyen Hoang Nghia. Combinatorics of φ-deformed stuffle Hopf algebras. arXiv:1302.5391v7.

[61] Tobias Dyckerhoff. Hall Algebras - Bonn, Wintersemester 14/15. Lecture notes, version February 5, 2015. https://web.archive.org/web/20150601115158/http://www.math.uni-bonn.de/people/dyckerho/notes.pdf

[62] Jack Edmonds. Submodular functions, matroids, and certain polyhedra. In: Combinatorial Structures and their Applications (Proc. Calgary Internat. Conf., Calgary, Alta., 1969), Gordon and Breach, New York, 1970, pp. 66–87; reprinted in Combinatorial optimization: Eureka, you shrink!, pp. 11–26, *Lecture Notes in Comput. Sci.* **2570**, Springer, Berlin, 2003. https://doi.org/10.1007/3-540-36478-1_2

[63] Eric S. Egge. *An Introduction to Symmetric Functions and Their Combinatorics*. Student Mathematical Library **91**, American Mathematical Society, 2019. https://bookstore.ams.org/stml-91

[64] Richard Ehrenborg. On posets and Hopf algebras. *Adv. Math.* **119** (1996), 1–25. https://doi.org/10.1006/aima.1996.0026

[65] David Eisenbud. Commutative Algebra with a View Toward Algebraic Geometry. *Graduate Texts in Mathematics* **150**, Springer 1995. https://doi.org/10.1007/978-1-4612-5350-1

[66] Sergi Elizalde, Justin M. Troyka. Exact and asymptotic enumeration of cyclic permutations according to descent set. *J. Combin. Theory, Ser. A* **165** (2019), 360–391. Also available at arXiv:1710.05103v3.

[67] Alexander P. Ellis, and Mikhail Khovanov. The Hopf algebra of odd symmetric functions. *Adv. Math.* **231** (2012), 965–999. A newer version is available as arXiv:1107.5610v2.

[68] Brittney Ellzey. On Chromatic Quasisymmetric Functions of Directed Graphs. PhD thesis, University of Miami, 2018. https://scholarlyrepository.miami.edu/oa_dissertations/2091

[69] Pavel Etingof, Oleg Golberg, Sebastian Hensel, Tiankai Liu, Alex Schwendner, Dmitry Vaintrob, and Elena Yudovina. Introduction to representation theory. *Student Mathematical Library* **59**, Amer. Math. Soc., Providence, RI, 2011. http://www-math.mit.edu/~etingof/repb.pdf . (Parts of this book appear in arXiv:0901.0827v5.) A newer version is available at http://www-math.mit.edu/~etingof/reprbook.pdf .

[70] Loic Foissy. Algèbres de Hopf combinatoires. http://loic.foissy.free.fr/pageperso/Hopf.pdf

[71] Loic Foissy. Free and cofree Hopf algebras. *Journal of Pure and Applied Algebra* **216**, Issue 2, February 2012, 480–494. https://doi.org/10.1016/j.jpaa.2011.07.010 . A preprint is arXiv:1010.5402v3.

[72] Harold Fredricksen, James Maiorana. Necklaces of beads in k colors and k-ary de Bruijn sequences. *Discrete Mathematics* **23** (1978), 207–210. https://doi.org/10.1016/0012-365X(78)90002-X

[73] William Fulton. Young Tableaux. *London Mathematical Society Student Texts* **35**, Cambridge University Press, Cambridge-New York, 1997. https://doi.org/10.1017/CBO9780511626241

[74] Adriano M. Garsia. Permutation q-enumeration with the Schur row adder. *PU. M. A. (Pure Mathematics and Applications)* **21** (2010), No. 2, 233–248. http://puma.dimai.unifi.it/21_2/7_Garsia.pdf (also mirrored at http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.432.8196&rep=rep1&type=pdf).

[75] Vesselin Gasharov. Incomparability graphs of (3+1)-free posets are s-positive. Proceedings of the 6th Conference on Formal Power Series and Algebraic Combinatorics (New Brunswick, NJ, 1994). *Discrete Math.* **157** (1996), 193–197. https://doi.org/10.1016/S0012-365X(96)83014-7

[76] Vesselin Gasharov. A Short Proof of the Littlewood-Richardson Rule. *European Journal of Combinatorics*, Volume 19, Issue 4, May 1998, Pages 451–453. https://doi.org/10.1006/eujc.1998.0212

[77] Israel M. Gelfand, Daniel Krob, Alain Lascoux, Bernard Leclerc, Vladimir S. Retakh, Jean-Yves Thibon. Noncommutative symmetric functions. *Adv. Math.* **112** (1995), 218–348. https://doi.org/10.1006/aima.1995.1032
A preprint is available as arXiv:hep-th/9407124v1.

[78] M. Gerstenhaber, S.D. Schack. The shuffle bialgebra and the cohomology of commutative algebras. *Journal of Pure and Applied Algebra* **70** (1991), 263–272. https://doi.org/10.1016/0022-4049(91)90073-B

[79] Ira M. Gessel. Multipartite P-partitions and inner products of skew Schur functions. Combinatorics and algebra (Boulder, Colo., 1983), 289–317, *Contemp. Math.* **34**, Amer. Math. Soc., Providence, RI, 1984. http://people.brandeis.edu/~gessel/homepage/papers/multipartite.pdf

[80] Ira M. Gessel. A Historical Survey of P-Partitions. 2015, arXiv:1506.03508v1. Published in: Patricia Hersh, Thomas Lam, Pavlo Pylyavskyy and Victor Reiner (eds.), *The Mathematical Legacy of Richard P. Stanley*, Amer. Math. Soc., Providence, RI, 2016, pp. 169–188.

[81] Ira M. Gessel, Antonio Restivo, Christophe Reutenauer. A Bijection between Words and Multisets of Necklaces. *European Journal of Combinatorics* **33** (2012), pp. 1537–1546. https://doi.org/10.1016/j.ejc.2012.03.016

[82] Ira M. Gessel, Christophe Reutenauer. Counting Permutations with Given Cycle Structure and Descent Set. *Journal of Combinatorial Theory, Series A* **64** (1993), 189–215. https://doi.org/10.1016/0097-3165(93)90095-P

[83] Ira M. Gessel, X.G. Viennot. Determinants, Paths, and Plane Partitions. preprint, 1989, http://people.brandeis.edu/~gessel/homepage/papers/pp.pdf

[84] Andrew Granville. Number Theory Revealed: A Masterclass. *Number Theory Revealed: The Series* **#1B**, American Mathematical Society 2019.

[85] Darij Grinberg. Double posets and the antipode of QSym. arXiv:1509.08355v3.

[86] Darij Grinberg. A constructive proof of Orzech's theorem. Preprint, 20 November 2016. https://www.cip.ifi.lmu.de/~grinberg/algebra/orzech.pdf

[87] Frank D. Grosshans, Gian-Carlo Rota, Joel A. Stein. Invariant Theory and Superalgebras. *CBMS Regional Conference Series in Mathematics* **69**, American Mathematical Society, 1987. https://bookstore.ams.org/cbms-69

[88] A.M. Hamel, I.P. Goulden. Planar Decompositions of Tableaux and Schur Function Determinants. *Europ. J. Combinatorics* **16** (1995), 461–477. https://doi.org/10.1016/0195-6698(95)90002-0

[89] Michiel Hazewinkel. The algebra of quasi-symmetric functions is free over the integers. *Adv. Math.* **164** (2001), 283–300. https://doi.org/10.1006/aima.2001.2017

[90] Michiel Hazewinkel. Witt vectors. Part 1. In: M. Hazewinkel (ed.), *Handbook of Algebra* **6**, Elsevier 2009. Also available at arXiv:0804.3888v1.

[91] Michiel Hazewinkel. The Leibniz-Hopf Algebra and Lyndon Words. *Preprint AM CWI* **9612** (1996). http://oai.cwi.nl/oai/asset/4828/04828D.pdf

[92] Michiel Hazewinkel. Chen-Fox-Lyndon Factorization for Words over Partially Ordered Sets. *Journal of Mathematical Sciences* **131** (12-2005), Issue 6, 6027–6031. https://doi.org/10.1007/s10958-005-0458-7

[93] Michiel Hazewinkel, Nadiya Gubareni, and Vladimir V. Kirichenko. Algebras, rings and modules. Lie algebras and Hopf algebras. *Mathematical Surveys and Monographs* **168**. American Mathematical Society, Providence, RI, 2010.

[94] Robert Henderson. The Algebra Of Multiple Zeta Values. http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.227.5432

[95] Lars Hesselholt. Lecture notes on Witt vectors. http://www.math.nagoya-u.ac.jp/~larsh/papers/s03/wittsurvey.ps

[96] Lars Hesselholt. The big de Rham–Witt complex. *Acta Math.* **214** (2015), 135–207. https://doi.org/10.1007/s11511-015-0124-y

[97] Florent Hivert. An introduction to combinatorial Hopf algebras: examples and realizations. *Nato Advanced Study Institute School on Physics and Computer Science, 2005, october, 17–29, Cargese, France.* http://www-igm.univ-mlv.fr/~hivert/PAPER/Cargese.pdf

[98] Florent Hivert, Jean-Christophe Novelli and Jean-Yves Thibon. Commutative combinatorial Hopf algebras. *J. Algebraic Combin.* **28** (2008), no. 1, 65–95. https://doi.org/10.1007/s10801-007-0077-0
Also available as arXiv:math/0605262v1.

[99] ———. The algebra of binary search trees. *Theoret. Comput. Sci.* **339** (2005), no. 1, 129–165. https://doi.org/10.1016/j.tcs.2005.01.012
A preprint appears as arXiv:math/0401089v2.

[100] ———. Trees, functional equations, and combinatorial Hopf algebras. *European J. Combin.* **29** (2008), no. 7, 1682–1695. https://doi.org/10.1016/j.ejc.2007.09.005
A preprint appears as arXiv:math/0701539v1.

[101] Michael E. Hoffman. Combinatorics of rooted trees and Hopf algebras. *Trans. AMS* **355** (2003), 3795–3811. https://doi.org/10.1090/S0002-9947-03-03317-8

[102] ———. A character on the quasi-symmetric functions coming from multiple zeta values. *The Electronic Journal of Combinatorics* **15** (2008), R97. http://www.combinatorics.org/ojs/index.php/eljc/article/view/v15i1r97

[103] Brandon Humpert, and Jeremy L. Martin. The incidence Hopf algebra of graphs. *SIAM Journal on Discrete Mathematics* **26**, no. 2 (2012), 555–570. Also available as arXiv:1012.4786v3.

[104] Gordon James and Martin Liebeck. Representations and characters of groups. 2nd edition, Cambridge University Press, Cambridge-New York, 2001.

[105] Emma Yu Jin. Outside nested decompositions of skew diagrams and Schur function determinants. *European Journal of Combinatorics* **67** (2018), 239–267. https://doi.org/10.1016/j.ejc.2017.08.007 . A preprint is available at http://www.emmayujin.at/Pubs/Jin18.pdf.

[106] S.A. Joni and Gian-Carlo Rota. Coalgebras and bialgebras in combinatorics. *Studies in Applied Mathematics* **61** (1979), 93–139. https://doi.org/10.1002/sapm197961293

[107] Christian Kassel. Quantum groups. *Graduate Texts in Mathematics* **155**. Springer, Berlin, 1995.

[108] Sergei V. Kerov. Asymptotic representation theory of the symmetric group and its applications in analysis. *Translations of Mathematical Monographs* **219**. American Mathematical Society, Providence, RI, 2003.

[109] Anatol N. Kirillov, Arkadiy D. Berenstein. Groups generated by involutions, Gelfand-Tsetlin patterns and the combinatorics of Young tableaux. *Algebra i Analiz* **7** (1995), issue 1, 92–152. A preprint is available at http://pages.uoregon.edu/arkadiy/bk1.pdf

[110] T. Klein. The multiplication of Schur-functions and extensions of p-modules. *J. London Math. Soc.* **43** (1968), 280–284. https://doi.org/10.1112/jlms/s1-43.1.280

[111] Donald E. Knuth. Permutations, matrices, and generalized Young tableaux. *Pacific J. Math.* **34**, Number 3 (1970), 709–727. https://projecteuclid.org/euclid.pjm/1102971948

[112] Donald E. Knuth. The Art of Computer Programming, Volume 4A: Combinatorial Algorithms, Part 1. Pearson 2011. See https://www-cs-faculty.stanford.edu/~knuth/taocp.html for errata.

[113] Donald Knutson. λ-Rings and the Representation Theory of the Symmetric Group. *Lecture Notes in Mathematics* **308**, Springer, Berlin-Heidelberg-New York 1973. https://doi.org/10.1007/BFb0069217

[114] Manfred Krause. A Simple Proof of the Gale-Ryser Theorem. *The American Mathematical Monthly* **103** (1996), 335–337. https://doi.org/10.2307/2975191

[115] Daniel Krob. Eléments de combinatoire. Magistère 1-ère année, Ecole Normale Supérieure, version 1.0, Novembre 1995. http://krob.cesames.net/IMG/ps/combi.ps

[116] Manfred Kufleitner. On Bijective Variants of the Burrows-Wheeler Transform. Presented at the Prague Stringology Conference 2009 (PSC 2009). arXiv:0908.0239v1.

[117] Andrius Kulikauskas, Jeffrey Remmel. Lyndon words and transition matrices between elementary, homogeneous and monomial symmetric functions. *Electronic Journal of Combinatorics* **13** (2006), Research Paper ♯R18. http://www.combinatorics.org/ojs/index.php/eljc/article/view/v13i1r18

[118] Kalle Kytölä. Introduction to Hopf algebras and representations. Lecture notes, Spring 2011. https://math.aalto.fi/~kkytola/files_KK/lectures_files_KK/Hopf-lecture_notes.pdf

[119] Dan Laksov, Alain Lascoux, Piotr Pragacz, and Anders Thorup. The LLPT Notes. Edited by A. Thorup, 1995–2018. http://web.math.ku.dk/noter/filer/sympol.pdf

[120] Thomas Lam, Aaron Lauve, and Frank Sottile. Skew Littlewood-Richardson rules from Hopf Algebras. *DMTCS Proceedings, 22nd International Conference on Formal Power Series and Algebraic Combinatorics (FPSAC 2010)* **2010**, 355–366. A preprint can also be found at arXiv:0908.3714v3.

[121] Thomas Lam, and Pavlo Pylyavskyy. Combinatorial Hopf algebras and K-homology of Grassmanians. *International Mathematics Research Notices*, **2007** (2007), rnm 125, 48 pages. A preprint is arXiv:0705.2189v1.

[122] Sergei K. Lando. On a Hopf Algebra in Graph Theory. *Journal of Combinatorial Theory, Series B* **80** (2000), 104–121. https://doi.org/10.1006/jctb.2000.1973

[123] Aaron D. Lauda, Heather M. Russell. Oddification of the cohomology of type A Springer varieties. *International Math Research Notices* **2014**, No. 17, 4822–4854. A preprint is arXiv:1203.0797v1.

[124] Hartmut Laue. Freie algebraische Strukturen. Lecture notes, Mathematisches Seminar der Universität Kiel 2013, version 16 Sep 2013. http://www.uni-kiel.de/math/algebra/laue/vorlesungen/frei/freiealgstr.pdf

[125] Aaron Lauve and Sarah K. Mason. QSym over Sym has a stable basis. FPSAC 2010, San Francisco, USA. *DMTCS proc. AN* **2010**, 367–378. Also available as arXiv:1003.2124v1.

[126] Marc van Leeuwen. Schur functions and alternating sums. *Electronic Journal of Combinatorics* **11(2)** A5 (2006). Also available at http://www-math.univ-poitiers.fr/~maavl/pdf/alt-Schur.pdf.

[127] Marc van Leeuwen. Flag varieties, and interpretations of Young tableau algorithms. *Journal of Algebra* **224** (2000). Also available at http://wwwmathlabo.univ-poitiers.fr/~maavl/pdf/geometry.pdf

[128] Marc van Leeuwen. An application of Hopf-Algebra techniques to representations of finite Classical Groups. *Journal of Algebra* **140**, Issue 1, 15 June 1991, pp. 210–246. Also available at http://wwwmathlabo.univ-poitiers.fr/~maavl/pdf/Hopf.pdf

[129] Marc van Leeuwen. The Littlewood-Richardson rule, and related combinatorics. *Math. Soc. of Japan Memoirs* **11**, Interaction of Combinatorics and Representation Theory. Also available at http://wwwmathlabo.univ-poitiers.fr/~maavl/pdf/lrr.pdf

[130] Marc van Leeuwen. The Robinson-Schensted and Schützenberger algorithms, an elementary approach. *Electronic Journal of Combinatorics*, Foata Festschrift, **3** (no. 2), R15 (1996). Also available at http://wwwmathlabo.univ-poitiers.fr/~maavl/pdf/foata-fest.pdf

[131] Ji Li. Prime Graphs and Exponential Composition of Species. *Journal of Combinatorial Theory, Series A* **115**, Issue 8, November 2008, 1374–1401. See arXiv:0705.0038v4 for a preprint.

[132] Ricky Ini Liu. A simplified Kronecker rule for one hook shape. *Proc. Amer. Math. Soc.* **145** (2017), pp. 3657–3664. https://doi.org/10.1090/proc/13692 See arXiv:1412.2180v1 for a preprint.

[133] Arunas Liulevicius. Arrows, symmetries and representation rings. *Journal of Pure and Applied Algebra* **19** (1980), 259–273. https://doi.org/10.1016/0022-4049(80)90103-6

[134] Jean-Louis Loday. Cyclic Homology. *Grundlehren der mathematischen Wissenschaften* **301**, 2nd edition, Springer, Berlin-Heidelberg 1998.

[135] Jean-Louis Loday. Série de Hausdorff, idempotents Eulériens et algèbres de Hopf. *Expo. Math.* **12** (1994), 165–178. http://www-irma.u-strasbg.fr/~loday/PAPERS/94Loday%28Eulerien%29.pdf

[136] Jean-Louis Loday and María O. Ronco. Combinatorial Hopf algebras. Quanta of maths, *Clay Math. Proc.* **11**, 347–383, Amer. Math. Soc., Providence, RI, 2010. http://www-irma.u-strasbg.fr/~loday/PAPERS/2011LodayRonco(CHA).pdf

[137] ______. Hopf algebra of the planar binary trees. *Adv. Math.* **139** (1998), no. 2, 293–309. https://doi.org/10.1006/aima.1998.1759

[138] Nicholas A. Loehr. Bijective Combinatorics. CRC Press, 2011. See http://www.math.vt.edu/people/nloehr/bijbook.html for errata.

[139] M. Lothaire. *Combinatorics on words*. Corrected printing, Cambridge University Press, 1997.

[140] Kurt Luoto, Stefan Mykytiuk, Stephanie van Willigenburg. An introduction to quasisymmetric Schur functions – Hopf algebras, quasisymmetric functions, and Young composition tableaux. Springer, May 23, 2013. http://www.math.ubc.ca/~steph/papers/QuasiSchurBook.pdf

[141] R.C. Lyndon. On Burnside's Problem. *Transactions of the AMS* **77**, 202–215. https://doi.org/10.1090/S0002-9947-1954-0064049-X

[142] Ian Grant Macdonald. Symmetric functions and Hall polynomials. 2nd edition, Oxford University Press, Oxford-New York, 1995.

[143] I.G. Macdonald. Schur functions : theme and variations. *Publ. I.R.M.A. Strasbourg*, **1992**, 498/S-27, Actes 28 e Seminaire Lotharingien, 5–39. https://www.emis.de/journals/SLC/opapers/s28macdonald.html

[144] Manuel Maia, Miguel Méndez. On the arithmetic product of combinatorial species. *Discrete Mathematics* **308**, Issue 23, 6 December 2008, 5407–5427. https://doi.org/10.1016/j.disc.2007.09.062 .
See arXiv:math/0503436v2 for a preprint.

[145] Claudia Malvenuto. Produits et coproduits des fonctions quasi-symétriques et de l'algèbre des descents. PhD dissertation, Univ. du Québéc à Montreal, 1993. http://lacim.uqam.ca/wp-content/uploads/Publications/16.pdf

[146] Clauda [sic] Malvenuto and Christophe Reutenauer. Duality between quasi-symmetric functions and the Solomon descent algebra. *J. Algebra* **177** (1995), 967–982. https://doi.org/10.1006/jabr.1995.1336

[147] Claudia Malvenuto and Christophe Reutenauer. Plethysm and conjugation of quasi-symmetric functions. *Discrete Mathematics* **193**, Issues 1–3, 28 November 1998, 225–233. https://doi.org/10.1016/S0012-365X(98)00142-3

[148] Claudia Malvenuto and Christophe Reutenauer. A self paired Hopf algebra on double posets and a Littlewood-Richardson rule. *Journal of Combinatorial Theory, Series A* **118** (2011), 1322–1333. https://doi.org/10.1016/j.jcta.2010.10.010

[149] Dominique Manchon. Hopf algebras, from basics to applications to renormalization. *Comptes Rendus des Rencontres Mathematiques de Glanon* 2001 (published in 2003). arXiv:math/0408405v2.

[150] Marco Manetti. A voyage round coalgebras. 27 June 2016. https://www1.mat.uniroma1.it/people/manetti/dispense/voyage.pdf

[151] Laurent Manivel. Chern classes of tensor products. arXiv:1012.0014v1.

[152] Peter R.W. McNamara, Ryan E. Ward. Equality of P-partition generating functions. arXiv:1210.2412v2.

[153] Pierre-Loïc Méliot. Representation Theory of Symmetric Groups. *Discrete Mathematics and its Applications*, CRC Press 2017.

[154] Anthony Mendes, Jeffrey Remmel. Counting with Symmetric Functions. *Developments in Mathematics* **43**, Springer 2015.

[155] Miguel Mendez. *MathOverflow answer #139482*. http://mathoverflow.net/a/139482/.

[156] John W. Milnor and John C. Moore. On the structure of Hopf algebras. *The Annals of Mathematics, Second Series* **81**, No. 2 (Mar., 1965), 211–264. https://doi.org/10.2307/1970615

[157] Susan Montgomery. Hopf algebras and their actions on rings. *Regional Conference Series in Mathematics* **82**, Amer. Math. Soc., Providence, RI, 2010. https://bookstore.ams.org/cbms-82

[158] Jack Morava. Homotopy-theoretically enriched categories of noncommutative motives. *Research in the Mathematical Sciences* **2** (2015), no. 8. https://doi.org/10.1186/s40687-015-0028-7

[159] Eduardo Moreno. On the theorem of Fredricksen and Maiorana about de Bruijn sequences. *Advances in Applied Mathematics* **33**, Issue 2, August 2004, 413–415. https://doi.org/10.1016/j.aam.2003.10.002

[160] Eduardo Moreno, Dominique Perrin. Corrigendum to "On the theorem of Fredricksen and Maiorana about de Bruijn sequences" [Adv. in Appl. Math. 33 (2) (2004) 413–415]. *Advances in Applied Mathematics* **62**, January 2015, Pages 184–187. http://www.sciencedirect.com/science/article/pii/S0196885814000918

[161] Jeremy L. Martin, Matthew Morin, Jennifer D. Wagner. On distinguishing trees by their chromatic symmetric functions. *Journal of Combinatorial Theory, Series A* **115**, Issue 2, February 2008, 237–253. https://doi.org/10.1016/j.jcta.2007.05.008

[162] Robert Morris. Umbral Calculus and Hopf Algebras. *Contemporary Mathematics* **6**, AMS, Providence 1982. https://bookstore.ams.org/conm-6

[163] Jakob Oesinghaus. Quasisymmetric functions and the Chow ring of the stack of expanded pairs. *Res. Math. Sci.* **6** (2019), no. 5. https://doi.org/10.1007/s40687-018-0168-7 . A preprint is arXiv:1806.10700v1.

[164] James Oxley. Matroid theory. Oxford University Press, Oxford-New York, 1992.

[165] Igor Pak, Alexander Postnikov. Oscillating Tableaux, $S_p \times S_q$-modules, and Robinson-Schensted-Knuth Correspondence. Updated (yet unfinished) version of FPSAC 1996 abstract, January 15, 1994. http://math.mit.edu/~apost/papers/osc.pdf

[166] Frédéric Patras. La décomposition en poids des algèbres de Hopf. *Annales de l'institut Fourier* **43**, no 4 (1993), 1067–1087. https://eudml.org/doc/75026

[167] F. Patras. L'algèbre des descentes d'une bigèbre graduée. *Journal of Algebra* **170** (1994), 547–566. https://doi.org/10.1006/jabr.1994.1352

[168] Frédéric Patras, Christophe Reutenauer. On Dynkin and Klyachko idempotents in graded bialgebras. *Advanced in Applied Mathematics* **28**, Issues 3–4, April 2002, 560–579. https://doi.org/10.1006/aama.2001.0795

[169] Frédéric Patras, Christophe Reutenauer. Higher Lie idempotents. *J. Algebra* **222** (1999), no. 1, 51–64. https://doi.org/10.1006/jabr.1999.7887

[170] Rebecca Patrias. Antipode formulas for combinatorial Hopf algebras. arXiv:1501.00710v2. Published in: *The Electronic Journal of Combinatorics* **23**, Issue 4 (2016), P4.30. http://www.combinatorics.org/ojs/index.php/eljc/article/view/v23i4p30/

[171] Victor Prasolov. Problems and theorems in linear algebra. *Translations of mathematical monographs* **134**, 1st edition 1994, AMS. http://www2.math.su.se/~mleites/books/prasolov-1994-problems.pdf

[172] Stéphane Poirier, Christophe Reutenauer. Algèbres de Hopf de tableaux. *Ann. Sci. Math. Québec* **19** (1995), no. 1, 79–90. http://www.lacim.uqam.ca/~christo/Publi%C3%A9s/1995/Alg%C3%A8bres%20de%20Hopf%20de%20tableaux.pdf

[173] Alexander Postnikov. Permutohedra, associahedra, and beyond. *Int. Math. Res. Notices* **2009**, No. 6, pp. 1026–1106. A preprint appears at https://math.mit.edu/~apost/papers/permutohedron.pdf and as arXiv:math/0507163v1.

[174] Amritanshu Prasad. An Introduction to Schur Polynomials. *Graduate J. Math.* **4** (2019), 62–84. https://www.gradmath.org/wp-content/uploads/2020/01/Prasad-GJM2019.pdf . A preprint appears at arXiv:1802.06073v2.

[175] Pavlo Pylyavskyy. Comparing products of Schur functions and quasisymmetric functions. PhD dissertation, MIT, 2007. https://dspace.mit.edu/handle/1721.1/38957

[176] David E. Radford. Hopf algebras. *Series on Knots and Everything* **49**. World Scientific, 2012. https://doi.org/10.1142/8055

[177] David E. Radford. A Natural Ring Basis for the Shuffle Algebra and an Application to Group Schemes. *Journal of Algebra* **58** (1979), 432–454. https://doi.org/10.1016/0021-8693(79)90171-6

[178] Nathan Reading. Lattice congruences, fans and Hopf algebras. *Journal of Combinatorial Theory, Series A* **110**, Issue 2, May 2005, pp. 237–273. https://doi.org/10.1016/j.jcta.2004.11.001 . A preprint is arXiv:math/0402063v1.

[179] Victor Reiner. Signed permutation statistics and cycle type. *European J. Combin.* **14** (1993), no. 6, 569–579. https://doi.org/10.1006/eujc.1993.1059

[180] Victor Reiner, Kristin M. Shaw, and Stephanie van Willigenburg. Coincidences among skew Schur functions. arXiv:math/0602634v4. (Update of a paper published in *Advances in Mathematics*, **216(1)**:118–152, 2007.)

[181] Jeffrey B. Remmel. The combinatorics of (k, ℓ)-hook Schur functions. In: C. Greene (ed.), *Combinatorics and algebra*, Proceedings of the AMS-IMS-SIAM joint summer research conference in the mathematical sciences on combinatorics and algebra, Colorado, Boulder, 1983, *Contemporary Mathematics* **34**, 1984, 253–287.

[182] Christophe Reutenauer. Free Lie Algebras. *London Mathematical Society Monographs, New Series* **7**. Clarendon Press, Oxford 1993.

[183] Mercedes H. Rosas. The Kronecker Product of Schur Functions Indexed by Two-Row Shapes or Hook Shapes. *Journal of Algebraic Combinatorics* **14** (2001), 153–173. https://doi.org/10.1023/A:1011942029902
A preprint is arXiv:math/0001084v1.

[184] Mercedes H. Rosas, Bruce E. Sagan. Symmetric functions in noncommuting variables. *Transactions of the American Mathematical Society* **358**, 183–214. https://doi.org/10.1090/S0002-9947-04-03623-2

[185] Joseph P.S. Kung, Gian-Carlo Rota. Gian-Carlo Rota on Combinatorics: Introductory Papers and Commentaries. Birkhäuser 1995.

[186] Bruce E. Sagan. The symmetric group: representations, combinatorial algorithms, and symmetric functions. 2nd edition, Springer, New York-Berlin-Heidelberg 2001. See https://users.math.msu.edu/users/bsagan/Books/Sym/errata.pdf for errata.

[187] Bruce E. Sagan. Combinatorics: The Art of Counting. Draft of a textbook, 2020. https://users.math.msu.edu/users/bsagan/Books/Aoc/aocAMS.pdf

[188] Bruce E. Sagan, Richard P. Stanley. Robinson-Schensted Algorithms for Skew Tableaux. *Journal of Combinatorial Theory, Series A* **55** (1990), 161–193. https://doi.org/10.1016/0097-3165(90)90066-6

[189] Steven V. Sam. Notes for Math 740 (Symmetric Functions), 27 April 2017. https://www.math.wisc.edu/~svs/740/notes.pdf

[190] Olivier Schiffmann. Lectures on Hall algebras. arXiv:math/0611617v2.

[191] William R. Schmitt. Incidence Hopf algebras. *Journal of Pure and Applied Algebra* **96** (1994), 299–330. https://doi.org/10.1016/0022-4049(94)90105-8 . A preprint appears at http://home.gwu.edu/~wschmitt/papers/iha.pdf

[192] William R. Schmitt. Antipodes and Incidence Coalgebras. *Journal of Combinatorial Theory, Series A* **46** (1987), 264–290. https://doi.org/10.1016/0097-3165(87)90006-9

[193] William R. Schmitt. Expository notes, specifically "A concrete introduction to category theory" and "Notes on modules and algebras". http://home.gwu.edu/~wschmitt/

[194] ———. Hopf algebras of combinatorial structures. *Canadian Journal of Mathematics* **45** (1993), 412–428. https://doi.org/10.4153/CJM-1993-021-5 . A preprint appears at http://home.gwu.edu/~wschmitt/papers/hacs.pdf

[195] I. Schur. Arithmetische Eigenschaften der Potenzsummen einer algebraischen Gleichung. *Compositio Mathematica* **4** (1937), 432–444. http://www.numdam.org/item?id=CM_1937__4__432_0

[196] Christoph Schweigert. Hopf algebras, quantum groups and topological field theory. Lecture notes, Winter term 2014/15, Hamburg. Version of 16 May 2015. http://www.math.uni-hamburg.de/home/schweigert/ws12/hskript.pdf

[197] Jean-Pierre Serre. Linear representations of finite groups. Springer, Berlin-Heidelberg-New York, 1977. https://doi.org/10.1007/978-1-4684-9458-7

[198] John Shareshian and Michelle L. Wachs. Chromatic quasisymmetric functions and Hessenberg varieties. In: A. Björner, F. Cohen, C. De Concini, C. Procesi, M. Salvetti (Eds.), Configuration Spaces, *Publications of the Scuola Normale Superiore* **14**, Springer, Berlin-Heidelberg-New York 2013. A preprint is arXiv:1106.4287v3.

[199] John Shareshian and Michelle L. Wachs. Chromatic quasisymmetric functions. *Advances in Mathematics* **295** (2016), pp. 497–551. A preprint is arXiv:1405.4629v2.

[200] John Shareshian and Michelle L. Wachs. Eulerian quasisymmetric functions. *Advances in Mathematics* **225** (2010), pp. 2921–2966. https://doi.org/10.1016/j.aim.2010.05.009 . A preprint is arXiv:0812.0764v2

[201] Seth Shelley-Abrahamson. Hopf Modules and Representations of Finite Groups of Lie Type. Honors thesis, Stanford, May 2013. http://mathematics.stanford.edu/wp-content/uploads/2013/08/Shelley-Abrahamson-Honors-Thesis-2013.pdf

[202] Anatolii I. Shirshov. On Free Lie Rings. *Mat. Sbornik N.S.* **45** (87), (1958), no. 2, 113–122. Original at: http://mi.mathnet.ru/msb4963. Translation in: L.A. Bokut, V. Latyshev, I. Shestakov, E. Zelmanov (eds.), *Selected works of A.I. Shirshov*, Birkhäuser 2009.

[203] Richard P. Stanley. Ordered structures and partitions. *Memoirs of the Amer. Math. Soc.* **119**, American Mathematical Society, Providence, R.I., 1972. http://www-math.mit.edu/~rstan/pubs/pubfiles/9.pdf

[204] ______. Acyclic orientations of graphs. *Discrete Math.* **5** (1973), 171–178. Reprinted in: *Discrete Math.* **306** (2006), 905–909. https://doi.org/10.1016/j.disc.2006.03.010

[205] ______. A symmetric function generalization of the chromatic polynomial of a graph. *Adv. Math.* **111** (1995), 166–194. https://doi.org/10.1006/aima.1995.1020

[206] ______. Enumerative Combinatorics, Volumes 1 and 2. *Cambridge Studies in Advanced Mathematics*, **49** and **62**. Cambridge University Press, Cambridge, 2nd edition 2011 (volume 1) and 1st edition 1999 (volume 2).

[207] Shishuo Fu, Victor Reiner, Dennis Stanton, Nathaniel Thiem. The negative q-binomial. *The Electronic Journal of Combinatorics* **19**, Issue 1 (2012), P36. http://www.combinatorics.org/ojs/index.php/eljc/article/view/v19i1p36

[208] R. Steinberg. A geometric approach to the representations of the full linear group over a Galois field. *Trans. Amer. Math. Soc.* **71**, (1951), 274–282. https://doi.org/10.1090/S0002-9947-1951-0043784-0

[209] Jacob Steinhardt. Permutations with Ascending and Descending Blocks. *The Electronic Journal of Combinatorics* **17** (2010), #R14. https://www.combinatorics.org/ojs/index.php/eljc/article/view/v17i1r14

[210] John R. Stembridge. A concise proof of the Littlewood-Richardson rule. *The Electronic Journal of Combinatorics* **9**, 2002, N5. http://www.combinatorics.org/ojs/index.php/eljc/article/view/v9i1n5

[211] John Stembridge. Multiplicity-Free Products of Schur Functions. *Annals of Combinatorics* **5** (2001), 113–121. http://www.math.lsa.umich.edu/~jrs/papers/mfree.ps.gz

[212] Gilbert Strang. The algebra of Elimination. http://www-math.mit.edu/~gs/papers/Paper7_ver8.pdf.

[213] Moss E. Sweedler. Hopf algebras. W.A. Benjamin, New York, 1969.

[214] Mitsuhiro Takeuchi. Free Hopf algebras generated by coalgebras. *J. Math. Soc. Japan* **23** (1971), 561–582. http://projecteuclid.org/euclid.jmsj/1259849779

[215] Harry Tamvakis. The theory of Schur polynomials revisited. *Enseign. Math.* **58** (2012), 147–163. A preprint appears at http://www2.math.umd.edu/~harryt/papers/schurrev.pdf

[216] Jean-Yves Thibon. An Introduction to Noncommutative Symmetric Functions. Cargese lecture, October 2005. J.-P. Gazeau, J. Nesetril, B. Rovan (eds.): From Numbers and Languages to (Quantum) Cryptography, *NATO Security through Science Series: Information and Communication Security* **7**, IOS Press, 2007. Available at http://igm.univ-mlv.fr/~jyt/ARTICLES/cargese_thibon.ps.

[217] Nathaniel Thiem and C. Ryan Vinroot. On the characteristic map of finite unitary groups. *Advances in Mathematics* **210**, Issue 2, 1 April 2007, pp. 707–732. https://doi.org/10.1016/j.aim.2006.07.018 . A preprint is http://www.math.wm.edu/~vinroot/charunitary.pdf

[218] Hugh Thomas, Alexander Yong. An S_3-symmetric Littlewood-Richardson rule. *Math. Res. Lett.* **15** (2008), no. 5, 1027–1037. arXiv:0704.0817v1.

[219] Stijn Vermeeren. Sequences and nets in topology. Version of 11 September 2013. http://stijnvermeeren.be/download/mathematics/nets.pdf

[220] Michelle L. Wachs. Flagged Schur Functions, Schubert Polynomials, and Symmetrizing Operators. *Journal of Combinatorial Theory, Series A* **40** (1985), 276–289. https://doi.org/10.1016/0097-3165(85)90091-3

[221] Bartel Leendert van der Waerden. Algebra, Volume I. Translation of the 7th (German) edition. Springer 2003.

[222] Peter Webb. A Course in Finite Group Representation Theory. 23 February 2016. http://www-users.math.umn.edu/~webb/RepBook/

[223] Mark Wildon. Representation theory of the symmetric group. 5 April 2018. http://www.ma.rhul.ac.uk/~uvah099/teaching.html

[224] Mark Wildon. An involutive introduction to symmetric functions. 8 May 2020. http://www.ma.rhul.ac.uk/~uvah099/teaching.html

[225] Robert Wisbauer. Coalgebras and Bialgebras. *The Egyptian Mathematical Society, The Mathematical Sciences Research Centre (MSRC) Technical Reports* **No. 1**, 2004. http://www.math.uni-duesseldorf.de/~wisbauer/

[226] Qimh Richey Xantcha. Binomial Rings: Axiomatisation, Transfer, and Classification. arXiv:1104.1931v4.

[227] Andrey V. Zelevinsky. Representations of finite classical groups: a Hopf algebra approach. *Lecture Notes in Mathematics* **869**. Springer-Verlag, Berlin-New York, 1981.

[228] Andrey V. Zelevinsky. A Generalization of the Littlewood-Richardson Rule and the Robinson-Schensted-Knuth Correspondence. *Journal of Algebra* **69** (1981), 82–94. https://doi.org/10.1016/0021-8693(81)90128-9

[229] G.-S. Zhou, D.-M. Lu. Lyndon words for Artin-Schelter regular algebras. arXiv:1403.0385v1.

An index goes here!

13. Solutions to the exercises

This chapter contains solutions to the exercises scattered throughout the text. These solutions vary in level of detail (some of them are detailed, some only outline the most important steps, and many lie inbetween these two extremes), and sometimes in notation (as they have been written over a long timespan). They also have seen far less quality control than the main text, so typos and worse are to be expected. Comments and alternative solutions are welcome!

13.1. **Solution to Exercise 1.2.3.**

First solution to Exercise 1.2.3. This is analogous to the well-known fact that any nonunital nonassociative[383] $\mathbf{k}$-algebra has at most one multiplicative identity. If you can prove the latter fact by pure abstract nonsense (i.e., without referring to elements), then the same proof serves as a solution to Exercise 1.2.3 once all arrows are reversed (and all m's and u's are replaced by Δ's and ϵ's). Let us see how this works.

How does one classically prove that every nonunital nonassociative $\mathbf{k}$-algebra has at most one multiplicative identity? Let A be a nonunital nonassociative $\mathbf{k}$-algebra, and let 1 and $1'$ be two elements of A which could both serve as multiplicative identities. That is, every $a \in A$ satisfies both $1a = a1 = a$ and $1'a = a1' = a$. Now, applying $a1 = a$ to $a = 1'$ yields $1' \cdot 1 = 1'$. But applying $1'a = a$ to $a = 1$ yields $1' \cdot 1 = 1$. Comparing $1' \cdot 1 = 1$ with $1' \cdot 1 = 1'$ yields $1 = 1'$, and thus the multiplicative identity is unique.

This argument made use of elements, which we need to get rid of in order to be able to reverse the arrows. The idea is to replace every element α of A by the linear map $\mathbf{k} \to A$ which sends $1 \in \mathbf{k}$ to α. In terms of commutative diagrams, a nonunital nonassociative $\mathbf{k}$-algebra is a $\mathbf{k}$-module A endowed with a $\mathbf{k}$-linear map $m : A \otimes A \to A$ which is not a priori required to satisfy any properties. A multiplicative identity of A then corresponds to a $\mathbf{k}$-linear map $u : \mathbf{k} \to A$ making the diagram (1.1.2) commute. Let u and u' be two such $\mathbf{k}$-linear maps $\mathbf{k} \to A$. Instead of applying $a1 = a$ to $a = 1'$, we now need to take the commutative diagram

$$\begin{tikzcd}
A \otimes \mathbf{k} \arrow[d, "\mathrm{id} \otimes u"'] & A \arrow[l] \arrow[d, "\mathrm{id}"'] \\
A \otimes A \arrow[r, "m"] & A
\end{tikzcd}$$

(which commutes because u makes (1.1.2) commute, and corresponds to the axiom $a1 = a$) and pre-compose it with the morphism u' (which corresponds to the multiplicative identity $1'$), thus obtaining

(13.1.1)
$$\begin{tikzcd}
 & & \mathbf{k} . \arrow[dl, "u'"] \\
A \otimes \mathbf{k} \arrow[d, "\mathrm{id} \otimes u"'] & A \arrow[l] \arrow[d, "\mathrm{id}"'] & \\
A \otimes A \arrow[r, "m"] & A &
\end{tikzcd}$$

Similarly, the element-free version of applying $1'a = a$ to $a = 1$ yields the commutative diagram

(13.1.2)
$$\begin{tikzcd}
\mathbf{k} \arrow[dr, "u"] & & . \\
 & A \arrow[r] \arrow[d, "\mathrm{id}"'] & \mathbf{k} \otimes A \arrow[d, "u' \otimes \mathrm{id}"] \\
 & A & A \otimes A \arrow[l, "m"']
\end{tikzcd}$$

[383] The word "nonassociative" does not prohibit associativity; it simply means that associativity is not required. Similarly, "nonunital" does not force the nonexistence of a multiplicative identity.

The path $\mathbf{k} \xrightarrow{u'} A \longrightarrow A \otimes \mathbf{k} \xrightarrow{\mathrm{id} \otimes u} A \otimes A$ through the diagram (13.1.1) and the path $\mathbf{k} \xrightarrow{u} A \longrightarrow \mathbf{k} \otimes A \xrightarrow{u' \otimes \mathrm{id}} A \otimes A$ through the diagram (13.1.2) give the same map, since they both correspond to the element $1 \otimes 1'$ of $A \otimes A$. How can this be seen without referring to elements? Being a tautology at the level of elements, this must follow from formal properties of tensor products (without using axioms like the commutativity of (1.1.2)). And so it does: In fact, the three small quadrilaterals in the diagram

$$\begin{tikzcd}
 & & \mathbf{k} \arrow[dll, "u'"'] \arrow[d] \arrow[drr, "u"] & & \\
A \arrow[d] & & \mathbf{k} \otimes \mathbf{k} \arrow[dll, "u' \otimes \mathrm{id}"] \arrow[drr, "\mathrm{id} \otimes u"'] & & A \arrow[d] \\
A \otimes \mathbf{k} \arrow[drr, "\mathrm{id} \otimes u"'] & & & & \mathbf{k} \otimes A \arrow[dll, "u' \otimes \mathrm{id}"] \\
 & & A \otimes A & &
\end{tikzcd}$$

commute, and therefore so does the whole diagram. We can now piece this diagram together with the diagrams (13.1.1) and (13.1.2) (this is the diagrammatic equivalent of comparing $1 \cdot 1' = 1$ with $1 \cdot 1' = 1'$), and obtain

$$\begin{tikzcd}
 & & \mathbf{k} \arrow[dll, "u'"'] \arrow[d] \arrow[drr, "u"] & & & . \\
A \arrow[d] \arrow[ddddrr, bend right=60, "\mathrm{id}"'] & & \mathbf{k} \otimes \mathbf{k} \arrow[dll, "u' \otimes \mathrm{id}"] \arrow[drr, "\mathrm{id} \otimes u"'] & & A \arrow[d] \arrow[ddddll, bend left=60, "\mathrm{id}"] & \\
A \otimes \mathbf{k} \arrow[drr, "\mathrm{id} \otimes u"'] & & & & \mathbf{k} \otimes A \arrow[dll, "u' \otimes \mathrm{id}"] & \\
 & & A \otimes A \arrow[d, "m"] & & & \\
 & & A & & &
\end{tikzcd}$$

Following the outside quadrilateral of this commutative diagram yields $\mathrm{id} \circ u = \mathrm{id} \circ u'$, so that $u = u'$, which shows that the two maps u and u' are equal. The purely diagrammatic argument that we made can now be easily transformed into a solution of Exercise 1.2.3 by merely reversing arrows and replacing m and u by Δ and ϵ.

Second solution to Exercise 1.2.3. The exercise can also be solved using Sweedler notation, which makes some good practice in using the latter. Here is how the solution goes:

We must prove that there exists *at most one* $\mathbf{k}$-linear map $\epsilon : C \to \mathbf{k}$ such that the diagram (1.2.2) commutes. In other words, we must show that if ϵ_1 and ϵ_2 are two such maps, then $\epsilon_1 = \epsilon_2$.

So let ϵ_1 and ϵ_2 be two such maps. We must then show that $\epsilon_1 = \epsilon_2$.

We know that ϵ_1 is a $\mathbf{k}$-linear map $\epsilon : C \to \mathbf{k}$ such that the diagram (1.2.2) commutes. Thus, the diagram

$$\begin{tikzcd}
C \otimes \mathbf{k} \arrow[r] & C & \mathbf{k} \otimes C \arrow[l] \\
C \otimes C \arrow[u, "\mathrm{id} \otimes \epsilon_1"] & C \arrow[l, "\Delta"] \arrow[r, "\Delta"'] \arrow[u, "\mathrm{id}"] & C \otimes C \arrow[u, "\epsilon_1 \otimes \mathrm{id}"']
\end{tikzcd} \tag{13.1.3}$$

commutes. Let us denote the top-left horizontal arrow of this diagram by κ; thus, $\kappa : C \otimes \mathbf{k} \to C$ is the canonical isomorphism (sending each $c \otimes \lambda \in C \otimes \mathbf{k}$ to $\lambda c \in \mathbf{k}$). The commutativity of the left square in the commutative diagram (13.1.3) thus says that

$$\mathrm{id} = \kappa \circ (\mathrm{id} \otimes \epsilon_1) \circ \Delta. \tag{13.1.4}$$

We shall use Sweedler notation as in (1.2.3) to abbreviate formulas involving Δ. For each $c \in C$, we have

$$\begin{aligned}
c &= \underbrace{\mathrm{id}}_{\substack{=\kappa\circ(\mathrm{id}\otimes\epsilon_1)\circ\Delta \\ \text{(by (13.1.4))}}} (c) = (\kappa \circ (\mathrm{id}\otimes\epsilon_1) \circ \Delta)(c) = \kappa\left((\mathrm{id}\otimes\epsilon_1)\left(\underbrace{\Delta(c)}_{=\sum_{(c)} c_1\otimes c_2}\right)\right) \\
&= \kappa\left(\underbrace{(\mathrm{id}\otimes\epsilon_1)\left(\sum_{(c)} c_1 \otimes c_2\right)}_{=\sum_{(c)} \mathrm{id}(c_1)\otimes\epsilon_1(c_2)}\right) = \kappa\left(\sum_{(c)} \underbrace{\mathrm{id}(c_1)}_{=c_1} \otimes \epsilon_1(c_2)\right) \\
&= \kappa\left(\sum_{(c)} c_1 \otimes \epsilon_1(c_2)\right) \\
&= \sum_{(c)} \epsilon_1(c_2)\, c_1 \qquad \text{(by the definition of } \kappa\text{)}. \qquad (13.1.5)
\end{aligned}$$

We have obtained this equality from the commutativity of the left square in the commutative diagram (13.1.3). Similarly, from the commutativity of the right square in the same diagram, we can obtain the equality

$$c = \sum_{(c)} \epsilon_1(c_1)\, c_2.$$

The same argument (with the roles of ϵ_1 and ϵ_2 interchanged) yields

$$c = \sum_{(c)} \epsilon_2(c_1)\, c_2.$$

Applying the map ϵ_1 to both sides of this equality, we find

$$\epsilon_1(c) = \epsilon_1\left(\sum_{(c)} \epsilon_2(c_1)\, c_2\right) = \sum_{(c)} \epsilon_2(c_1)\, \epsilon_1(c_2) \qquad \text{(since the map } \epsilon_1 \text{ is } \mathbf{k}\text{-linear)}$$

On the other hand, applying the map ϵ_2 to both sides of the equality (13.1.5), we find

$$\begin{aligned}
\epsilon_2(c) &= \epsilon_2\left(\sum_{(c)} \epsilon_1(c_2)\, c_1\right) = \sum_{(c)} \underbrace{\epsilon_1(c_2)\, \epsilon_2(c_1)}_{\substack{=\epsilon_2(c_1)\epsilon_1(c_2) \\ \text{(since the multiplication} \\ \text{in } \mathbf{k} \text{ is commutative)}}} \qquad \text{(since the map } \epsilon_2 \text{ is } \mathbf{k}\text{-linear)} \\
&= \sum_{(c)} \epsilon_2(c_1)\, \epsilon_1(c_2).
\end{aligned}$$

Comparing these two equalities, we find $\epsilon_1(c) = \epsilon_2(c)$.

Forget that we fixed c. We thus have shown that $\epsilon_1(c) = \epsilon_2(c)$ for each $c \in C$. In other words, $\epsilon_1 = \epsilon_2$. This is precisely what we set out to prove. Thus, Exercise 1.2.3 is solved again.

13.2. **Solution to Exercise 1.3.4.** *Solution to Exercise 1.3.4.* (a) Exercise 1.3.4(a) is a classical result about algebras, and proven in various textbooks. Nevertheless, we shall give two solutions to it: one solution by elementwise computation, and another by formally manipulating homomorphisms (this is essentially what is called "diagram chasing", except that we are not going to draw any diagrams). Both solutions have their advantages, and it is useful to see them both.

First solution to Exercise 1.3.4(a). Let ξ be the canonical $\mathbf{k}$-module isomorphism $\mathbf{k} \to \mathbf{k} \otimes \mathbf{k}$. We have defined the multiplication map $m_{A\otimes B}$ of the $\mathbf{k}$-algebra $A \otimes B$ by

$$m_{A\otimes B} = (m_A \otimes m_B) \circ (\mathrm{id}_A \otimes T \otimes \mathrm{id}_B),$$

and the unit map $u_{A\otimes B}$ of the $\mathbf{k}$-algebra $A \otimes B$ by

$$u_{A\otimes B} = (u_A \otimes u_B) \circ \xi.$$

We now must prove that the $\mathbf{k}$-module $A \otimes B$, equipped with these two maps $m_{A\otimes B}$ and $u_{A\otimes B}$, is indeed a $\mathbf{k}$-algebra. In other words, we must show that the two diagrams

(13.2.1)
$$\begin{tikzcd}
 & A\otimes B\otimes A\otimes B\otimes A\otimes B \arrow[dl, "m_{A\otimes B}\otimes \mathrm{id}_{A\otimes B}"'] \arrow[dr, "\mathrm{id}_{A\otimes B}\otimes m_{A\otimes B}"] & \\
A\otimes B\otimes A\otimes B \arrow[dr, "m_{A\otimes B}"] & & A\otimes B\otimes A\otimes B \arrow[dl, "m_{A\otimes B}"'] \\
 & A\otimes B &
\end{tikzcd}$$

and

(13.2.2)
$$\begin{tikzcd}
A\otimes B\otimes \mathbf{k} \arrow[d, "\mathrm{id}_{A\otimes B}\otimes u_{A\otimes B}"'] & A\otimes B \arrow[l] \arrow[r] \arrow[d, "\mathrm{id}_{A\otimes B}"'] & \mathbf{k}\otimes A\otimes B \arrow[d, "u_{A\otimes B}\otimes \mathrm{id}_{A\otimes B}"] \\
A\otimes B\otimes A\otimes B \arrow[r, "m_{A\otimes B}"] & A\otimes B & A\otimes B\otimes A\otimes B \arrow[l, "m_{A\otimes B}"']
\end{tikzcd}$$

[384] commute (since Definition 1.1.1 yields that $A \otimes B$ is a $\mathbf{k}$-algebra if and only if these two diagrams commute).

We shall only prove that the diagram (13.2.1) commutes. The commutativity of the diagram (13.2.2) is proven similarly (but with less work), and so is left to the reader.

So we need to prove that the diagram (13.2.1) commutes. In other words, we need to prove that

$$m_{A\otimes B} \circ (m_{A\otimes B} \otimes \mathrm{id}_{A\otimes B}) = m_{A\otimes B} \circ (\mathrm{id}_{A\otimes B} \circ m_{A\otimes B}).$$

We first notice that

$$m_{A\otimes B}(p \otimes q \otimes p' \otimes q') = pp' \otimes qq' \tag{13.2.3}$$

for any $p \in A$, $q \in B$, $p' \in A$ and $q' \in B$ [385].

[384] where the maps $A \otimes B \to A \otimes B \otimes \mathbf{k}$ and $A \otimes B \to \mathbf{k} \otimes A \otimes B$ are the isomorphisms sending each $a \in A \otimes B$ to $a \otimes 1$ and to $1 \otimes a$, respectively

[385] *Proof of (13.2.3):* Let $p \in A$, $q \in B$, $p' \in A$ and $q' \in B$. Then,

$$\begin{aligned}
&\underbrace{m_{A\otimes B}}_{=(m_A\otimes m_B)\circ(\mathrm{id}_A\otimes T\otimes \mathrm{id}_B)} (p\otimes q\otimes p'\otimes q') \\
&= ((m_A\otimes m_B)\circ(\mathrm{id}_A\otimes T\otimes \mathrm{id}_B))(p\otimes q\otimes p'\otimes q') \\
&= (m_A\otimes m_B)\left(\underbrace{(\mathrm{id}_A\otimes T\otimes \mathrm{id}_B)(p\otimes q\otimes p'\otimes q')}_{=\mathrm{id}_A(p)\otimes T(q\otimes p')\otimes \mathrm{id}_B(q')}\right) \\
&= (m_A\otimes m_B)\left(\underbrace{\mathrm{id}_A(p)}_{=p}\otimes \underbrace{T(q\otimes p')}_{\substack{=p'\otimes q \\ \text{(by the definition of } T)}} \otimes \underbrace{\mathrm{id}_B(q')}_{=q'}\right) \\
&= (m_A\otimes m_B)(p\otimes p'\otimes q\otimes q') = \underbrace{m_A(p\otimes p')}_{\substack{=pp' \\ \text{(since } m_A \text{ is the} \\ \text{multiplication map of } A)}} \otimes \underbrace{m_B(q\otimes q')}_{\substack{=qq' \\ \text{(since } m_B \text{ is the} \\ \text{multiplication map of } B)}} \\
&= pp'\otimes qq'.
\end{aligned}$$

Qed.

Let $x \in A \otimes B \otimes A \otimes B \otimes A \otimes B$. Thus, x (like any tensor in $A \otimes B \otimes A \otimes B \otimes A \otimes B$) must be a $\mathbf{k}$-linear combination of pure tensors.

We want to show the equality

$$(m_{A\otimes B} \circ (m_{A\otimes B} \otimes \operatorname{id}_{A\otimes B}))(x) = (m_{A\otimes B} \circ (\operatorname{id}_{A\otimes B} \circ m_{A\otimes B}))(x). \tag{13.2.4}$$

This equality is $\mathbf{k}$-linear in x (since all maps that appear in it are $\mathbf{k}$-linear). Hence, we can WLOG assume that x is a pure tensor (since x is a $\mathbf{k}$-linear combination of pure tensors). Assume this. Thus, $x = a \otimes b \otimes a' \otimes b' \otimes a'' \otimes b''$ for some $a \in A$, $b \in B$, $a' \in A$, $b' \in B$, $a'' \in A$ and $b'' \in B$. Consider these a, b, a', b', a'' and b''.

Now,

$$\begin{aligned}
(m_{A\otimes B} \circ (m_{A\otimes B} \otimes \operatorname{id}_{A\otimes B}))(x) &= m_{A\otimes B}\left((m_{A\otimes B} \otimes \operatorname{id}_{A\otimes B})\left(\underbrace{x}_{=a\otimes b\otimes a'\otimes b'\otimes a''\otimes b''}\right)\right) \\
&= m_{A\otimes B}\left(\underbrace{(m_{A\otimes B} \otimes \operatorname{id}_{A\otimes B})(a \otimes b \otimes a' \otimes b' \otimes a'' \otimes b'')}_{=m_{A\otimes B}(a\otimes b\otimes a'\otimes b')\otimes \operatorname{id}_{A\otimes B}(a''\otimes b'')}\right) \\
&= m_{A\otimes B}\left(\underbrace{m_{A\otimes B}(a \otimes b \otimes a' \otimes b')}_{\substack{=aa'\otimes bb' \\ \text{(by (13.2.3))}}} \otimes \underbrace{\operatorname{id}_{A\otimes B}(a'' \otimes b'')}_{=a''\otimes b''}\right) \\
&= m_{A\otimes B}(aa' \otimes bb' \otimes a'' \otimes b'') = \underbrace{(aa')\,a''}_{=aa'a''} \otimes \underbrace{(bb')\,b''}_{=bb'b''} \qquad \text{(by (13.2.3))} \\
&= aa'a'' \otimes bb'b''.
\end{aligned}$$

A similar computation shows that

$$(m_{A\otimes B} \circ (\operatorname{id}_{A\otimes B} \otimes m_{A\otimes B}))(x) = aa'a'' \otimes bb'b''.$$

Comparing these two equalities, we obtain $(m_{A\otimes B} \circ (m_{A\otimes B} \otimes \operatorname{id}_{A\otimes B}))(x) = (m_{A\otimes B} \circ (\operatorname{id}_{A\otimes B} \otimes m_{A\otimes B}))(x)$. Thus, the equality (13.2.4) is proven.

Now, forget that we fixed x. We thus have proven the equality (13.2.4) for every $x \in A\otimes B\otimes A\otimes B\otimes A\otimes B$. In other words, we have $m_{A\otimes B} \circ (m_{A\otimes B} \otimes \operatorname{id}_{A\otimes B}) = m_{A\otimes B} \circ (\operatorname{id}_{A\otimes B} \circ m_{A\otimes B})$. In other words, the diagram (13.2.1) commutes.

It remains to prove that the diagram (13.2.2) commutes. We leave this to the reader, as the proof is similar to (but simpler than) the proof of the commutativity of (13.2.1) given above. Hence, both diagrams (13.2.1) and (13.2.2) commute. In other words, the $\mathbf{k}$-module $A \otimes B$, equipped with the two maps $m_{A\otimes B}$ and $u_{A\otimes B}$, is a $\mathbf{k}$-algebra. In other words, the $\mathbf{k}$-algebra $A \otimes B$ introduced in Definition 1.3.3 is actually well-defined. This solves Exercise 1.3.4(a).

Second solution to Exercise 1.3.4(a). For any two $\mathbf{k}$-modules U and V, let $T_{U,V} : U \otimes V \to V \otimes U$ be the twist map (i.e., the $\mathbf{k}$-linear map $U \otimes V \to V \otimes U$ sending every $u \otimes v$ to $v \otimes u$). A simple linear-algebraic fact says that if U, V, U' and V' are four $\mathbf{k}$-modules and $x : U \to U'$ and $y : V \to V'$ are two $\mathbf{k}$-linear maps, then

$$(y \otimes x) \circ T_{U,V} = T_{U',V'} \circ (x \otimes y). \tag{13.2.5}$$

For every $\mathbf{k}$-module U, let $\operatorname{kan}_{1,U} : U \to U \otimes \mathbf{k}$ and $\operatorname{kan}_{2,U} : U \to \mathbf{k} \otimes U$ be the canonical $\mathbf{k}$-module isomorphisms. Every $\mathbf{k}$-modules U and V satisfy the identities

$$\operatorname{id}_U \otimes \operatorname{kan}_{1,V} = \operatorname{kan}_{1,U\otimes V}, \tag{13.2.6}$$

$$\operatorname{kan}_{2,V} \otimes \operatorname{id}_U = \operatorname{kan}_{2,V\otimes U}, \tag{13.2.7}$$

$$\operatorname{id}_U \otimes \operatorname{kan}_{1,V}^{-1} = \operatorname{kan}_{1,U\otimes V}^{-1}, \tag{13.2.8}$$

$$\operatorname{kan}_{2,V}^{-1} \otimes \operatorname{id}_U = \operatorname{kan}_{2,V\otimes U}^{-1}. \tag{13.2.9}$$

(These identities are well-known and straightforward to check.)

Recall that the $\mathbf{k}$-module A, equipped with the maps m_A and u_A, is a $\mathbf{k}$-algebra. In other words, the two diagrams

(13.2.10)
$$\begin{tikzcd}
& A \otimes A \otimes A \arrow[dl, "m_A \otimes \mathrm{id}_A"'] \arrow[dr, "\mathrm{id}_A \otimes m_A"] & \\
A \otimes A \arrow[dr, "m_A"] & & A \otimes A \arrow[dl, "m_A"'] \\
& A &
\end{tikzcd}$$

and

(13.2.11)
$$\begin{tikzcd}
A \otimes \mathbf{k} \arrow[d, "\mathrm{id}_A \otimes u_A"'] & A \arrow[l, "\mathrm{kan}_{1,A}"'] \arrow[r, "\mathrm{kan}_{2,A}"] \arrow[d, "\mathrm{id}_A"'] & \mathbf{k} \otimes A \arrow[d, "u_A \otimes \mathrm{id}_A"] \\
A \otimes A \arrow[r, "m_A"] & A & A \otimes A \arrow[l, "m_A"']
\end{tikzcd}$$

commute (since Definition 1.1.1 yields that A is a $\mathbf{k}$-algebra if and only if these two diagrams commute).

The diagram (13.2.10) commutes. In other words, we have

$$m_A \circ (m_A \otimes \mathrm{id}_A) = m_A \circ (\mathrm{id}_A \otimes m_A) . \tag{13.2.12}$$

The same argument (applied to B instead of A) shows that

$$m_B \circ (m_B \otimes \mathrm{id}_B) = m_B \circ (\mathrm{id}_B \otimes m_B) . \tag{13.2.13}$$

The diagram (13.2.11) commutes. In other words, we have

$$\mathrm{id}_A = m_A \circ (\mathrm{id}_A \otimes u_A) \circ \mathrm{kan}_{1,A} \qquad \text{and} \tag{13.2.14}$$

$$\mathrm{id}_A = m_A \circ (u_A \otimes \mathrm{id}_A) \circ \mathrm{kan}_{2,A} . \tag{13.2.15}$$

The same argument (applied to B instead of A) shows that

$$\mathrm{id}_B = m_B \circ (\mathrm{id}_B \otimes u_B) \circ \mathrm{kan}_{1,B} \qquad \text{and} \tag{13.2.16}$$

$$\mathrm{id}_B = m_B \circ (u_B \otimes \mathrm{id}_B) \circ \mathrm{kan}_{2,B} . \tag{13.2.17}$$

Let ξ be the canonical $\mathbf{k}$-module isomorphism $\mathbf{k} \to \mathbf{k} \otimes \mathbf{k}$. According to Definition 1.3.3, we define the map $m_{A\otimes B} : A \otimes B \otimes A \otimes B \to A \otimes B$ by

$$m_{A\otimes B} = (m_A \otimes m_B) \circ (\mathrm{id}_A \otimes T_{B,A} \otimes \mathrm{id}_B) ,$$

and we define the map $u_{A\otimes B} : \mathbf{k} \to A \otimes B$ by

$$u_{A\otimes B} = (u_A \otimes u_B) \circ \xi.$$

We now must prove that the $\mathbf{k}$-module $A \otimes B$, equipped with these two maps $m_{A\otimes B}$ and $u_{A\otimes B}$, is indeed a $\mathbf{k}$-algebra. In other words, we must show that the two diagrams

(13.2.18)
$$\begin{tikzcd}
& A \otimes B \otimes A \otimes B \otimes A \otimes B \arrow[dl, "m_{A\otimes B} \otimes \mathrm{id}_{A\otimes B}"'] \arrow[dr, "\mathrm{id}_{A\otimes B} \otimes m_{A\otimes B}"] & \\
A \otimes B \otimes A \otimes B \arrow[dr, "m_{A\otimes B}"] & & A \otimes B \otimes A \otimes B \arrow[dl, "m_{A\otimes B}"'] \\
& A \otimes B &
\end{tikzcd}$$

and

(13.2.19)
$$\begin{tikzcd}
A \otimes B \otimes \mathbf{k} \arrow[d, "\mathrm{id}_{A\otimes B} \otimes u_{A\otimes B}"'] & A \otimes B \arrow[l, "\mathrm{kan}_{1,A\otimes B}"'] \arrow[r, "\mathrm{kan}_{2,A,\otimes B}"] \arrow[d, "\mathrm{id}_{A\otimes B}"'] & \mathbf{k} \otimes A \otimes B \arrow[d, "u_{A\otimes B} \otimes \mathrm{id}_{A\otimes B}"] \\
A \otimes B \otimes A \otimes B \arrow[r, "m_{A\otimes B}"] & A \otimes B & A \otimes B \otimes A \otimes B \arrow[l, "m_{A\otimes B}"']
\end{tikzcd}$$

commute (since Definition 1.1.1 yields that $A \otimes B$ is a $\mathbf{k}$-algebra if and only if these two diagrams commute).

Let us first prove that the diagram (13.2.18) commutes. In other words, let us prove that

$$m_{A\otimes B}\circ\left(m_{A\otimes B}\otimes\operatorname{id}_{A\otimes B}\right)=m_{A\otimes B}\circ\left(\operatorname{id}_{A\otimes B}\otimes m_{A\otimes B}\right).$$

Define a $\mathbf{k}$-linear map $Q : B\otimes A\otimes B\otimes A\to A\otimes A\otimes B\otimes B$ by

$$Q=\left(\operatorname{id}_A\otimes T_{B,A}\otimes\operatorname{id}_B\right)\circ\left(T_{B,A}\otimes T_{B,A}\right). \tag{13.2.20}$$

It is easy to see that

$$Q=\left(\operatorname{id}_A\otimes T_{B\otimes B,A}\right)\circ\left(T_{B,A}\otimes\operatorname{id}_B\otimes\operatorname{id}_A\right). \tag{13.2.21}$$

[*Proof of (13.2.21):* Let $b\in B$, $a\in A$, $b'\in B$ and $a'\in A$ be arbitrary. Applying both sides of the equality (13.2.20) to $b\otimes a\otimes b'\otimes a'$, we obtain

$$\begin{aligned}
Q\left(b\otimes a\otimes b'\otimes a'\right)&=\left(\left(\operatorname{id}_A\otimes T_{B,A}\otimes\operatorname{id}_B\right)\circ\left(T_{B,A}\otimes T_{B,A}\right)\right)\left(b\otimes a\otimes b'\otimes a'\right)\\
&=\left(\operatorname{id}_A\otimes T_{B,A}\otimes\operatorname{id}_B\right)\left(\underbrace{\left(T_{B,A}\otimes T_{B,A}\right)\left(b\otimes a\otimes b'\otimes a'\right)}_{=T_{B,A}(b\otimes a)\otimes T_{B,A}(b'\otimes a')}\right)\\
&=\left(\operatorname{id}_A\otimes T_{B,A}\otimes\operatorname{id}_B\right)\left(\underbrace{T_{B,A}\left(b\otimes a\right)}_{\substack{=a\otimes b\\ \text{(by the definition of }T_{B,A})}}\otimes\underbrace{T_{B,A}\left(b'\otimes a'\right)}_{\substack{=a'\otimes b'\\ \text{(by the definition of }T_{B,A})}}\right)\\
&=\left(\operatorname{id}_A\otimes T_{B,A}\otimes\operatorname{id}_B\right)\left(a\otimes b\otimes a'\otimes b'\right)=\underbrace{\operatorname{id}_A\left(a\right)}_{=a}\otimes\underbrace{T_{B,A}\left(b\otimes a'\right)}_{\substack{=a'\otimes b\\ \text{(by the definition of }T_{B,A})}}\otimes\underbrace{\operatorname{id}_B\left(b'\right)}_{=b'}\\
&=a\otimes a'\otimes b\otimes b'. \qquad (13.2.22)
\end{aligned}$$

Comparing this with

$$\begin{aligned}
&\left(\left(\operatorname{id}_A\otimes T_{B\otimes B,A}\right)\circ\left(T_{B,A}\otimes\operatorname{id}_B\otimes\operatorname{id}_A\right)\right)\left(b\otimes a\otimes b'\otimes a'\right)\\
&=\left(\operatorname{id}_A\otimes T_{B\otimes B,A}\right)\left(\underbrace{\left(T_{B,A}\otimes\operatorname{id}_B\otimes\operatorname{id}_A\right)\left(b\otimes a\otimes b'\otimes a'\right)}_{=T_{B,A}(b\otimes a)\otimes\operatorname{id}_B(b')\otimes\operatorname{id}_A(a')}\right)\\
&=\left(\operatorname{id}_A\otimes T_{B\otimes B,A}\right)\left(\underbrace{T_{B,A}\left(b\otimes a\right)}_{\substack{=a\otimes b\\ \text{(by the definition of }T_{B,A})}}\otimes\underbrace{\operatorname{id}_B\left(b'\right)}_{=b'}\otimes\underbrace{\operatorname{id}_A\left(a'\right)}_{=a'}\right)\\
&=\left(\operatorname{id}_A\otimes T_{B\otimes B,A}\right)\left(a\otimes b\otimes b'\otimes a'\right)=\underbrace{\operatorname{id}_A\left(a\right)}_{=a}\otimes\underbrace{T_{B\otimes B,A}\left(b\otimes b'\otimes a'\right)}_{\substack{=a'\otimes b\otimes b'\\ \text{(by the definition of }T_{B\otimes B,A})}}\\
&=a\otimes a'\otimes b\otimes b',
\end{aligned}$$

we obtain

$$Q\left(b\otimes a\otimes b'\otimes a'\right)=\left(\left(\operatorname{id}_A\otimes T_{B\otimes B,A}\right)\circ\left(T_{B,A}\otimes\operatorname{id}_B\otimes\operatorname{id}_A\right)\right)\left(b\otimes a\otimes b'\otimes a'\right). \tag{13.2.23}$$

Now, forget that we fixed b,a,b',a'. We thus have shown that every $b\in B$, $a\in A$, $b'\in B$ and $a'\in A$ satisfy (13.2.23). In other words, the two maps Q and $\left(\operatorname{id}_A\otimes T_{B\otimes B,A}\right)\circ\left(T_{B,A}\otimes\operatorname{id}_B\otimes\operatorname{id}_A\right)$ are equal to each other on each pure tensor. Since these two maps are $\mathbf{k}$-linear, we thus conclude that these two maps are identical (because if two $\mathbf{k}$-linear maps from a tensor product are equal to each other on each pure tensor, then these two maps are identical). In other words,

$$Q=\left(\operatorname{id}_A\otimes T_{B\otimes B,A}\right)\circ\left(T_{B,A}\otimes\operatorname{id}_B\otimes\operatorname{id}_A\right).$$

This proves (13.2.21).]

Now,

$$\begin{aligned}
&\underbrace{\mathrm{id}_A}_{=\mathrm{id}_A\circ\mathrm{id}_A} \otimes \underbrace{Q}_{\substack{=(\mathrm{id}_A\otimes T_{B\otimes B,A})\circ(T_{B,A}\otimes\mathrm{id}_B\otimes\mathrm{id}_A)\\ \text{(by (13.2.21))}}} \otimes \underbrace{\mathrm{id}_B}_{=\mathrm{id}_B\circ\mathrm{id}_B}\\
&= (\mathrm{id}_A\circ\mathrm{id}_A)\otimes((\mathrm{id}_A\otimes T_{B\otimes B,A})\circ(T_{B,A}\otimes\mathrm{id}_B\otimes\mathrm{id}_A))\otimes(\mathrm{id}_B\circ\mathrm{id}_B)\\
&= \underbrace{(\mathrm{id}_A\otimes(\mathrm{id}_A\otimes T_{B\otimes B,A})\otimes\mathrm{id}_B)}_{=\mathrm{id}_A\otimes\mathrm{id}_A\otimes T_{B\otimes B,A}\otimes\mathrm{id}_B}\circ\underbrace{(\mathrm{id}_A\otimes(T_{B,A}\otimes\mathrm{id}_B\otimes\mathrm{id}_A)\otimes\mathrm{id}_B)}_{=\mathrm{id}_A\otimes T_{B,A}\otimes\mathrm{id}_B\otimes\mathrm{id}_A\otimes\mathrm{id}_B}\\
&= (\mathrm{id}_A\otimes\mathrm{id}_A\otimes T_{B\otimes B,A}\otimes\mathrm{id}_B)\circ(\mathrm{id}_A\otimes T_{B,A}\otimes\mathrm{id}_B\otimes\mathrm{id}_A\otimes\mathrm{id}_B)\,.
\end{aligned}\tag{13.2.24}$$

Now, we shall show that

$$(\mathrm{id}_A\otimes T_{B,A}\otimes\mathrm{id}_B)\circ(m_{A\otimes B}\otimes\mathrm{id}_{A\otimes B})=(m_A\otimes\mathrm{id}_A\otimes m_B\otimes\mathrm{id}_B)\circ(\mathrm{id}_A\otimes Q\otimes\mathrm{id}_B)\tag{13.2.25}$$

(as maps from $A\otimes B\otimes A\otimes B\otimes A\otimes B$ to $A\otimes A\otimes B\otimes B$).

[*Proof of (13.2.25):* We have

$$\begin{aligned}
&\underbrace{m_{A\otimes B}}_{=(m_A\otimes m_B)\circ(\mathrm{id}_A\otimes T_{B,A}\otimes\mathrm{id}_B)} \otimes \underbrace{\mathrm{id}_{A\otimes B}}_{=\mathrm{id}_{A\otimes B}\circ\mathrm{id}_{A\otimes B}}\\
&= ((m_A\otimes m_B)\circ(\mathrm{id}_A\otimes T_{B,A}\otimes\mathrm{id}_B))\otimes(\mathrm{id}_{A\otimes B}\circ\mathrm{id}_{A\otimes B})\\
&= \underbrace{((m_A\otimes m_B)\otimes\mathrm{id}_{A\otimes B})}_{=m_A\otimes m_B\otimes\mathrm{id}_{A\otimes B}}\circ\underbrace{((\mathrm{id}_A\otimes T_{B,A}\otimes\mathrm{id}_B)\otimes\mathrm{id}_{A\otimes B})}_{=\mathrm{id}_A\otimes T_{B,A}\otimes\mathrm{id}_B\otimes\mathrm{id}_{A\otimes B}}\\
&= (m_A\otimes m_B\otimes\mathrm{id}_{A\otimes B})\circ(\mathrm{id}_A\otimes T_{B,A}\otimes\mathrm{id}_B\otimes\mathrm{id}_{A\otimes B})\,.
\end{aligned}\tag{13.2.26}$$

The equality (13.2.5) (applied to $B\otimes B$, A, B, A, m_B and id_A instead of U, V, U', V', x and y) yields

$$(\mathrm{id}_A\otimes m_B)\circ T_{B\otimes B,A}=T_{B,A}\circ(m_B\otimes\mathrm{id}_A)\,.$$

Hence,

$$\begin{aligned}
&(\mathrm{id}_A\circ m_A)\otimes\underbrace{((\mathrm{id}_A\otimes m_B)\circ T_{B\otimes B,A})}_{=T_{B,A}\circ(m_B\otimes\mathrm{id}_A)}\otimes(\mathrm{id}_B\circ\mathrm{id}_B)\\
&= (\mathrm{id}_A\circ m_A)\otimes(T_{B,A}\circ(m_B\otimes\mathrm{id}_A))\otimes(\mathrm{id}_B\circ\mathrm{id}_B)\\
&= (\mathrm{id}_A\otimes T_{B,A}\otimes\mathrm{id}_B)\circ\underbrace{(m_A\otimes(m_B\otimes\mathrm{id}_A)\otimes\mathrm{id}_B)}_{\substack{=m_A\otimes m_B\otimes\mathrm{id}_A\otimes\mathrm{id}_B=m_A\otimes m_B\otimes\mathrm{id}_{A\otimes B}\\ \text{(since }\mathrm{id}_A\otimes\mathrm{id}_B=\mathrm{id}_A\otimes B\text{)}}}\\
&= (\mathrm{id}_A\otimes T_{B,A}\otimes\mathrm{id}_B)\circ(m_A\otimes m_B\otimes\mathrm{id}_{A\otimes B})\,.
\end{aligned}$$

Thus,

$$\begin{aligned}
&(\mathrm{id}_A\otimes T_{B,A}\otimes\mathrm{id}_B)\circ(m_A\otimes m_B\otimes\mathrm{id}_{A\otimes B})\\
&= \underbrace{(\mathrm{id}_A\circ m_A)}_{=m_A=m_A\circ\mathrm{id}_{A\otimes A}}\otimes((\mathrm{id}_A\otimes m_B)\circ T_{B\otimes B,A})\otimes(\mathrm{id}_B\circ\mathrm{id}_B)\\
&= (m_A\circ\mathrm{id}_{A\otimes A})\otimes((\mathrm{id}_A\otimes m_B)\circ T_{B\otimes B,A})\otimes(\mathrm{id}_B\circ\mathrm{id}_B)\\
&= \underbrace{(m_A\otimes(\mathrm{id}_A\otimes m_B)\otimes\mathrm{id}_B)}_{=m_A\otimes\mathrm{id}_A\otimes m_B\otimes\mathrm{id}_B}\circ\left(\underbrace{\mathrm{id}_{A\otimes A}}_{=\mathrm{id}_A\otimes\mathrm{id}_A}\otimes T_{B\otimes B,A}\otimes\mathrm{id}_B\right)\\
&= (m_A\otimes\mathrm{id}_A\otimes m_B\otimes\mathrm{id}_B)\circ(\mathrm{id}_A\otimes\mathrm{id}_A\otimes T_{B\otimes B,A}\otimes\mathrm{id}_B)\,.
\end{aligned}\tag{13.2.27}$$

Now,

$$\begin{aligned}
&(\operatorname{id}_A \otimes T_{B,A} \otimes \operatorname{id}_B) \circ \underbrace{(m_{A\otimes B} \otimes \operatorname{id}_{A\otimes B})}_{\substack{=(m_A\otimes m_B\otimes \operatorname{id}_{A\otimes B})\circ(\operatorname{id}_A \otimes T_{B,A}\otimes \operatorname{id}_B \otimes \operatorname{id}_{A\otimes B}) \\ (\text{by } (13.2.26))}} \\
&= \underbrace{(\operatorname{id}_A \otimes T_{B,A} \otimes \operatorname{id}_B) \circ (m_A \otimes m_B \otimes \operatorname{id}_{A\otimes B})}_{\substack{=(m_A\otimes \operatorname{id}_A \otimes m_B\otimes \operatorname{id}_B)\circ(\operatorname{id}_A \otimes \operatorname{id}_A \otimes T_{B\otimes B,A}\otimes \operatorname{id}_B) \\ (\text{by } (13.2.27))}} \circ \left(\operatorname{id}_A \otimes T_{B,A} \otimes \operatorname{id}_B \otimes \underbrace{\operatorname{id}_{A\otimes B}}_{=\operatorname{id}_A \otimes \operatorname{id}_B} \right) \\
&= (m_A \otimes \operatorname{id}_A \otimes m_B \otimes \operatorname{id}_B) \circ \underbrace{(\operatorname{id}_A \otimes \operatorname{id}_A \otimes T_{B\otimes B,A} \otimes \operatorname{id}_B) \circ (\operatorname{id}_A \otimes T_{B,A} \otimes \operatorname{id}_B \otimes \operatorname{id}_A \otimes \operatorname{id}_B)}_{\substack{=\operatorname{id}_A \otimes Q\otimes \operatorname{id}_B \\ (\text{by } (13.2.24))}} \\
&= (m_A \otimes \operatorname{id}_A \otimes m_B \otimes \operatorname{id}_B) \circ (\operatorname{id}_A \otimes Q \otimes \operatorname{id}_B) .
\end{aligned}$$

This proves (13.2.25).]

Now,

$$\begin{aligned}
&\underbrace{m_{A\otimes B}}_{=(m_A\otimes m_B)\circ(\operatorname{id}_A \otimes T_{B,A}\otimes \operatorname{id}_B)} \circ (m_{A\otimes B} \otimes \operatorname{id}_{A\otimes B}) \\
&= (m_A \otimes m_B) \circ \underbrace{(\operatorname{id}_A \otimes T_{B,A} \otimes \operatorname{id}_B) \circ (m_{A\otimes B} \otimes \operatorname{id}_{A\otimes B})}_{\substack{=(m_A\otimes \operatorname{id}_A \otimes m_B\otimes \operatorname{id}_B)\circ(\operatorname{id}_A \otimes Q\otimes \operatorname{id}_B) \\ (\text{by } (13.2.25))}} \\
&= (m_A \otimes m_B) \circ \underbrace{(m_A \otimes \operatorname{id}_A \otimes m_B \otimes \operatorname{id}_B)}_{=(m_A\otimes \operatorname{id}_A)\otimes(m_B\otimes \operatorname{id}_B)} \circ (\operatorname{id}_A \otimes Q \otimes \operatorname{id}_B) \\
&= \underbrace{(m_A \otimes m_B) \circ ((m_A \otimes \operatorname{id}_A) \otimes (m_B \otimes \operatorname{id}_B))}_{=(m_A\circ(m_A\otimes \operatorname{id}_A))\otimes(m_B\circ(m_B\otimes \operatorname{id}_B))} \circ (\operatorname{id}_A \otimes Q \otimes \operatorname{id}_B) \\
&= \left(\underbrace{(m_A \circ (m_A \otimes \operatorname{id}_A))}_{\substack{=m_A\circ(\operatorname{id}_A \otimes m_A) \\ (\text{by } (13.2.12))}} \otimes \underbrace{(m_B \circ (m_B \otimes \operatorname{id}_B))}_{\substack{=m_B\circ(\operatorname{id}_B \otimes m_B) \\ (\text{by } (13.2.13))}} \right) \circ (\operatorname{id}_A \otimes Q \otimes \operatorname{id}_B) \\
&= \underbrace{((m_A \circ (\operatorname{id}_A \otimes m_A)) \otimes (m_B \circ (\operatorname{id}_B \otimes m_B)))}_{=(m_A\otimes m_B)\circ((\operatorname{id}_A \otimes m_A)\otimes(\operatorname{id}_B \otimes m_B))} \circ (\operatorname{id}_A \otimes Q \otimes \operatorname{id}_B) \\
&= (m_A \otimes m_B) \circ \underbrace{((\operatorname{id}_A \otimes m_A) \otimes (\operatorname{id}_B \otimes m_B))}_{=\operatorname{id}_A \otimes m_A\otimes \operatorname{id}_B \otimes m_B} \circ (\operatorname{id}_A \otimes Q \otimes \operatorname{id}_B) \\
&= (m_A \otimes m_B) \circ (\operatorname{id}_A \otimes m_A \otimes \operatorname{id}_B \otimes m_B) \circ (\operatorname{id}_A \otimes Q \otimes \operatorname{id}_B) . \qquad (13.2.28)
\end{aligned}$$

On the other hand, it is easy to see that

$$Q = (T_{B,A\otimes A} \otimes \operatorname{id}_B) \circ (\operatorname{id}_B \otimes \operatorname{id}_A \otimes T_{B,A}) . \tag{13.2.29}$$

[*Proof of (13.2.29):* This proof is similar to the proof of (13.2.21), so we omit it.]

Now,

$$\begin{aligned}
&\underbrace{\operatorname{id}_A}_{=\operatorname{id}_A \circ \operatorname{id}_A} \otimes \underbrace{Q}_{\substack{=(T_{B,A\otimes A}\otimes \operatorname{id}_B)\circ(\operatorname{id}_B \otimes \operatorname{id}_A \otimes T_{B,A}) \\ (\text{by } (13.2.29))}} \otimes \underbrace{\operatorname{id}_B}_{=\operatorname{id}_B \circ \operatorname{id}_B} \\
&= (\operatorname{id}_A \circ \operatorname{id}_A) \otimes ((T_{B,A\otimes A} \otimes \operatorname{id}_B) \circ (\operatorname{id}_B \otimes \operatorname{id}_A \otimes T_{B,A})) \otimes (\operatorname{id}_B \circ \operatorname{id}_B) \\
&= \underbrace{(\operatorname{id}_A \otimes (T_{B,A\otimes A} \otimes \operatorname{id}_B) \otimes \operatorname{id}_B)}_{=\operatorname{id}_A \otimes T_{B,A\otimes A}\otimes \operatorname{id}_B \otimes \operatorname{id}_B} \circ \underbrace{(\operatorname{id}_A \otimes (\operatorname{id}_B \otimes \operatorname{id}_A \otimes T_{B,A}) \otimes \operatorname{id}_B)}_{=\operatorname{id}_A \otimes \operatorname{id}_B \otimes \operatorname{id}_A \otimes T_{B,A}\otimes \operatorname{id}_B} \\
&= (\operatorname{id}_A \otimes T_{B,A\otimes A} \otimes \operatorname{id}_B \otimes \operatorname{id}_B) \circ (\operatorname{id}_A \otimes \operatorname{id}_B \otimes \operatorname{id}_A \otimes T_{B,A} \otimes \operatorname{id}_B) . \qquad (13.2.30)
\end{aligned}$$

Now, we shall show that

$$(\mathrm{id}_A \otimes T_{B,A} \otimes \mathrm{id}_B) \circ (\mathrm{id}_{A\otimes B} \otimes m_{A,B}) = (\mathrm{id}_A \otimes m_A \otimes \mathrm{id}_B \otimes m_B) \circ (\mathrm{id}_A \otimes Q \otimes \mathrm{id}_B) \tag{13.2.31}$$

(as maps from $A \otimes B \otimes A \otimes B \otimes A \otimes B$ to $A \otimes A \otimes B \otimes B$).

[*Proof of (13.2.31):* We have

$$\begin{aligned}
&\underbrace{\mathrm{id}_{A\otimes B}}_{=\mathrm{id}_{A\otimes B} \circ \mathrm{id}_{A\otimes B}} \otimes \underbrace{m_{A\otimes B}}_{=(m_A \otimes m_B) \circ (\mathrm{id}_A \otimes T_{B,A} \otimes \mathrm{id}_B)} \\
&= (\mathrm{id}_{A\otimes B} \circ \mathrm{id}_{A\otimes B}) \otimes ((m_A \otimes m_B) \circ (\mathrm{id}_A \otimes T_{B,A} \otimes \mathrm{id}_B)) \\
&= \underbrace{(\mathrm{id}_{A\otimes B} \otimes (m_A \otimes m_B))}_{=\mathrm{id}_{A\otimes B} \otimes m_A \otimes m_B} \circ \underbrace{(\mathrm{id}_{A\otimes B} \otimes (\mathrm{id}_A \otimes T_{B,A} \otimes \mathrm{id}_B))}_{=\mathrm{id}_{A\otimes B} \otimes \mathrm{id}_A \otimes T_{B,A} \otimes \mathrm{id}_B}
\end{aligned}$$

$$= (\mathrm{id}_{A\otimes B} \otimes m_A \otimes m_B) \circ (\mathrm{id}_{A\otimes B} \otimes \mathrm{id}_A \otimes T_{B,A} \otimes \mathrm{id}_B) . \tag{13.2.32}$$

The equality (13.2.5) (applied to B, $A \otimes A$, B, A, id_B and m_A instead of U, V, U', V', x and y) yields

$$(m_A \otimes \mathrm{id}_B) \circ T_{B,A\otimes A} = T_{B,A} \circ (\mathrm{id}_B \otimes m_A) .$$

Hence,

$$\begin{aligned}
&(\mathrm{id}_A \circ \mathrm{id}_A) \otimes \underbrace{((m_A \otimes \mathrm{id}_B) \circ T_{B,A\otimes A})}_{=T_{B,A} \circ (\mathrm{id}_B \otimes m_A)} \otimes (\mathrm{id}_B \circ m_B) \\
&= (\mathrm{id}_A \circ \mathrm{id}_A) \otimes (T_{B,A} \circ (\mathrm{id}_B \otimes m_A)) \otimes (\mathrm{id}_B \circ m_B) \\
&= (\mathrm{id}_A \otimes T_{B,A} \otimes \mathrm{id}_B) \circ \underbrace{(\mathrm{id}_A \otimes (\mathrm{id}_B \otimes m_A) \otimes m_B)}_{\substack{=\mathrm{id}_A \otimes \mathrm{id}_B \otimes m_A \otimes m_B = \mathrm{id}_{A\otimes B} \otimes m_A \otimes m_B \\ (\text{since } \mathrm{id}_A \otimes \mathrm{id}_B = \mathrm{id}_{A\otimes B})}} \\
&= (\mathrm{id}_A \otimes T_{B,A} \otimes \mathrm{id}_B) \circ (\mathrm{id}_{A\otimes B} \otimes m_A \otimes m_B) .
\end{aligned}$$

Thus,

$$\begin{aligned}
&(\mathrm{id}_A \otimes T_{B,A} \otimes \mathrm{id}_B) \circ (\mathrm{id}_{A\otimes B} \otimes m_A \otimes m_B) \\
&= (\mathrm{id}_A \circ \mathrm{id}_A) \otimes ((m_A \otimes \mathrm{id}_B) \circ T_{B,A\otimes A}) \otimes \underbrace{(\mathrm{id}_B \circ m_B)}_{=m_B = m_B \circ \mathrm{id}_{B\otimes B}} \\
&= (\mathrm{id}_A \circ \mathrm{id}_A) \otimes ((m_A \otimes \mathrm{id}_B) \circ T_{B,A\otimes A}) \otimes (m_B \circ \mathrm{id}_{B\otimes B}) \\
&= \underbrace{(\mathrm{id}_A \otimes (m_A \otimes \mathrm{id}_B) \otimes m_B)}_{=\mathrm{id}_A \otimes m_A \otimes \mathrm{id}_B \otimes m_B} \circ \left(\mathrm{id}_A \otimes T_{B,A\otimes A} \otimes \underbrace{\mathrm{id}_{B\otimes B}}_{=\mathrm{id}_B \otimes \mathrm{id}_B} \right)
\end{aligned}$$

$$= (\mathrm{id}_A \otimes m_A \otimes \mathrm{id}_B \otimes m_B) \circ (\mathrm{id}_A \otimes T_{B,A\otimes A} \otimes \mathrm{id}_B \otimes \mathrm{id}_B) . \tag{13.2.33}$$

Now,

$$\begin{aligned}
&(\mathrm{id}_A \otimes T_{B,A} \otimes \mathrm{id}_B) \circ \underbrace{(\mathrm{id}_{A\otimes B} \otimes m_{A\otimes B})}_{\substack{=(\mathrm{id}_{A\otimes B} \otimes m_A \otimes m_B) \circ (\mathrm{id}_{A\otimes B} \otimes \mathrm{id}_A \otimes T_{B,A} \otimes \mathrm{id}_B) \\ (\text{by } (13.2.32))}} \\
&= \underbrace{(\mathrm{id}_A \otimes T_{B,A} \otimes \mathrm{id}_B) \circ (\mathrm{id}_{A\otimes B} \otimes m_A \otimes m_B)}_{\substack{=(\mathrm{id}_A \otimes m_A \otimes \mathrm{id}_B \otimes m_B) \circ (\mathrm{id}_A \otimes T_{B,A\otimes A} \otimes \mathrm{id}_B \otimes \mathrm{id}_B) \\ (\text{by } (13.2.33))}} \circ \left(\underbrace{\mathrm{id}_{A\otimes B}}_{=\mathrm{id}_A \otimes \mathrm{id}_B} \otimes \mathrm{id}_A \otimes T_{B,A} \otimes \mathrm{id}_B \right) \\
&= (\mathrm{id}_A \otimes m_A \otimes \mathrm{id}_B \otimes m_B) \circ \underbrace{(\mathrm{id}_A \otimes T_{B,A\otimes A} \otimes \mathrm{id}_B \otimes \mathrm{id}_B) \circ (\mathrm{id}_A \otimes \mathrm{id}_B \otimes \mathrm{id}_A \otimes T_{B,A} \otimes \mathrm{id}_B)}_{\substack{=\mathrm{id}_A \otimes Q \otimes \mathrm{id}_B \\ (\text{by } (13.2.30))}} \\
&= (\mathrm{id}_A \otimes m_A \otimes \mathrm{id}_B \otimes m_B) \circ (\mathrm{id}_A \otimes Q \otimes \mathrm{id}_B) .
\end{aligned}$$

This proves (13.2.31).]

Now,

$$\begin{aligned}&\underbrace{m_{A\otimes B}}_{=(m_A\otimes m_B)\circ(\operatorname{id}_A\otimes T_{B,A}\otimes\operatorname{id}_B)}\circ(\operatorname{id}_{A\otimes B}\otimes m_{A\otimes B})\\&=(m_A\otimes m_B)\circ\underbrace{(\operatorname{id}_A\otimes T_{B,A}\otimes\operatorname{id}_B)\circ(\operatorname{id}_{A\otimes B}\otimes m_{A\otimes B})}_{\substack{=(\operatorname{id}_A\otimes m_A\otimes\operatorname{id}_B\otimes m_B)\circ(\operatorname{id}_A\otimes Q\otimes\operatorname{id}_B)\\ \text{(by (13.2.31))}}}\\&=(m_A\otimes m_B)\circ(\operatorname{id}_A\otimes m_A\otimes\operatorname{id}_B\otimes m_B)\circ(\operatorname{id}_A\otimes Q\otimes\operatorname{id}_B).\end{aligned}$$

Comparing this with (13.2.28), we obtain

$$m_{A\otimes B}\circ(m_{A\otimes B}\otimes\operatorname{id}_{A\otimes B})=m_{A\otimes B}\circ(\operatorname{id}_{A\otimes B}\otimes m_{A\otimes B}).$$

In other words, the diagram (13.2.18) commutes.

Let us next prove that the diagram (13.2.19) commutes.

We first observe that

$$(\operatorname{id}_A\otimes T_{B,\mathbf{k}}\otimes\operatorname{id}_{\mathbf{k}})\circ(\operatorname{id}_{A\otimes B}\otimes\xi)\circ\operatorname{kan}_{1,A\otimes B}=\operatorname{kan}_{1,A}\otimes\operatorname{kan}_{1,B}. \tag{13.2.34}$$

[*Proof of (13.2.34):* This can be proven along the same lines as the proof of proof of (13.2.21): We fix $a\in A$ and $b\in B$, and apply both sides of (13.2.34) to $a\otimes b$, then check that the results are the same.]

The equality (13.2.6) (applied to $U=A$ and $V=B$) yields

$$\operatorname{id}_A\otimes\operatorname{kan}_{1,B}=\operatorname{kan}_{1,A\otimes B}. \tag{13.2.35}$$

From

$$\begin{aligned}&(\operatorname{id}_A\circ\operatorname{id}_A)\otimes\underbrace{((u_A\otimes\operatorname{id}_B)\circ T_{B,\mathbf{k}})}_{\substack{=T_{B,A}\circ(\operatorname{id}_B\otimes u_A)\\ \text{(by (13.2.5) (applied}\\ \text{to } B,\ \mathbf{k},\ B,\ A,\ \operatorname{id}_B \text{ and } u_A\\ \text{instead of } U,\ V,\ U',\ V',\ x \text{ and } y))}}\otimes\underbrace{(u_B\circ\operatorname{id}_{\mathbf{k}})}_{=u_B=\operatorname{id}_B\circ u_B}\\&=(\operatorname{id}_A\circ\operatorname{id}_A)\otimes(T_{B,A}\circ(\operatorname{id}_B\otimes u_A))\otimes(\operatorname{id}_B\circ u_B)=(\operatorname{id}_A\otimes T_{B,A}\otimes\operatorname{id}_B)\circ\underbrace{(\operatorname{id}_A\otimes(\operatorname{id}_B\otimes u_A)\otimes u_B)}_{=\operatorname{id}_A\otimes\operatorname{id}_B\otimes u_A\otimes u_B}\\&=(\operatorname{id}_A\otimes T_{B,A}\otimes\operatorname{id}_B)\circ\left(\underbrace{\operatorname{id}_A\otimes\operatorname{id}_B}_{=\operatorname{id}_{A\otimes B}}\otimes u_A\otimes u_B\right)=(\operatorname{id}_A\otimes T_{B,A}\otimes\operatorname{id}_B)\circ(\operatorname{id}_{A\otimes B}\otimes u_A\otimes u_B),\end{aligned}$$

we obtain

$$\begin{aligned}&(\operatorname{id}_A\otimes T_{B,A}\otimes\operatorname{id}_B)\circ(\operatorname{id}_{A\otimes B}\otimes u_A\otimes u_B)\\&=(\operatorname{id}_A\circ\operatorname{id}_A)\otimes((u_A\otimes\operatorname{id}_B)\circ T_{B,\mathbf{k}})\otimes(u_B\circ\operatorname{id}_{\mathbf{k}})\\&=\underbrace{(\operatorname{id}_A\otimes(u_A\otimes\operatorname{id}_B)\otimes u_B)}_{=\operatorname{id}_A\otimes u_A\otimes\operatorname{id}_B\otimes u_B}\circ(\operatorname{id}_A\otimes T_{B,\mathbf{k}}\otimes\operatorname{id}_{\mathbf{k}})\\&=(\operatorname{id}_A\otimes u_A\otimes\operatorname{id}_B\otimes u_B)\circ(\operatorname{id}_A\otimes T_{B,\mathbf{k}}\otimes\operatorname{id}_{\mathbf{k}}).\end{aligned} \tag{13.2.36}$$

Next, we observe that

$$\begin{aligned}&\underbrace{\operatorname{id}_{A\otimes B}}_{=\operatorname{id}_{A\otimes B}\circ\operatorname{id}_{A\otimes B}}\otimes\underbrace{u_{A\otimes B}}_{=(u_A\otimes u_B)\circ\xi}\\&=(\operatorname{id}_{A\otimes B}\circ\operatorname{id}_{A\otimes B})\otimes((u_A\otimes u_B)\circ\xi)=\underbrace{(\operatorname{id}_{A\otimes B}\otimes(u_A\otimes u_B))}_{=\operatorname{id}_{A\otimes B}\otimes u_A\otimes u_B}\circ(\operatorname{id}_{A\otimes B}\otimes\xi)\\&=(\operatorname{id}_{A\otimes B}\otimes u_A\otimes u_B)\circ(\operatorname{id}_{A\otimes B}\otimes\xi).\end{aligned}$$

Hence,

$$\begin{aligned}
&\underbrace{m_{A\otimes B}}_{=(m_A\otimes m_B)\circ(\mathrm{id}_A\otimes T_{B,A}\otimes\mathrm{id}_B)} \circ \underbrace{(\mathrm{id}_{A\otimes B}\otimes u_{A\otimes B})}_{=(\mathrm{id}_{A\otimes B}\otimes u_A\otimes u_B)\circ(\mathrm{id}_{A\otimes B}\otimes\xi)} \\
&= (m_A\otimes m_B)\circ \underbrace{(\mathrm{id}_A\otimes T_{B,A}\otimes\mathrm{id}_B)\circ(\mathrm{id}_{A\otimes B}\otimes u_A\otimes u_B)}_{\substack{=(\mathrm{id}_A\otimes u_A\otimes\mathrm{id}_B\otimes u_B)\circ(\mathrm{id}_A\otimes T_{B,\mathbf{k}}\otimes\mathrm{id}_{\mathbf{k}}) \\ \text{(by (13.2.36))}}} \circ (\mathrm{id}_{A\otimes B}\otimes\xi) \\
&= (m_A\otimes m_B)\circ(\mathrm{id}_A\otimes u_A\otimes\mathrm{id}_B\otimes u_B)\circ(\mathrm{id}_A\otimes T_{B,\mathbf{k}}\otimes\mathrm{id}_{\mathbf{k}})\circ(\mathrm{id}_{A\otimes B}\otimes\xi).
\end{aligned}$$

Hence,

$$\begin{aligned}
&\underbrace{m_{A\otimes B}\circ(\mathrm{id}_{A\otimes B}\otimes u_{A\otimes B})}_{=(m_A\otimes m_B)\circ(\mathrm{id}_A\otimes u_A\otimes\mathrm{id}_B\otimes u_B)\circ(\mathrm{id}_A\otimes T_{B,\mathbf{k}}\otimes\mathrm{id}_{\mathbf{k}})\circ(\mathrm{id}_{A\otimes B}\otimes\xi)} \circ \mathrm{kan}_{1,A\otimes B} \\
&= (m_A\otimes m_B)\circ \underbrace{(\mathrm{id}_A\otimes u_A\otimes\mathrm{id}_B\otimes u_B)}_{=(\mathrm{id}_A\otimes u_A)\otimes(\mathrm{id}_B\otimes u_B)} \circ \underbrace{(\mathrm{id}_A\otimes T_{B,\mathbf{k}}\otimes\mathrm{id}_{\mathbf{k}})\circ(\mathrm{id}_{A\otimes B}\otimes\xi)\circ\mathrm{kan}_{1,A\otimes B}}_{\substack{=\mathrm{kan}_{1,A}\otimes\mathrm{kan}_{1,B} \\ \text{(by (13.2.34))}}} \\
&= (m_A\otimes m_B)\circ((\mathrm{id}_A\otimes u_A)\otimes(\mathrm{id}_B\otimes u_B))\circ(\mathrm{kan}_{1,A}\otimes\mathrm{kan}_{1,B}).
\end{aligned}$$

Comparing this with

$$\begin{aligned}
\mathrm{id}_{A\otimes B} &= \underbrace{\mathrm{id}_A}_{\substack{=m_A\circ(\mathrm{id}_A\otimes u_A)\circ\mathrm{kan}_{1,A} \\ \text{(by (13.2.14))}}} \otimes \underbrace{\mathrm{id}_B}_{\substack{=m_B\circ(\mathrm{id}_B\otimes u_B)\circ\mathrm{kan}_{1,B} \\ \text{(by (13.2.16))}}} \\
&= (m_A\circ(\mathrm{id}_A\otimes u_A)\circ\mathrm{kan}_{1,A})\otimes(m_B\circ(\mathrm{id}_B\otimes u_B)\circ\mathrm{kan}_{1,B}) \\
&= (m_A\otimes m_B)\circ((\mathrm{id}_A\otimes u_A)\otimes(\mathrm{id}_B\otimes u_B))\circ(\mathrm{kan}_{1,A}\otimes\mathrm{kan}_{1,B}),
\end{aligned}$$

we obtain

$$m_{A\otimes B}\circ(\mathrm{id}_{A\otimes B}\otimes u_{A\otimes B})\circ\mathrm{kan}_{1,A\otimes B} = \mathrm{id}_{A\otimes B}.$$

In other words, the left rectangle of the diagram (13.2.19) commutes.

A similar argument shows that the right rectangle of the diagram (13.2.19) commutes[386]. Thus, the whole diagram (13.2.19) commutes.

We have now shown that the two diagrams (13.2.18) and (13.2.19) commute. Thus, the **k**-module $A\otimes B$, equipped with the two maps $m_{A\otimes B}$ and $u_{A\otimes B}$, is a **k**-algebra (since Definition 1.1.1 yields that $A\otimes B$ is a **k**-algebra if and only if the two diagrams (13.2.18) and (13.2.19) commute). In other words, the **k**-algebra $A\otimes B$ introduced in Definition 1.3.3 is actually well-defined. This solves Exercise 1.3.4(a) again.

(b) Exercise 1.3.4(b) is the "dual" statement to Exercise 1.3.4(a). We shall sketch two solutions to it: one solution by elementwise computation (similar to the first solution to Exercise 1.3.4(a), but somewhat more complicated due to the many sums involved), and another by formally manipulating homomorphisms. The second solution will be very brief, because we will not elaborate on it; we will merely explain how it can be obtained by "reversing arrows" from the second solution to Exercise 1.3.4(a).

First solution to Exercise 1.3.4(b). Let θ be the canonical **k**-module isomorphism $\mathbf{k}\otimes\mathbf{k}\to\mathbf{k}$. We have defined the comultiplication map $\Delta_{C\otimes D}$ of the **k**-coalgebra $C\otimes D$ by

$$\Delta_{C\otimes D} = (\mathrm{id}_C\otimes T\otimes\mathrm{id}_D)\circ(\Delta_C\otimes\Delta_D), \qquad (13.2.38)$$

and the counit map $\epsilon_{C\otimes D}$ of the **k**-coalgebra $C\otimes D$ by

$$\epsilon_{C\otimes D} = \theta\circ(\epsilon_C\otimes\epsilon_D).$$

[386] We leave the details of this argument to the reader. Let us just mention that it uses the following equality (analogous to the equality (13.2.34) used in the proof of the commutativity of the left diagram):

$$(\mathrm{id}_{\mathbf{k}}\otimes T_{\mathbf{k},A}\otimes\mathrm{id}_B)\circ(\xi\otimes\mathrm{id}_{A\otimes B})\circ\mathrm{kan}_{2,A\otimes B} = \mathrm{kan}_{2,A}\otimes\mathrm{kan}_{2,B}. \qquad (13.2.37)$$

We now must prove that the $\mathbf{k}$-module $C \otimes D$, equipped with these two maps $\Delta_{C\otimes D}$ and $\epsilon_{C\otimes D}$, is indeed a $\mathbf{k}$-coalgebra. In other words, we must show that the two diagrams

(13.2.39)
$$\begin{tikzcd}
 & C \otimes D \otimes C \otimes D \otimes C \otimes D & \\
C \otimes D \otimes C \otimes D \arrow[ur, "\Delta_{C\otimes D} \otimes \mathrm{id}_{C\otimes D}"] & & C \otimes D \otimes C \otimes D \arrow[ul, "\mathrm{id}_{C\otimes D} \otimes \Delta_{C\otimes D}"'] \\
 & C \otimes D \arrow[ul, "\Delta_{C\otimes D}"] \arrow[ur, "\Delta_{C\otimes D}"'] &
\end{tikzcd}$$

and

(13.2.40)
$$\begin{tikzcd}
C \otimes D \otimes \mathbf{k} \arrow[r] & C \otimes D & \mathbf{k} \otimes C \otimes D \arrow[l] \\
C \otimes D \otimes C \otimes D \arrow[u, "\mathrm{id}_{C\otimes D} \otimes \epsilon_{C\otimes D}"] & C \otimes D \arrow[l, "\Delta_{C\otimes D}"'] \arrow[r, "\Delta_{C\otimes D}"] \arrow[u, "\mathrm{id}_{C\otimes D}"] & C \otimes D \otimes C \otimes D \arrow[u, "\epsilon_{C\otimes D} \otimes \mathrm{id}_{C\otimes D}"']
\end{tikzcd}$$

[387] commute (since Definition 1.2.1 yields that $C \otimes D$ is a $\mathbf{k}$-coalgebra if and only if these two diagrams commute).

We shall only prove that the diagram (13.2.39) commutes. The commutativity of the diagram (13.2.40) is proven similarly (but with less work), and so is left to the reader.

So we need to prove that the diagram (13.2.39) commutes. In other words, we need to prove that

$$(\Delta_{C\otimes D} \otimes \mathrm{id}_{C\otimes D}) \circ \Delta_{C\otimes D} = (\mathrm{id}_{C\otimes D} \otimes \Delta_{C\otimes D}) \circ \Delta_{C\otimes D}.$$

We shall use the Sweedler notation, by writing $\sum_{(x)} x_1 \otimes x_2$ for $\Delta(x)$ whenever x is an element of a coalgebra. This is a neat opportunity to practice the use of the Sweedler notation. But if you are uncomfortable with the Sweedler notation, you can easily exorcise it from the following argument as follows:

- Whenever an element $e \in C$ is defined, fix a decomposition $\Delta_C(e) = \sum_{a=1}^{b} r_a \otimes s_a$ of $\Delta_C(e)$ into a sum of pure tensors.
- Whenever an element $f \in D$ is defined, fix a decomposition $\Delta_D(f) = \sum_{a'=1}^{b'} r'_{a'} \otimes s'_{a'}$ of $\Delta_D(f)$ into a sum of pure tensors.
- Whenever an element $c \in C$ is defined, fix a decomposition $\Delta_C(c) = \sum_{i=1}^{n} p_i \otimes q_i$ of $\Delta_C(c)$ into a sum of pure tensors, and furthermore:
  - For each $i \in \{1, 2, \ldots, n\}$, fix a decomposition $\Delta_C(p_i) = \sum_{j=1}^{k_i} p'_{i,j} \otimes p''_{i,j}$ of $\Delta_C(p_i)$ into a sum of pure tensors.
  - For each $i \in \{1, 2, \ldots, n\}$, fix a decomposition $\Delta_C(q_i) = \sum_{h=1}^{\ell_i} q'_{i,h} \otimes q''_{i,h}$ of $\Delta_C(q_i)$ into a sum of pure tensors.
- Whenever an element $d \in D$ is defined, fix a decomposition $\Delta_D(d) = \sum_{i'=1}^{n'} x_{i'} \otimes y_{i'}$ of $\Delta_D(d)$ into a sum of pure tensors, and furthermore:
  - For each $i' \in \{1, 2, \ldots, n'\}$, fix a decomposition $\Delta_D(x_{i'}) = \sum_{j'=1}^{k'_{i'}} x'_{i',j'} \otimes x''_{i',j'}$ of $\Delta_D(x_{i'})$ into a sum of pure tensors.
  - For each $i' \in \{1, 2, \ldots, n'\}$, fix a decomposition $\Delta_D(y_{i'}) = \sum_{h'=1}^{\ell'_{i'}} y'_{i',h'} \otimes y''_{i',h'}$ of $\Delta_D(y_{i'})$ into a sum of pure tensors.

Once these decompositions are chosen, it remains to replace each appearance of one of the symbols

$$\begin{aligned}
&\sum_{(e)}, \quad e_1, \quad e_2, \quad \sum_{(f)}, \quad f_1, \quad f_2, \\
&\sum_{(c)}, \quad c_1, \quad c_2, \quad \sum_{(c_1)}, \quad (c_1)_1, \quad (c_1)_2, \quad \sum_{(c_2)}, \quad (c_2)_1, \quad (c_2)_2, \\
&\sum_{(d)}, \quad d_1, \quad d_2, \quad \sum_{(d_1)}, \quad (d_1)_1, \quad (d_1)_2, \quad \sum_{(d_2)}, \quad (d_2)_1, \quad (d_2)_2
\end{aligned}$$

[387] where the maps $C \otimes D \otimes \mathbf{k} \to C \otimes D$ and $\mathbf{k} \otimes C \otimes D \to C \otimes D$ are the isomorphisms sending each $a \otimes \lambda \in C \otimes D \otimes \mathbf{k}$ with $a \in C \otimes D$ and $\lambda \in \mathbf{k}$ (resp., each $\lambda \otimes a \in \mathbf{k} \otimes C \otimes D$ with $\lambda \in \mathbf{k}$ and $a \in C \otimes D$) to λa.

by the symbol

$$\begin{array}{llllllllll}
\sum_{a=1}^{b}, & r_a, & s_a, & \sum_{a'=1}^{b'}, & r'_{a'}, & s'_{a'}, & & & \\
\sum_{i=1}^{n}, & p_i, & q_i, & \sum_{j=1}^{k_i}, & p'_{i,j}, & p''_{i,j}, & \sum_{h=1}^{\ell_i}, & q'_{i,h}, & q''_{i,h}, \\
\sum_{i'=1}^{n'}, & x_{i'}, & y_{i'}, & \sum_{j'=1}^{k'_{i'}}, & x'_{i',j'}, & x''_{i',j'}, & \sum_{h'=1}^{\ell'_{i'}}, & y'_{i',h'}, & y''_{i',h'},
\end{array}$$

respectively[388]. For example, these replacements transform the expression

$$\sum_{(c)} \sum_{(c_2)} \sum_{(d)} \sum_{(d_2)} c_1 \otimes d_1 \otimes (c_2)_1 \otimes (d_2)_1 \otimes (c_2)_2 \otimes (d_2)_2$$

into

$$\sum_{i=1}^{n} \sum_{h=1}^{\ell_i} \sum_{i'=1}^{n'} \sum_{h'=1}^{\ell'_{i'}} p_i \otimes x_{i'} \otimes q'_{i,h} \otimes y'_{i',h'} \otimes q''_{i,h} \otimes y''_{i',h'}.$$

Once these replacements are all done, the argument we give below becomes a perfectly valid argument that does not use the Sweedler notation.

So let us come to the actual argument.

We first notice that

$$\Delta_{C\otimes D}(e \otimes f) = \sum_{(e)} \sum_{(f)} e_1 \otimes f_1 \otimes e_2 \otimes f_2 \tag{13.2.41}$$

for every $e \in C$ and $f \in D$.

[388] Some care must be taken here: For example, if the symbol "c_1" appears inside "$(c_1)_1$", then it should not be replaced by "p_i", but instead the whole "$(c_1)_1$" should be replaced by "$p'_{i,j}$".

[*Proof of (13.2.41):* Let $e \in C$ and $f \in D$. Applying both sides of the equality (13.2.38) to $e \otimes f$, we obtain

$$\begin{aligned}
\Delta_{C\otimes D}(e \otimes f) &= ((\operatorname{id}_C \otimes T \otimes \operatorname{id}_D) \circ (\Delta_C \otimes \Delta_D))(e \otimes f) = (\operatorname{id}_C \otimes T \otimes \operatorname{id}_D) \underbrace{((\Delta_C \otimes \Delta_D)(e \otimes f))}_{=\Delta_C(e)\otimes\Delta_D(f)} \\
&= (\operatorname{id}_C \otimes T \otimes \operatorname{id}_D) \left(\underbrace{\Delta_C(e)}_{=\sum_{(e)} e_1 \otimes e_2} \otimes \underbrace{\Delta_D(f)}_{=\sum_{(f)} f_1 \otimes f_2} \right) \\
&= (\operatorname{id}_C \otimes T \otimes \operatorname{id}_D) \left(\underbrace{\left(\sum_{(e)} e_1 \otimes e_2 \right) \otimes \left(\sum_{(f)} f_1 \otimes f_2 \right)}_{=\sum_{(e)} \sum_{(f)} e_1 \otimes e_2 \otimes f_1 \otimes f_2} \right) \\
&= (\operatorname{id}_C \otimes T \otimes \operatorname{id}_D) \left(\sum_{(e)} \sum_{(f)} e_1 \otimes e_2 \otimes f_1 \otimes f_2 \right) \\
&= \sum_{(e)} \sum_{(f)} \underbrace{(\operatorname{id}_C \otimes T \otimes \operatorname{id}_D)(e_1 \otimes e_2 \otimes f_1 \otimes f_2)}_{=\operatorname{id}_C(e_1)\otimes T(e_2\otimes f_1)\otimes \operatorname{id}_D(f_2)} \\
&= \sum_{(e)} \sum_{(f)} \underbrace{\operatorname{id}_C(e_1)}_{=e_1} \otimes \underbrace{T(e_2 \otimes f_1)}_{\substack{=f_1 \otimes e_2 \\ \text{(by the definition of } T)}} \otimes \underbrace{\operatorname{id}_D(f_2)}_{=f_2} \\
&= \sum_{(e)} \sum_{(f)} e_1 \otimes f_1 \otimes e_2 \otimes f_2.
\end{aligned}$$

This proves (13.2.41).]

Furthermore, every $c \in C$ satisfies

$$\sum_{(c)} \sum_{(c_1)} (c_1)_1 \otimes (c_1)_2 \otimes c_2 = \sum_{(c)} \sum_{(c_2)} c_1 \otimes (c_2)_1 \otimes (c_2)_2 . \tag{13.2.42}$$

[*Proof of (13.2.42):* Let $c \in C$. Recall that C is a **k**-coalgebra. Thus, the diagram (1.2.1) commutes (by the definition of a **k**-coalgebra). In other words, we have $(\Delta_C \otimes \operatorname{id}_C) \circ \Delta_C = (\operatorname{id}_C \otimes \Delta_C) \circ \Delta_C$. Applying both sides of this equality to c, we obtain

$$((\Delta_C \otimes \operatorname{id}_C) \circ \Delta_C)(c) = ((\operatorname{id}_C \otimes \Delta_C) \circ \Delta_C)(c) .$$

In light of

$$\begin{aligned}
((\Delta_C \otimes \operatorname{id}_C) \circ \Delta_C)(c) &= (\Delta_C \otimes \operatorname{id}_C) \left(\underbrace{\Delta_C(c)}_{=\sum_{(c)} c_1 \otimes c_2} \right) = (\Delta_C \otimes \operatorname{id}_C) \left(\sum_{(c)} c_1 \otimes c_2 \right) \\
&= \sum_{(c)} \underbrace{\Delta_C(c_1)}_{=\sum_{(c_1)} (c_1)_1 \otimes (c_1)_2} \otimes \underbrace{\operatorname{id}_C(c_2)}_{=c_2} = \sum_{(c)} \left(\sum_{(c_1)} (c_1)_1 \otimes (c_1)_2 \right) \otimes c_2 \\
&= \sum_{(c)} \sum_{(c_1)} (c_1)_1 \otimes (c_1)_2 \otimes c_2
\end{aligned}$$

and

$$\begin{aligned}\left(\left(\operatorname{id}_C \otimes \Delta_C\right) \circ \Delta_C\right)(c) &= \left(\operatorname{id}_C \otimes \Delta_C\right)\left(\underbrace{\Delta_C(c)}_{=\sum_{(c)} c_1 \otimes c_2}\right) = \left(\operatorname{id}_C \otimes \Delta_C\right)\left(\sum_{(c)} c_1 \otimes c_2\right) \\ &= \sum_{(c)} \underbrace{\operatorname{id}_C(c_1)}_{=c_1} \otimes \underbrace{\Delta_C(c_2)}_{=\sum_{(c_2)} (c_2)_1 \otimes (c_2)_2} = \sum_{(c)} c_1 \otimes \left(\sum_{(c_2)} (c_2)_1 \otimes (c_2)_2\right) \\ &= \sum_{(c)} \sum_{(c_2)} c_1 \otimes (c_2)_1 \otimes (c_2)_2,\end{aligned}$$

this rewrites as

$$\sum_{(c)} \sum_{(c_1)} (c_1)_1 \otimes (c_1)_2 \otimes c_2 = \sum_{(c)} \sum_{(c_2)} c_1 \otimes (c_2)_1 \otimes (c_2)_2 .$$

This proves (13.2.42).]

Also, every $d \in D$ satisfies

$$\sum_{(d)} \sum_{(d_1)} (d_1)_1 \otimes (d_1)_2 \otimes d_2 = \sum_{(d)} \sum_{(d_2)} d_1 \otimes (d_2)_1 \otimes (d_2)_2 . \tag{13.2.43}$$

[*Proof of (13.2.43):* This proof is analogous to the proof of (13.2.42).]

Let $z \in C \otimes D$. Thus, z (like any tensor in $C \otimes D$) must be a $\mathbf{k}$-linear combination of pure tensors.

We want to show the equality

$$\left(\left(\Delta_{C\otimes D} \otimes \operatorname{id}_{C\otimes D}\right) \circ \Delta_{C\otimes D}\right)(z) = \left(\left(\operatorname{id}_{C\otimes D} \otimes \Delta_{C\otimes D}\right) \circ \Delta_{C\otimes D}\right)(z) . \tag{13.2.44}$$

This equality is $\mathbf{k}$-linear in z (since all maps that appear in it are $\mathbf{k}$-linear). Hence, we can WLOG assume that z is a pure tensor (since z is a $\mathbf{k}$-linear combination of pure tensors). Assume this. Thus, $z = c \otimes d$ for some $c \in C$ and $d \in D$. Consider these c and d.

Taking the tensor product of the equalities (13.2.42) and (13.2.43), we obtain

$$\begin{aligned}&\left(\sum_{(c)} \sum_{(c_1)} (c_1)_1 \otimes (c_1)_2 \otimes c_2\right) \otimes \left(\sum_{(d)} \sum_{(d_1)} (d_1)_1 \otimes (d_1)_2 \otimes d_2\right) \\ &= \left(\sum_{(c)} \sum_{(c_2)} c_1 \otimes (c_2)_1 \otimes (c_2)_2\right) \otimes \left(\sum_{(d)} \sum_{(d_2)} d_1 \otimes (d_2)_1 \otimes (d_2)_2\right) .\end{aligned}$$

In other words,

$$\begin{aligned}&\sum_{(c)} \sum_{(d)} \sum_{(c_1)} \sum_{(d_1)} (c_1)_1 \otimes (c_1)_2 \otimes c_2 \otimes (d_1)_1 \otimes (d_1)_2 \otimes d_2 \\ &= \sum_{(c)} \sum_{(d)} \sum_{(c_2)} \sum_{(d_2)} c_1 \otimes (c_2)_1 \otimes (c_2)_2 \otimes d_1 \otimes (d_2)_1 \otimes (d_2)_2 .\end{aligned} \tag{13.2.45}$$

Applying the $\mathbf{k}$-linear map

$$C \otimes C \otimes C \otimes D \otimes D \otimes D \to C \otimes D \otimes C \otimes D \otimes C \otimes D,$$
$$\gamma_1 \otimes \gamma_2 \otimes \gamma_3 \otimes \delta_1 \otimes \delta_2 \otimes \delta_3 \mapsto \gamma_1 \otimes \delta_1 \otimes \gamma_2 \otimes \delta_2 \otimes \gamma_3 \otimes \delta_3$$

to both sides of this equality, we obtain

$$\begin{aligned}&\sum_{(c)} \sum_{(d)} \sum_{(c_1)} \sum_{(d_1)} (c_1)_1 \otimes (d_1)_1 \otimes (c_1)_2 \otimes (d_1)_2 \otimes c_2 \otimes d_2 \\ &= \sum_{(c)} \sum_{(d)} \sum_{(c_2)} \sum_{(d_2)} c_1 \otimes d_1 \otimes (c_2)_1 \otimes (d_2)_1 \otimes (c_2)_2 \otimes (d_2)_2 .\end{aligned} \tag{13.2.46}$$

Now,

$$\begin{aligned}
&\left(\left(\Delta_{C\otimes D}\otimes\operatorname{id}_{C\otimes D}\right)\circ\Delta_{C\otimes D}\right)(z)\\
&=\left(\Delta_{C\otimes D}\otimes\operatorname{id}_{C\otimes D}\right)\left(\Delta_{C\otimes D}\left(\underbrace{z}_{=c\otimes d}\right)\right)\\
&=\left(\Delta_{C\otimes D}\otimes\operatorname{id}_{C\otimes D}\right)\left(\underbrace{\Delta_{C\otimes D}\left(c\otimes d\right)}_{\substack{=\sum_{(c)}\sum_{(d)}c_1\otimes d_1\otimes c_2\otimes d_2\\ \text{(by (13.2.41) (applied to } e=c \text{ and } f=d))}}\right)\\
&=\left(\Delta_{C\otimes D}\otimes\operatorname{id}_{C\otimes D}\right)\left(\sum_{(c)}\sum_{(d)}c_1\otimes d_1\otimes c_2\otimes d_2\right)\\
&=\sum_{(c)}\sum_{(d)}\underbrace{\left(\Delta_{C\otimes D}\otimes\operatorname{id}_{C\otimes D}\right)\left(c_1\otimes d_1\otimes c_2\otimes d_2\right)}_{=\Delta_{C\otimes D}(c_1\otimes d_1)\otimes\operatorname{id}_{C\otimes D}(c_2\otimes d_2)}\\
&=\sum_{(c)}\sum_{(d)}\underbrace{\Delta_{C\otimes D}\left(c_1\otimes d_1\right)}_{\substack{=\sum_{(c_1)}\sum_{(d_1)}(c_1)_1\otimes(d_1)_1\otimes(c_1)_2\otimes(d_1)_2\\ \text{(by (13.2.41) (applied to } e=c_1 \text{ and } f=d_1))}}\otimes\underbrace{\operatorname{id}_{C\otimes D}\left(c_2\otimes d_2\right)}_{=c_2\otimes d_2}\\
&=\sum_{(c)}\sum_{(d)}\left(\sum_{(c_1)}\sum_{(d_1)}\left(c_1\right)_1\otimes\left(d_1\right)_1\otimes\left(c_1\right)_2\otimes\left(d_1\right)_2\right)\otimes c_2\otimes d_2\\
&=\sum_{(c)}\sum_{(d)}\sum_{(c_1)}\sum_{(d_1)}\left(c_1\right)_1\otimes\left(d_1\right)_1\otimes\left(c_1\right)_2\otimes\left(d_1\right)_2\otimes c_2\otimes d_2; \qquad (13.2.47)
\end{aligned}$$

an analogous computation shows that

$$\begin{aligned}
&\left(\left(\operatorname{id}_{C\otimes D}\otimes\Delta_{C\otimes D}\right)\circ\Delta_{C\otimes D}\right)(z)\\
&=\sum_{(c)}\sum_{(d)}\sum_{(c_2)}\sum_{(d_2)}c_1\otimes d_1\otimes\left(c_2\right)_1\otimes\left(d_2\right)_1\otimes\left(c_2\right)_2\otimes\left(d_2\right)_2. \qquad (13.2.48)
\end{aligned}$$

In light of (13.2.47) and (13.2.48), the equality (13.2.46) rewrites as

$$\left(\left(\Delta_{C\otimes D}\otimes\operatorname{id}_{C\otimes D}\right)\circ\Delta_{C\otimes D}\right)(z)=\left(\left(\operatorname{id}_{C\otimes D}\otimes\Delta_{C\otimes D}\right)\circ\Delta_{C\otimes D}\right)(z).$$

Thus, the equality (13.2.44) is proven.

Now, forget that we fixed z. We thus have proven the equality (13.2.44) for every $z\in C\otimes D$. In other words, we have $\left(\Delta_{C\otimes D}\otimes\operatorname{id}_{C\otimes D}\right)\circ\Delta_{C\otimes D}=\left(\operatorname{id}_{C\otimes D}\otimes\Delta_{C\otimes D}\right)\circ\Delta_{C\otimes D}$. In other words, the diagram (13.2.39) commutes.

It remains to prove that the diagram (13.2.40) commutes. We leave this to the reader, as the proof is similar to (but simpler than) the proof of the commutativity of (13.2.39) given above. Hence, both diagrams (13.2.39) and (13.2.40) commute. In other words, the $\mathbf{k}$-module $C\otimes D$, equipped with the two maps $\Delta_{C\otimes D}$ and $\epsilon_{C\otimes D}$, is a $\mathbf{k}$-coalgebra. In other words, the $\mathbf{k}$-coalgebra $C\otimes D$ introduced in Definition 1.3.3 is actually well-defined. This solves Exercise 1.3.4(b).

Second solution to Exercise 1.3.4(b). Let θ be the canonical $\mathbf{k}$-module isomorphism $\mathbf{k}\otimes\mathbf{k}\to\mathbf{k}$.

A solution to Exercise 1.3.4(b) can now be obtained in a straightforward fashion from the Second solution to Exercise 1.3.4(a), by making the following modifications:

(1) Replace every appearance of any of the terms

$$A,\ B,\ m_A,\ m_B,\ m_{A\otimes B},\ u_A,\ u_B,\ u_{A\otimes B},\ \xi,\ \operatorname{kan}_{1,U},\ \operatorname{kan}_{2,U},\ T_{U,V}$$

(for any $\mathbf{k}$-modules U and V) by

$$C,\ D,\ \Delta_C,\ \Delta_D,\ \Delta_{C\otimes D},\ u_C,\ u_D,\ u_{C\otimes D},\ \theta,\ \operatorname{kan}_{1,U}^{-1},\ \operatorname{kan}_{2,U}^{-1},\ T_{V,U},$$

respectively.

(For example, the equation (13.2.25) becomes

$$(\operatorname{id}_C \otimes T_{C,D} \otimes \operatorname{id}_D) \circ (\Delta_{C\otimes D} \otimes \operatorname{id}_{C\otimes D}) = (\Delta_C \otimes \operatorname{id}_C \otimes \Delta_D \otimes \operatorname{id}_D) \circ (\operatorname{id}_C \otimes Q \otimes \operatorname{id}_D) . \tag{13.2.49}$$

This new equation makes no sense, because (for example) the maps $\operatorname{id}_C \otimes T_{D,C} \otimes \operatorname{id}_D$ and $\Delta_{C\otimes D} \otimes \operatorname{id}_{C\otimes D}$ cannot be composed; but this is fine, since we shall make a further modification which will turn this equation into a meaningful one.

For another example, the map $Q : B \otimes A \otimes B \otimes A \to A \otimes A \otimes B \otimes B$ becomes a map $Q : D \otimes C \otimes D \otimes C \to C \otimes C \otimes D \otimes D$.)

(2) Reverse the direction of all arrows[389]. In other words, any map which used to go from a set X to a set Y shall now go from Y to X.

(For example, the map $Q : D \otimes C \otimes D \otimes C \to C \otimes C \otimes D \otimes D$ becomes a map $Q : C \otimes C \otimes D \otimes D \to D \otimes C \otimes D \otimes C$.)

(3) In any composition of $\mathbf{k}$-linear maps, reverse the order of the maps being composed. In other words, replace any composition $f_1 \circ f_2 \circ \cdots \circ f_k$ by $f_k \circ f_{k-1} \circ \cdots \circ f_1$.

(For example, the meaningless equality (13.2.49) thus becomes

$$(\Delta_{C\otimes D} \otimes \operatorname{id}_{C\otimes D}) \circ (\operatorname{id}_C \otimes T_{C,D} \otimes \operatorname{id}_D) = (\operatorname{id}_C \otimes Q \otimes \operatorname{id}_D) \circ (\Delta_C \otimes \operatorname{id}_C \otimes \Delta_D \otimes \operatorname{id}_D) .$$

This equality is meaningful and correct.)

(4) Any part of our solution that involved elements of A and B (as opposed to mere computations with maps) must be redone from scratch. For example, the proof of (13.2.21) involved elements of A and B (because we picked $b \in B$, $a \in A$, $b' \in B$ and $a' \in A$ in that proof), and thus must be redone, whereas the proof of (13.2.31) did not involve elements of A and B and therefore needs not be modified any further.

Fortunately, very few parts of our solution involved elements of A and B. To wit, these parts are the proofs of the equalities (13.2.21), (13.2.29), (13.2.34) and (13.2.37). After the above modifications, these equalities have become

$$\begin{aligned}
&Q = (T_{C,D} \otimes \operatorname{id}_D \otimes \operatorname{id}_C) \circ (\operatorname{id}_C \otimes T_{C,D\otimes D}) , \\
&Q = (\operatorname{id}_D \otimes \operatorname{id}_C \otimes T_{C,D}) \circ (T_{C\otimes C,D} \otimes \operatorname{id}_D) , \\
&\operatorname{kan}_{1,C\otimes D}^{-1} \circ (\operatorname{id}_{C\otimes D} \otimes \theta) \circ (\operatorname{id}_C \otimes T_{\mathbf{k},D} \otimes \operatorname{id}_{\mathbf{k}}) = \operatorname{kan}_{1,C}^{-1} \otimes \operatorname{kan}_{1,D}^{-1}, \qquad \text{and} \\
&\operatorname{kan}_{2,C\otimes D}^{-1} \circ (\theta \otimes \operatorname{id}_{C\otimes D}) \circ (\operatorname{id}_{\mathbf{k}} \otimes T_{C,\mathbf{k}} \otimes \operatorname{id}_D) = \operatorname{kan}_{2,C}^{-1} \otimes \operatorname{kan}_{2,D}^{-1},
\end{aligned}$$

respectively. So these four equalities must be proven. Fortunately, these proofs are completely straightforward[390]; the reader can easily come up with them.

These four modifications are sufficient to transform our Second solution to Exercise 1.3.4(a) into a solution to Exercise 1.3.4(b). Thus, Exercise 1.3.4(b) is solved again.

[*Remark:* The second and the third modifications made above are usually subsumed under the concept of "reversing all arrows".]

13.3. **Solution to Exercise 1.3.6.** *Solution to Exercise 1.3.6.*

(b) Let C, C', D and D' be four $\mathbf{k}$-coalgebras. Let $f : C \to C'$ and $g : D \to D'$ be two $\mathbf{k}$-coalgebra homomorphisms. We need to prove that $f \otimes g : C \otimes D \to C' \otimes D'$ is a $\mathbf{k}$-coalgebra homomorphism.

Recall that (by the definition of a "$\mathbf{k}$-coalgebra homomorphism") the map $f \otimes g : C \otimes D \to C' \otimes D'$ is a $\mathbf{k}$-coalgebra homomorphism if and only if the two diagrams

(13.3.1)
$$\begin{tikzcd}
C \otimes D \arrow[r, "f\otimes g"] \arrow[d, "\Delta_{C\otimes D}"'] & C' \otimes D' \arrow[d, "\Delta_{C'\otimes D'}"] \\
(C \otimes D) \otimes (C \otimes D) \arrow[r, "(f\otimes g)\otimes(f\otimes g)"] & (C' \otimes D') \otimes (C' \otimes D')
\end{tikzcd}$$

[389] This includes both the arrows in the description of maps and the arrows in commutative diagrams.

[390] These equalities are properties of tensor products of $\mathbf{k}$-modules; they make no use of the coalgebra structures on C and D.

and

(13.3.2)
$$\begin{tikzcd} C \otimes D \arrow[rr, "f \otimes g"] \arrow[dr, "\epsilon_{C \otimes D}"'] & & C' \otimes D' \arrow[dl, "\epsilon_{C' \otimes D'}"] \\ & \mathbf{k} & \end{tikzcd}$$

commute. We shall now prove that these diagrams indeed commute.

We know that the map $f : C \to C'$ is a $\mathbf{k}$-coalgebra homomorphism. By the definition of a $\mathbf{k}$-coalgebra homomorphism, this means that the two diagrams

(13.3.3)
$$\begin{tikzcd} C \arrow[r, "f"] \arrow[d, "\Delta_C"'] & C' \arrow[d, "\Delta_{C'}"] \\ C \otimes C \arrow[r, "f \otimes f"] & C' \otimes C' \end{tikzcd}$$

and

(13.3.4)
$$\begin{tikzcd} C \arrow[rr, "f"] \arrow[dr, "\epsilon_C"'] & & C' \arrow[dl, "\epsilon_{C'}"] \\ & \mathbf{k} & \end{tikzcd}$$

commute.

For any two $\mathbf{k}$-modules U and V, let $T_{U,V} : U \otimes V \to V \otimes U$ be the twist map (i.e., the $\mathbf{k}$-linear map $U \otimes V \to V \otimes U$ sending every $u \otimes v$ to $v \otimes u$). A simple linear-algebraic fact says that if U, V, U' and V' are four $\mathbf{k}$-modules and $x : U \to U'$ and $y : V \to V'$ are two $\mathbf{k}$-linear maps, then

$$(y \otimes x) \circ T_{U,V} = T_{U',V'} \circ (x \otimes y) . \tag{13.3.5}$$

The definition of the $\mathbf{k}$-coalgebra $C \otimes D$ yields

$$\Delta_{C \otimes D} = (\mathrm{id}_C \otimes T_{C,D} \otimes \mathrm{id}_D) \circ (\Delta_C \otimes \Delta_D) . \tag{13.3.6}$$

Similarly,

$$\Delta_{C' \otimes D'} = (\mathrm{id}_{C'} \otimes T_{C',D'} \otimes \mathrm{id}_{D'}) \circ (\Delta_{C'} \otimes \Delta_{D'}) . \tag{13.3.7}$$

But $(f \otimes f) \circ \Delta_C = \Delta_{C'} \circ f$ (since the diagram (13.3.3) commutes), and similarly $(g \otimes g) \circ \Delta_D = \Delta_{D'} \circ g$.

Now,

$$
\begin{aligned}
&\underbrace{((f\otimes g)\otimes(f\otimes g))}_{=f\otimes(g\otimes f)\otimes g}\circ \underbrace{\Delta_{C\otimes D}}_{=(\mathrm{id}_C\otimes T_{C,D}\otimes\mathrm{id}_D)\circ(\Delta_C\otimes\Delta_D)}\\
&= \underbrace{(f\otimes(g\otimes f)\otimes g)\circ(\mathrm{id}_C\otimes T_{C,D}\otimes\mathrm{id}_D)}_{=(f\circ\mathrm{id}_C)\otimes((g\otimes f)\circ T_{C,D})\otimes(g\circ\mathrm{id}_D)}\circ(\Delta_C\otimes\Delta_D)\\
&= \left(\underbrace{(f\circ\mathrm{id}_C)}_{=f=\mathrm{id}_{C'}\circ f}\otimes\underbrace{((g\otimes f)\circ T_{C,D})}_{\substack{=T_{C',D'}\circ(f\otimes g)\\ \text{(by (13.3.5))}}}\otimes\underbrace{(g\circ\mathrm{id}_D)}_{=g=\mathrm{id}_{D'}\circ g}\right)\circ(\Delta_C\otimes\Delta_D)\\
&= \underbrace{((\mathrm{id}_{C'}\circ f)\otimes(T_{C',D'}\circ(f\otimes g))\otimes(\mathrm{id}_{D'}\circ g))}_{=(\mathrm{id}_{C'}\otimes T_{C',D'}\otimes\mathrm{id}_{D'})\circ(f\otimes(f\otimes g)\otimes g)}\circ(\Delta_C\otimes\Delta_D)\\
&= (\mathrm{id}_{C'}\otimes T_{C',D'}\otimes\mathrm{id}_{D'})\circ\underbrace{(f\otimes(f\otimes g)\otimes g)}_{=(f\otimes f)\otimes(g\otimes g)}\circ(\Delta_C\otimes\Delta_D)\\
&= (\mathrm{id}_{C'}\otimes T_{C',D'}\otimes\mathrm{id}_{D'})\circ\underbrace{((f\otimes f)\otimes(g\otimes g))\circ(\Delta_C\otimes\Delta_D)}_{=((f\otimes f)\circ\Delta_C)\otimes((g\otimes g)\circ\Delta_D)}\\
&= (\mathrm{id}_{C'}\otimes T_{C',D'}\otimes\mathrm{id}_{D'})\circ\left(\underbrace{((f\otimes f)\circ\Delta_C)}_{=\Delta_{C'}\circ f}\otimes\underbrace{((g\otimes g)\circ\Delta_D)}_{=\Delta_{D'}\circ g}\right)\\
&= (\mathrm{id}_{C'}\otimes T_{C',D'}\otimes\mathrm{id}_{D'})\circ\underbrace{((\Delta_{C'}\circ f)\otimes(\Delta_{D'}\circ g))}_{=(\Delta_{C'}\otimes\Delta_{D'})\circ(f\otimes g)}\\
&= \underbrace{(\mathrm{id}_{C'}\otimes T_{C',D'}\otimes\mathrm{id}_{D'})\circ(\Delta_{C'}\otimes\Delta_{D'})}_{\substack{=\Delta_{C'\otimes D'}\\ \text{(by (13.3.7))}}}\circ(f\otimes g)=\Delta_{C'\otimes D'}\circ(f\otimes g)\,.
\end{aligned}
$$

In other words, the diagram (13.3.1) commutes.

We have $\epsilon_{C'}\circ f=\epsilon_C$ (since the diagram (13.3.4) commutes) and $\epsilon_{D'}\circ g=\epsilon_D$ (similarly).

Now, let θ be the canonical **k**-module isomorphism $\mathbf{k}\otimes\mathbf{k}\to\mathbf{k}$. Then, the definition of the **k**-coalgebra $C\otimes D$ yields

$$\epsilon_{C\otimes D}=\theta\circ(\epsilon_C\otimes\epsilon_D)\,. \tag{13.3.8}$$

Similarly,

$$\epsilon_{C'\otimes D'}=\theta\circ(\epsilon_{C'}\otimes\epsilon_{D'})\,. \tag{13.3.9}$$

Now,

$$
\begin{aligned}
\underbrace{\epsilon_{C'\otimes D'}}_{=\theta\circ(\epsilon_{C'}\otimes\epsilon_{D'})}\circ(f\otimes g)&=\theta\circ\underbrace{(\epsilon_{C'}\otimes\epsilon_{D'})\circ(f\otimes g)}_{=(\epsilon_{C'}\circ f)\otimes(\epsilon_{D'}\circ g)}\\
&=\theta\circ\left(\underbrace{(\epsilon_{C'}\circ f)}_{=\epsilon_C}\otimes\underbrace{(\epsilon_{D'}\circ g)}_{=\epsilon_D}\right)=\theta\circ(\epsilon_C\otimes\epsilon_D)=\epsilon_{C\otimes D}
\end{aligned}
$$

(by (13.3.8)). In other words, the diagram (13.3.2) commutes.

We now know that the two diagrams (13.3.1) and (13.3.2) commute. Thus, the map $f\otimes g:C\otimes D\to C'\otimes D'$ is a **k**-coalgebra homomorphism (because we know that the map $f\otimes g:C\otimes D\to C'\otimes D'$ is a **k**-coalgebra homomorphism if and only if the two diagrams (13.3.1) and (13.3.2) commute). This solves Exercise 1.3.6(b).

(a) In order to obtain a solution to Exercise 1.3.6(a), it is enough to reverse all arrows in the above solution to Exercise 1.3.6(b) (and, of course, replace C, D, Δ_C etc. by A, B, m_A etc.).

13.4. **Solution to Exercise 1.3.13.**

Solution to Exercise 1.3.13. (a) Assume that f is surjective. It is a known fact from linear algebra[391] that if U, V, U' and V' are four $\mathbf{k}$-modules, and $\phi : U \to U'$ and $\psi : V \to V'$ are two surjective $\mathbf{k}$-linear maps, then the kernel of $\phi \otimes \psi : U \otimes V \to U' \otimes V'$ is

$$\ker(\phi \otimes \psi) = (\ker \phi) \otimes V + U \otimes (\ker \psi).$$

Applying this to $U = A$, $U' = C$, $V = A$, $V' = C$, $\phi = f$ and $\psi = f$, we obtain $\ker(f \otimes f) = (\ker f) \otimes A + A \otimes (\ker f)$.

But f is a coalgebra homomorphism, so that $\epsilon = \epsilon \circ f$. Hence, every $x \in \ker f$ satisfies $\underbrace{\epsilon}_{=\epsilon\circ f}(x) = (\epsilon \circ f)(x) = \epsilon\left(\underbrace{f(x)}_{\substack{=0\\ (\text{since } x\in\ker f)}}\right) = \epsilon(0) = 0$. In other words, $\epsilon(\ker f) = 0$. Also,

$$\ker f \subset \ker \underbrace{(\Delta \circ f)}_{\substack{=(f\otimes f)\circ\Delta \\ (\text{since } f \text{ is a coalgebra} \\ \text{homomorphism})}} = \ker((f \otimes f) \circ \Delta) = \Delta^{-1}(\ker(f \otimes f)),$$

so that

$$\Delta(\ker f) \subset \ker(f \otimes f) = (\ker f) \otimes A + A \otimes (\ker f).$$

Combined with $\epsilon(\ker f) = 0$, this shows that $\ker f$ is a two-sided coideal of A. Thus, Exercise 1.3.13(a) is solved.

(b) Assume that $\mathbf{k}$ is a field. Then, it is a known fact from linear algebra[392] that if U, V, U' and V' are four $\mathbf{k}$-modules, and $\phi : U \to U'$ and $\psi : V \to V'$ are two $\mathbf{k}$-linear maps, then the kernel of $\phi \otimes \psi : U \otimes V \to U' \otimes V'$ is

$$\ker(\phi \otimes \psi) = (\ker \phi) \otimes V + U \otimes (\ker \psi).$$

Starting from this point, we can continue arguing as in the solution of part (a). Thus, Exercise 1.3.13(b) is solved.

13.5. **Solution to Exercise 1.3.18.**

Solution to Exercise 1.3.18.

It is not hard to solve Exercise 1.3.18 directly; we shall take a longer but somewhat more elegant approach. We begin with a definition:

Definition 13.5.1. Let $V = \bigoplus_{n\in\mathbb{N}} V_n$ be a graded $\mathbf{k}$-module (where the V_n are the homogeneous components of V). Let $n \in \mathbb{N}$. Then, we shall let $\pi_{n,V} : V \to V_n$ denote the canonical projection from V to its n-th graded component V_n.

Let us show a few properties of these projections:

Lemma 13.5.2. *Let V and W be two graded $\mathbf{k}$-modules. Let $f : V \to W$ be a graded $\mathbf{k}$-linear map. Let $m \in \mathbb{N}$. Let v be a homogeneous element of V. Then, $\pi_{m,W}(f(v)) = f(\pi_{m,V}(v))$.*

Proof of Lemma 13.5.2. Write V as $V = \bigoplus_{n\in\mathbb{N}} V_n$ (where the V_n are the homogeneous components of V). Write W as $W = \bigoplus_{n\in\mathbb{N}} W_n$ (where the W_n are the homogeneous components of W).

The map $\pi_{m,V}$ is the canonical projection from V to its m-th graded component V_m. Hence, $\pi_{m,V}$ fixes every element of V_m. In other words, we have

$$\pi_{m,V}(p) = p \qquad \text{for each } p \in V_m. \tag{13.5.1}$$

The same argument (applied to W and W_m instead of V and V_m) yields

$$\pi_{m,W}(p) = p \qquad \text{for each } p \in W_m. \tag{13.5.2}$$

[391] proven, e.g., in Keith Conrad's [40, “Tensor Products II”, Thm. 2.19]

[392] proven, e.g., in Keith Conrad's [40, “Tensor Products II”, Thm. 5.5]

The map $\pi_{m,V}$ is the canonical projection from V to its m-th graded component V_m. Hence, $\pi_{m,V}$ annihilates all graded components of V other than V_m. In other words,

$$\pi_{m,V}\left(V_n\right)=0 \qquad \text{for every } n\in\mathbb{N} \text{ satisfying } n\neq m. \tag{13.5.3}$$

The same argument (applied to W and W_m instead of V and V_m) yields

$$\pi_{m,W}\left(W_n\right)=0 \qquad \text{for every } n\in\mathbb{N} \text{ satisfying } n\neq m. \tag{13.5.4}$$

Recall that v is a homogeneous element of V. Thus, there exists some $n\in\mathbb{N}$ such that $v\in V_n$. Consider this n. From $v\in V_n$, we obtain $f\left(v\right)\in f\left(V_n\right)\subset W_n$ (since the map f is graded). We must show that $\pi_{m,W}\left(f\left(v\right)\right)=f\left(\pi_{m,V}\left(v\right)\right)$.

We are in one of the following two cases:

Case 1: We have $n=m$.

Case 2: We have $n\neq m$.

Let us first consider Case 1 first. In this case, we have $n=m$. Now, $v\in V_n=V_m$ (since $n=m$). Thus, (13.5.1) (applied to $p=v$) yields $\pi_{m,V}\left(v\right)=v$. Hence, $v=\pi_{m,V}\left(v\right)$. Also, $f\left(v\right)\in W_n=W_m$ (since $n=m$). Hence, (13.5.2) (applied to $p=f\left(v\right)$) yields

$$\pi_{m,W}\left(f\left(v\right)\right)=f\left(\underbrace{v}_{=\pi_{m,V}\left(v\right)}\right)=f\left(\pi_{m,V}\left(v\right)\right).$$

Thus, $\pi_{m,W}\left(f\left(v\right)\right)=f\left(\pi_{m,V}\left(v\right)\right)$ is proved in Case 1.

Let us next consider Case 2. In this case, we have $n\neq m$. But $v\in V_n$, so that $\pi_{m,V}\left(v\right)\in\pi_{m,V}\left(V_n\right)=0$ (by (13.5.3)). Hence, $\pi_{m,V}\left(v\right)=0$. Therefore, $f\left(\underbrace{\pi_{m,V}\left(v\right)}_{=0}\right)=f\left(0\right)=0$ (since the map f is $\mathbf{k}$-linear). On the other hand, $\pi_{m,W}\left(\underbrace{f\left(v\right)}_{\in W_n}\right)\in\pi_{m,W}\left(W_n\right)=0$ (by (13.5.4)), so that

$$\pi_{m,W}\left(f\left(v\right)\right)=0=f\left(\pi_{m,V}\left(v\right)\right) \qquad \left(\text{since } f\left(\pi_{m,V}\left(v\right)\right)=0\right).$$

Thus, $\pi_{m,W}\left(f\left(v\right)\right)=f\left(\pi_{m,V}\left(v\right)\right)$ is proved in Case 2.

We have now proved $\pi_{m,W}\left(f\left(v\right)\right)=f\left(\pi_{m,V}\left(v\right)\right)$ in each of these two Cases 1 and 2. Since these two Cases cover all possibilities, we thus conclude that $\pi_{m,W}\left(f\left(v\right)\right)=f\left(\pi_{m,V}\left(v\right)\right)$ always holds.

This proves Lemma 13.5.2. □

Proposition 13.5.3. *Let V and W be two graded $\mathbf{k}$-modules. Let $f:V\to W$ be a graded $\mathbf{k}$-linear map. Let $m\in\mathbb{N}$. Let $v\in V$ be arbitrary. Then, $\pi_{m,W}\left(f\left(v\right)\right)=f\left(\pi_{m,V}\left(v\right)\right)$.*

Proof of Proposition 13.5.3. Every element of V is a sum of homogeneous elements (since V is graded). Thus, v is a sum of homogeneous elements (since v is an element of V). In other words, we can write v in the form $v=\sum_{i=1}^{k}v_i$, where $k\in\mathbb{N}$, and where $v_1,v_2,\ldots,v_k$ are homogeneous elements of V. Consider this k and these $v_1,v_2,\ldots,v_k$.

For each $i\in\{1,2,\ldots,k\}$, the element v_i of V is homogeneous (since $v_1,v_2,\ldots,v_k$ are homogeneous elements of V) and therefore satisfies

$$\pi_{m,W}\left(f\left(v_i\right)\right)=f\left(\pi_{m,V}\left(v_i\right)\right) \tag{13.5.5}$$

(by Lemma 13.5.2, applied to v_i instead of v). Now,

$$\pi_{m,W}\left(f\left(\underbrace{v}_{=\sum_{i=1}^{k} v_i}\right)\right) = \pi_{m,W}\left(\underbrace{f\left(\sum_{i=1}^{k} v_i\right)}_{\substack{=\sum_{i=1}^{k} f(v_i) \\ \text{(since the map } f \text{ is } \mathbf{k}\text{-linear)}}}\right) = \pi_{m,W}\left(\sum_{i=1}^{k} f(v_i)\right)$$
$$= \sum_{i=1}^{k} \underbrace{\pi_{m,W}(f(v_i))}_{\substack{=f(\pi_{m,V}(v_i)) \\ \text{(by (13.5.5))}}} \qquad \text{(since the map } \pi_{m,W} \text{ is } \mathbf{k}\text{-linear)}$$
$$= \sum_{i=1}^{k} f(\pi_{m,V}(v_i)).$$

Comparing this with

$$f\left(\pi_{m,V}\left(\underbrace{v}_{=\sum_{i=1}^{k} v_i}\right)\right) = f\left(\underbrace{\pi_{m,V}\left(\sum_{i=1}^{k} v_i\right)}_{\substack{=\sum_{i=1}^{k} \pi_{m,V}(v_i) \\ \text{(since the map } \pi_{m,V} \text{ is } \mathbf{k}\text{-linear)}}}\right) = f\left(\sum_{i=1}^{k} \pi_{m,V}(v_i)\right)$$
$$= \sum_{i=1}^{k} f(\pi_{m,V}(v_i)) \qquad \text{(since the map } f \text{ is } \mathbf{k}\text{-linear)},$$

we obtain $\pi_{m,W}(f(v)) = f(\pi_{m,V}(v))$. This proves Proposition 13.5.3. □

Let us now come to the solution of Exercise 1.3.18. Let $f : V \to W$ be an invertible graded $\mathbf{k}$-linear map. We must prove that its inverse $f^{-1} : W \to V$ is also graded. In other words, we must prove that $f^{-1}(W_n) \subset V_n$ for each $n \in \mathbb{N}$.

So let $n \in \mathbb{N}$. Let $v \in f^{-1}(W_n)$. Thus, $v \in V$ and $f(v) \in W_n$.

The map $\pi_{n,W}$ is the canonical projection from W to its n-th graded component W_n. Hence, $\pi_{n,W}$ fixes every element of W_n. In other words, we have $\pi_{n,W}(p) = p$ for each $p \in W_n$. Applying this to $p = f(v)$, we obtain $\pi_{n,W}(f(v)) = f(v)$ (since $f(v) \in W_n$). But Proposition 13.5.3 (applied to $m = n$) yields $\pi_{n,W}(f(v)) = f(\pi_{n,V}(v))$. Comparing these two equalities, we obtain $f(v) = f(\pi_{n,V}(v))$.

But the map f is invertible, thus injective. In other words, if $a, b \in V$ are two elements satisfying $f(a) = f(b)$, then $a = b$. Applying this to $a = v$ and $b = \pi_{n,V}(v)$, we find $v = \pi_{n,V}(v)$ (since $f(v) = f(\pi_{n,V}(v))$).

But the map $\pi_{n,V}$ is the canonical projection from V to its n-th graded component V_n. Thus, $\pi_{n,V}(V) \subset V_n$. Hence, $\pi_{n,V}(v) \in V_n$ (since $v \in V$). Thus, $v = \pi_{n,V}(v) \in V_n$.

Forget that we fixed v. We thus have shown that $v \in V_n$ for each $v \in f^{-1}(W_n)$. In other words, $f^{-1}(W_n) \subset V_n$.

Forget that we fixed n. We thus have shown that $f^{-1}(W_n) \subset V_n$ for each $n \in \mathbb{N}$. In other words, the map f^{-1} is graded. This solves Exercise 1.3.18.

13.6. **Solution to Exercise 1.3.19.** *Solution to Exercise 1.3.19.*

(a) Define a map $\widetilde{\Delta} : A \to A \otimes A$ by

$$\widetilde{\Delta}(x) = \Delta(x) - (x \otimes 1 + 1 \otimes x) \qquad \text{for all } x \in A.$$

It is easily seen that $\widetilde{\Delta}$ is a homomorphism of graded $\mathbf{k}$-modules. Hence, the kernel $\ker \widetilde{\Delta}$ of $\widetilde{\Delta}$ is a graded $\mathbf{k}$-submodule of A (since the kernel of a homomorphism of graded $\mathbf{k}$-modules always is a graded $\mathbf{k}$-submodule of the domain).

We have defined $\mathfrak{p}$ as the set of all primitive elements of A. In other words,

$$\begin{aligned}\mathfrak{p} &= \left\{x \in A \ \mid \ \Delta(x) = x \otimes 1 + 1 \otimes x\right\} \\ &= \left\{ x \in A \ \mid \ \underbrace{\Delta(x) - (x \otimes 1 + 1 \otimes x)}_{\substack{=\widetilde{\Delta}(x) \\ (\text{since } \widetilde{\Delta}(x) = \Delta(x) - (x \otimes 1 + 1 \otimes x))}} = 0 \right\} \\ &= \left\{x \in A \ \mid \ \widetilde{\Delta}(x) = 0\right\} = \ker \widetilde{\Delta}.\end{aligned}$$

Thus, $\mathfrak{p}$ is a graded $\mathbf{k}$-submodule of A (since we know that $\ker \widetilde{\Delta}$ is a graded $\mathbf{k}$-submodule of A). This solves Exercise 1.3.19 (a).

(b) We notice first that

$$\epsilon(x) = 0 \qquad \text{for every } x \in \mathfrak{p}. \tag{13.6.1}$$

[*Proof of* (13.6.1). We know that A is a $\mathbf{k}$-bialgebra. Thus, ϵ is a $\mathbf{k}$-algebra homomorphism (by the definition of a $\mathbf{k}$-bialgebra). Hence, $\epsilon(1_A) = 1_{\mathbf{k}}$ and $\epsilon(0) = 0$.

Let $x \in \mathfrak{p}$. Thus, x is primitive, so that $\Delta(x) = x \otimes 1 + 1 \otimes x$ (by the definition of a primitive element). Applying the map $\epsilon \otimes \mathrm{id}$ to both sides of this equality, we obtain

$$(\epsilon \otimes \mathrm{id})(\Delta(x)) = (\epsilon \otimes \mathrm{id})(x \otimes 1 + 1 \otimes x) = \epsilon(x) \cdot \mathrm{id}(1) + \epsilon(1) \cdot \mathrm{id}(x)$$

(where we are identifying $\mathbf{k} \otimes A$ with A along the canonical isomorphism). Compared with $(\epsilon \otimes \mathrm{id})(\Delta(x)) = x$ (this is a consequence of the axioms of a coalgebra), this yields

$$x = \epsilon(x) \cdot \underbrace{\mathrm{id}(1)}_{=1=1_A} + \underbrace{\epsilon(1)}_{=\epsilon(1_A)=1_{\mathbf{k}}} \cdot \underbrace{\mathrm{id}(x)}_{=x} = \epsilon(x) \cdot 1_A + x.$$

Subtracting x from this equality, we obtain $0 = \epsilon(x) \cdot 1_A$. Applying the map ϵ to this equality, we find $\epsilon(0) = \epsilon(\epsilon(x) \cdot 1_A) = \epsilon(x) \cdot \underbrace{\epsilon(1_A)}_{=1_{\mathbf{k}}} = \epsilon(x)$. Hence, $\epsilon(x) = \epsilon(0) = 0$. This proves (13.6.1).]

(*Note:* We will reprove (13.6.1) below in Proposition 1.4.17.)

Now, for every $x \in \mathfrak{p}$, we have

$$\begin{aligned}\Delta(x) &= \underbrace{x}_{\in \mathfrak{p}} \otimes \underbrace{1}_{\in A} + \underbrace{1}_{\in A} \otimes \underbrace{x}_{\in \mathfrak{p}} \qquad (\text{since } x \in \mathfrak{p}, \text{ so that } x \text{ is primitive}) \\ &\in \mathfrak{p} \otimes A + A \otimes \mathfrak{p}.\end{aligned}$$

In other words, $\Delta(\mathfrak{p}) \subset \mathfrak{p} \otimes A + A \otimes \mathfrak{p}$. Combined with $\epsilon(\mathfrak{p}) = 0$ (this follows from (13.6.1)), this yields that $\mathfrak{p}$ is a two-sided coideal of A. This solves Exercise 1.3.19 (b).

13.7. Solution to Exercise 1.3.20.

Solution to Exercise 1.3.20. (a) The unit map $u : \mathbf{k} \to A$ is graded (since A is a graded algebra). Hence, $u(\mathbf{k}_0) \subset A_0$, where $\mathbf{k}_0$ denotes the 0-th graded component of $\mathbf{k}$. But the grading on $\mathbf{k}$ is such that $\mathbf{k}_0 = \mathbf{k}$. Thus, $u(\mathbf{k}_0) = u(\mathbf{k})$, so that $u(\mathbf{k}) = u(\mathbf{k}_0) \subset A_0$. But

$$u(\mathbf{k}) = \left\{ \underbrace{u(\lambda)}_{\substack{=\lambda \cdot 1_A \\ (\text{by the definition of } u)}} \ : \ \lambda \in \mathbf{k} \right\} = \left\{\lambda \cdot 1_A \ : \ \lambda \in \mathbf{k}\right\} = \mathbf{k} \cdot 1_A.$$

Hence, $\mathbf{k} \cdot 1_A = u(\mathbf{k}) \subset A_0$, so that part (a) is solved.

(b) Since A is connected, we have $A_0 \cong \mathbf{k}$. In other words, there exists a $\mathbf{k}$-module isomorphism $\phi : A_0 \to \mathbf{k}$. Consider this ϕ. Since ϕ is a $\mathbf{k}$-module isomorphism $A_0 \to \mathbf{k}$, the inverse ϕ^{-1} of ϕ is a well-defined $\mathbf{k}$-module isomorphism $\mathbf{k} \to A_0$.

We saw in the proof of part (a) that $u(\mathbf{k}) \subset A_0$. Hence, u restricts to a map $\mathbf{k} \to A_0$. Denote this map $\mathbf{k} \to A_0$ by u'. Then, u' is a restriction of u (more precisely, a corestriction of u, because we are restricting the target rather than the domain).

Let also ϵ' denote the restriction of ϵ to A_0. Since ϵ' and u' are restrictions of ϵ and u, we have $\epsilon' \circ u' = \epsilon \circ u = \mathrm{id}_{\mathbf{k}}$ (by the axioms of a coalgebra). Now, $\left(\epsilon' \circ \phi^{-1}\right) \circ (\phi \circ u') = \epsilon' \circ u' = \mathrm{id}_{\mathbf{k}}$. Hence, the $\mathbf{k}$-linear map $\epsilon' \circ \phi^{-1} : \mathbf{k} \to \mathbf{k}$ has a right inverse. Thus, the $\mathbf{k}$-linear map $\epsilon' \circ \phi^{-1} : \mathbf{k} \to \mathbf{k}$ is surjective. Since every surjective $\mathbf{k}$-linear map $\mathbf{k} \to \mathbf{k}$ is an isomorphism[393], this shows that the $\mathbf{k}$-linear map $\epsilon' \circ \phi^{-1} : \mathbf{k} \to \mathbf{k}$ is an isomorphism. Since ϕ also is an isomorphism, the map $\epsilon' \circ \phi^{-1} \circ \phi$ is a composition of two isomorphisms, and thus an isomorphism. In other words, ϵ' is an isomorphism (since $\epsilon' \circ \phi^{-1} \circ \phi = \epsilon'$). Hence, the inverse map of ϵ' is well-defined. This inverse map must be u' (since $\epsilon' \circ u' = \mathrm{id}_{\mathbf{k}}$), and so we conclude that u' is an isomorphism. In other words, the restriction of u to a map $\mathbf{k} \to A_0$ is an isomorphism. This solves part (b).

(c) Part (b) shows that the map $\mathbf{k} \overset{u}{\to} A_0$ is an isomorphism. Hence, this map $\mathbf{k} \overset{u}{\to} A_0$ is bijective, and thus also surjective. In other words, we have $A_0 = u(\mathbf{k})$. Hence,

$$A_0 = u\left(\underbrace{\mathbf{k}}_{=\mathbf{k}\cdot 1}\right) = u(\mathbf{k}\cdot 1) = \mathbf{k} \cdot \underbrace{u(1)}_{\substack{=1\cdot 1_A \\ \text{(by the definition of } u)}} \qquad \text{(since the map } u \text{ is } \mathbf{k}\text{-linear)}$$
$$= \underbrace{\mathbf{k}\cdot 1}_{=\mathbf{k}} \cdot 1_A = \mathbf{k} \cdot 1_A.$$

This proves part (c).

(e) In the proof of part (b), we showed that u' is the inverse map of ϵ'. Hence, ϵ' is the inverse map of u'. In other words, the restriction of ϵ to A_0 is the inverse map of the restriction of u to a map $\mathbf{k} \to A_0$. This solves part (e).

(d) The counit map ϵ is graded (since A is a graded coalgebra). Hence, every $n \geq 0$ satisfies $\epsilon(A_n) \subset \mathbf{k}_n$ (where $\mathbf{k}_n$ denotes the n-th graded component of $\mathbf{k}$). For every positive n, this shows that $A_n \subset \ker \epsilon$ [394]. Hence, $\bigoplus_{n>0} A_n = \sum_{n>0} \underbrace{A_n}_{\subset \ker \epsilon} \subset \sum_{n>0} \ker \epsilon \subset \ker \epsilon$ (since $\ker \epsilon$ is a $\mathbf{k}$-submodule of A).

Now, let $a \in \ker \epsilon$ be arbitrary. Then, $a \in A$ satisfies $\epsilon(a) = 0$. We have $a \in A = \bigoplus_{n \geq 0} A_n = A_0 \oplus \bigoplus_{n>0} A_n$. Hence, we can write a in the form $a = a' + a''$ for $a' \in A_0$ and $a'' \in \bigoplus_{n>0} A_n$. Consider these a' and a''. We have $a'' \in \bigoplus_{n>0} A_n \subset \ker \epsilon$, so that $\epsilon(a'') = 0$. Since $a = a' + a''$, we have $\epsilon(a) = \epsilon(a' + a'') = \epsilon(a') + \underbrace{\epsilon(a'')}_{=0} = \epsilon(a')$, thus $\epsilon(a') = \epsilon(a) = 0$. Since ϵ restricted to A_0 is injective (in fact, part (e) of this problem shows that ϵ restricted to A_0 is an isomorphism), this yields that $a' = 0$ (because $a' \in A_0$). Hence, $a = \underbrace{a'}_{=0} + a'' = a'' \in \bigoplus_{n>0} A_n$.

Now forget that we fixed a. We thus have seen that every $a \in \ker \epsilon$ satisfies $a \in \bigoplus_{n>0} A_n$. In other words, $\ker \epsilon \subset \bigoplus_{n>0} A_n$. Combined with $\bigoplus_{n>0} A_n \subset \ker \epsilon$, this yields $\ker \epsilon = \bigoplus_{n>0} A_n$. This solves part (d).

(f) Let $x \in A$. We have $A = \bigoplus_{n \geq 0} A_n = \underbrace{A_0}_{=\mathbf{k}\cdot 1_A} \oplus \underbrace{\bigoplus_{n>0} A_n}_{=I} = \mathbf{k} \cdot 1_A \oplus I = \mathbf{k} \cdot 1_A + I$ and

$$\Delta(x) \in A \otimes \underbrace{A}_{=\mathbf{k}\cdot 1_A + I} = A \otimes (\mathbf{k} \cdot 1_A + I) = A \otimes (\mathbf{k} \cdot 1_A) + A \otimes I.$$

[393] *Proof.* Let α be a surjective $\mathbf{k}$-linear map $\mathbf{k} \to \mathbf{k}$. We need to show that α is an isomorphism.

Let $\lambda \in \ker \alpha$. Then, $\lambda \in \mathbf{k}$ satisfies $\alpha(\lambda) = 0$. But α is surjective, so that $\mathbf{k} = \alpha(\mathbf{k})$. Hence, $1 \in \mathbf{k} = \alpha(\mathbf{k})$. Thus, there exists a $\mu \in \mathbf{k}$ such that $1 = \alpha(\mu)$. Consider this μ. Then, the $\mathbf{k}$-linearity of α yields $\alpha(\lambda\mu) = \lambda \underbrace{\alpha(\mu)}_{=1} = \lambda$. But the $\mathbf{k}$-linearity of α also shows that $\alpha(\mu\lambda) = \mu \underbrace{\alpha(\lambda)}_{=0} = 0$. Thus, $0 = \alpha(\mu\lambda) = \alpha(\lambda\mu) = \lambda$, so that $\lambda = 0$. We thus have shown that every $\lambda \in \ker \alpha$ satisfies $\lambda = 0$. Hence, $\ker \alpha = 0$, so that the map α is injective. Since α is injective and surjective, we see that α is bijective, thus an isomorphism, qed.

[394] *Proof.* Let n be a positive integer. Then, $\epsilon(A_n) \subset \mathbf{k}_n = 0$ (because of how the grading on $\mathbf{k}$ is constructed), so that $\epsilon(A_n) = 0$ and thus $A_n \subset \ker \epsilon$, qed.

Hence, there exist $y \in A \otimes (\mathbf{k} \cdot 1_A)$ and $z \in A \otimes I$ such that $\Delta(x) = y + z$. Consider these y and z. We will show that $y = x \otimes 1_A$.

Since $y \in A \otimes (\mathbf{k} \cdot 1_A) = \mathbf{k} \cdot (A \otimes 1_A) = A \otimes 1_A$, we can write y in the form $y = y' \otimes 1_A$ for some $y' \in A$. Consider this y'.

By the commutativity of (1.2.2), we have $(\operatorname{id}_A \otimes \epsilon)(\Delta(x)) = x$ (where we identify $A \otimes \mathbf{k}$ with A). Hence,

$$\begin{aligned} x &= (\operatorname{id}_A \otimes \epsilon)\left(\underbrace{\Delta(x)}_{=y+z}\right) = (\operatorname{id}_A \otimes \epsilon)(y+z) = (\operatorname{id}_A \otimes \epsilon)\left(\underbrace{y}_{=y'\otimes 1_A}\right) + (\operatorname{id}_A \otimes \epsilon)\left(\underbrace{z}_{\in A \otimes I}\right) \\ &\in \underbrace{(\operatorname{id}_A \otimes \epsilon)(y' \otimes 1_A)}_{=y'\epsilon(1_A)} + \underbrace{(\operatorname{id}_A \otimes \epsilon)(A \otimes I)}_{=\operatorname{id}_A(A)\epsilon(I)} = y' \underbrace{\epsilon(1_A)}_{=1} + \operatorname{id}_A(A) \underbrace{\epsilon(I)}_{\substack{=0 \\ (\text{since } I = \ker \epsilon \text{ by part (d)})}} \\ &= y'1 + \operatorname{id}_A(A)\,0 = y' + 0 = y', \end{aligned}$$

which shows that $x = y'$. Now, $y = \underbrace{y'}_{=x} \otimes 1_A = x \otimes 1_A$ and

$$x = \underbrace{y}_{=x\otimes 1_A} + \underbrace{z}_{\in A \otimes I} \in x \otimes \underbrace{1_A}_{=1} + A \otimes I = x \otimes 1 + A \otimes I.$$

This solves part (f).

(g) Let $x \in I$. Thus, $x \in I = \ker \epsilon$ (by part (d)), so that $\epsilon(x) = 0$.

Let us introduce y and z as in the solution to part (f). As we saw in that solution, we have $y = x \otimes 1_A$.

We have

$$z \in \underbrace{A}_{=\mathbf{k}\cdot 1_A + I} \otimes I = (\mathbf{k} \cdot 1_A + I) \otimes I = (\mathbf{k} \cdot 1_A) \otimes I + I \otimes I.$$

Hence, there exist $u \in (\mathbf{k} \cdot 1_A) \otimes I$ and $v \in I \otimes I$ such that $z = u + v$. Consider these u and v. We will show that $u = 1_A \otimes x$.

We have $v \in I \otimes I$, so that $(\epsilon \otimes \operatorname{id}_A)(v) \in (\epsilon \otimes \operatorname{id}_A)(I \otimes I) = \underbrace{\epsilon(I)}_{\substack{=0 \\ (\text{since } I = \ker \epsilon)}} \operatorname{id}_A(I) = 0$. Thus, $(\epsilon \otimes \operatorname{id}_A)(v) = 0$.

Since $u \in (\mathbf{k} \cdot 1_A) \otimes I = \mathbf{k} \cdot (1_A \otimes I) = 1_A \otimes I$, we can write u in the form $u = 1_A \otimes u'$ for some $u' \in I$. Consider this u'.

By the commutativity of (1.2.2), we have $(\epsilon \otimes \operatorname{id}_A)(\Delta(x)) = x$ (where we identify $\mathbf{k} \otimes A$ with A). Hence,

$$\begin{aligned} x &= (\epsilon \otimes \operatorname{id}_A)\left(\underbrace{\Delta(x)}_{=y+z}\right) = (\epsilon \otimes \operatorname{id}_A)(y+z) = (\epsilon \otimes \operatorname{id}_A)\left(\underbrace{y}_{=x\otimes 1_A}\right) + (\epsilon \otimes \operatorname{id}_A)\left(\underbrace{z}_{=u+v}\right) \\ &= \underbrace{(\epsilon \otimes \operatorname{id}_A)(x \otimes 1_A)}_{\substack{=\epsilon(x)\operatorname{id}_A(1_A)=0 \\ (\text{since } \epsilon(x)=0)}} + (\epsilon \otimes \operatorname{id}_A)(u+v) = (\epsilon \otimes \operatorname{id}_A)(u+v) = (\epsilon \otimes \operatorname{id}_A)(u) + \underbrace{(\epsilon \otimes \operatorname{id}_A)(v)}_{=0} \\ &= (\epsilon \otimes \operatorname{id}_A)\left(\underbrace{u}_{=1_A \otimes u'}\right) = (\epsilon \otimes \operatorname{id}_A)(1_A \otimes u') = \underbrace{\epsilon(1_A)}_{=1} u' = u'. \end{aligned}$$

Now, $u = 1_A \otimes \underbrace{u'}_{=x} = 1_A \otimes x$. Now,

$$\begin{aligned} \Delta(x) &= \underbrace{y}_{=x\otimes 1_A} + \underbrace{z}_{=u+v} = x \otimes 1_A + \underbrace{u}_{=1_A \otimes x} + \underbrace{v}_{\in I \otimes I} \\ &\in x \otimes \underbrace{1_A}_{=1} + \underbrace{1_A}_{=1} \otimes x + I \otimes I = x \otimes 1 + 1 \otimes x + I \otimes I = 1 \otimes x + x \otimes 1 + I \otimes I. \end{aligned}$$

In other words, $\Delta(x) = 1 \otimes x + x \otimes 1 + \Delta_+(x)$ for some $\Delta_+(x) \in I \otimes I$. This solves part (g).

(h) The definition of I yields

$$I = \bigoplus_{n>0} A_n = \bigoplus_{\ell>0} A_\ell = \sum_{\ell>0} A_\ell. \tag{13.7.1}$$

Now, let $n > 0$ and $x \in A_n$. We must show that $\Delta(x) = 1 \otimes x + x \otimes 1 + \Delta_+(x)$, where $\Delta_+(x)$ lies in $\sum_{k=1}^{n-1} A_k \otimes A_{n-k}$.

Part (g) yields $\Delta(x) = 1 \otimes x + x \otimes 1 + \Delta_+(x)$ for some $\Delta_+(x) \in I \otimes I$. Consider this $\Delta_+(x)$. It clearly suffices to show that this $\Delta_+(x)$ satisfies $\Delta_+(x) \in \sum_{k=1}^{n-1} A_k \otimes A_{n-k}$.

Regard $I \otimes I$ as a $\mathbf{k}$-submodule of $A \otimes A$. From $I = \sum_{\ell>0} A_\ell$, we obtain

$$I \otimes I = \left(\sum_{\ell>0} A_\ell\right) \otimes \left(\sum_{\ell>0} A_\ell\right) = \sum_{i>0,\ j>0} A_i \otimes A_j. \tag{13.7.2}$$

Let $\pi_n : A \otimes A \to A \otimes A$ be the projection of the graded $\mathbf{k}$-module $A \otimes A$ onto its n-th graded component $(A \otimes A)_n$. Then:

- The map π_n annihilates the k-th graded component $(A \otimes A)_k$ for every $k \neq n$. In other words,

$$\pi_n\left((A \otimes A)_k\right) = 0 \qquad \text{for every } k \in \mathbb{N} \text{ satisfying } k \neq n. \tag{13.7.3}$$

  Hence, every $i \in \mathbb{N}$ and $j \in \mathbb{N}$ satisfying $i + j \neq n$ satisfy

$$\pi_n\left(A_i \otimes A_j\right) = 0 \tag{13.7.4}$$

  [395].

- The map π_n acts as the identity on the n-th graded component $(A \otimes A)_n$. In other words,

$$\pi_n(z) = z \qquad \text{for each } z \in (A \otimes A)_n. \tag{13.7.5}$$

  Therefore, every $i \in \mathbb{N}$ and $j \in \mathbb{N}$ satisfying $i + j = n$ satisfy

$$\pi_n\left(A_i \otimes A_j\right) = A_i \otimes A_j \tag{13.7.6}$$

  [396].

Now, $x \in A_n$, so that $\Delta(x) \in \Delta(A_n) \subset (A \otimes A)_n$ (since the map Δ is graded). Also, $\underbrace{1}_{\in A_0} \otimes \underbrace{x}_{\in A_n} \in A_0 \otimes A_n \subset (A \otimes A)_{0+n}$ (by the definition of the grading on $A \otimes A$); this rewrites as $1 \otimes x \in (A \otimes A)_n$. Similarly, $x \otimes 1 \in (A \otimes A)_n$. From $\Delta(x) = 1 \otimes x + x \otimes 1 + \Delta_+(x)$, we obtain

$$\Delta_+(x) = \underbrace{\Delta(x)}_{\in (A \otimes A)_n} - \underbrace{1 \otimes x}_{\in (A \otimes A)_n} - \underbrace{x \otimes 1}_{\in (A \otimes A)_n} \in (A \otimes A)_n - (A \otimes A)_n - (A \otimes A)_n \subset (A \otimes A)_n.$$

[395] *Proof of (13.7.4):* Let $i \in \mathbb{N}$ and $j \in \mathbb{N}$ be such that $i + j \neq n$. The definition of the grading on $A \otimes A$ yields $A_i \otimes A_j \subset (A \otimes A)_{i+j}$. Applying the map π_n to both sides of this relation, we find

$$\pi_n\left(A_i \otimes A_j\right) \subset \pi_n\left((A \otimes A)_{i+j}\right) = 0$$

(by (13.7.3), applied to $k = i + j$). Hence, $\pi_n(A_i \otimes A_j) = 0$. This proves (13.7.4).

[396] *Proof of (13.7.6):* Let $i \in \mathbb{N}$ and $j \in \mathbb{N}$ be such that $i + j = n$. The definition of the grading on $A \otimes A$ yields $A_i \otimes A_j \subset (A \otimes A)_{i+j} = (A \otimes A)_n$ (since $i + j = n$). Hence, each $z \in A_i \otimes A_j$ satisfies $z \in (A \otimes A)_n$ and therefore $\pi_n(z) = z$ (by (13.7.5)). In other words, the map π_n acts as the identity on the set $A_i \otimes A_j$. Therefore, $\pi_n(A_i \otimes A_j) = A_i \otimes A_j$. This proves (13.7.6).

Hence, (13.7.5) (applied to $z = \Delta_+(x)$) yields $\pi_n(\Delta_+(x)) = \Delta_+(x)$. Thus,

$$\begin{aligned}
\Delta_+(x) &= \pi_n\left(\underbrace{\Delta_+(x)}_{\in I\otimes I}\right) \in \pi_n(I\otimes I) = \pi_n\left(\sum_{i>0,\ j>0} A_i \otimes A_j\right) \qquad \text{(by (13.7.2))} \\
&= \sum_{i>0,\ j>0} \pi_n(A_i\otimes A_j) \qquad \text{(since the map } \pi_n \text{ is } \mathbf{k}\text{-linear)} \\
&= \sum_{\substack{i>0,\ j>0;\\ i+j=n}} \underbrace{\pi_n(A_i\otimes A_j)}_{\substack{=A_i\otimes A_j\\ \text{(by (13.7.6))}}} + \sum_{\substack{i>0,\ j>0;\\ i+j\neq n}} \underbrace{\pi_n(A_i\otimes A_j)}_{\substack{=0\\ \text{(by (13.7.4))}}} \\
&= \sum_{\substack{i>0,\ j>0;\\ i+j=n}} A_i\otimes A_j + \underbrace{\sum_{\substack{i>0,\ j>0;\\ i+j\neq n}} 0}_{=0} = \sum_{\substack{i>0,\ j>0;\\ i+j=n}} A_i\otimes A_j = \sum_{k=1}^{n-1} A_k\otimes A_{n-k}.
\end{aligned}$$

We thus have shown that $\Delta(x) = 1\otimes x + x\otimes 1 + \Delta_+(x)$, where $\Delta_+(x)$ lies in $\sum_{k=1}^{n-1} A_k\otimes A_{n-k}$. This solves Exercise 1.3.20(h).

13.8. **Solution to Exercise 1.3.24.** *Solution to Exercise 1.3.24.*

It is easy to solve Exercise 1.3.24 by direct verification of all the axioms on homogeneous elements. However, we shall take a different (more "functorial") route instead. The first step on that route is to generalize the definition of D_q given in the exercise:

Definition 13.8.1. Fix $q \in \mathbf{k}$. Let $V = \bigoplus_{n\in\mathbb{N}} V_n$ be a graded $\mathbf{k}$-module (where the V_n are the homogeneous components of V). Let $D_{q,V} : V \to V$ be the $\mathbf{k}$-module endomorphism of V defined by setting

$$D_{q,V}(a) = q^n a \qquad \text{for each } n\in\mathbb{N} \text{ and each } a \in V_n.$$

(It is easy to see that this is well-defined; equivalently, $D_{q,V}$ could be defined as the direct sum $\bigoplus_{n\in\mathbb{N}} (q^n \cdot \mathrm{id}_{V_n}) : \bigoplus_{n\in\mathbb{N}} V_n \to \bigoplus_{n\in\mathbb{N}} V_n$ of the maps $q^n \cdot \mathrm{id}_{V_n} : V_n \to V_n$.)

Let us now state a few basic properties of these endomorphisms $D_{q,V}$:

Proposition 13.8.2. *Fix $q\in\mathbf{k}$. Let V and W be two graded $\mathbf{k}$-modules. Let $f : V \to W$ be a graded $\mathbf{k}$-linear map. Then, $D_{q,W}\circ f = f\circ D_{q,V}$.*

Proof of Proposition 13.8.2. Write V as $V = \bigoplus_{n\in\mathbb{N}} V_n$ (where the V_n are the homogeneous components of V). Write W as $W = \bigoplus_{n\in\mathbb{N}} W_n$ (where the W_n are the homogeneous components of W).

Every element of V is a $\mathbf{k}$-linear combination of homogeneous elements of V (since V is graded). Thus, if two $\mathbf{k}$-linear maps from V agree on each homogeneous element of V, then these two maps must be equal.

Let v be a homogeneous element of V. Thus, there exists some $n\in\mathbb{N}$ such that $v \in V_n$. Consider this n. From $v\in V_n$, we obtain $D_{q,V}(v) = q^n v$ (by the definition of $D_{q,V}$).

But from $v \in V_n$, we also obtain $f(v) \in f(V_n) \subset W_n$ (since the map f is graded).

But the definition of $D_{q,W}$ shows that $D_{q,W}(a) = q^n a$ for each $a\in W_n$. Applying this to $a = f(v)$, we find $D_{q,W}(f(v)) = q^n f(v)$ (since $f(v)\in W_n$). Now, comparing

$$(D_{q,W}\circ f)(v) = D_{q,W}(f(v)) = q^n f(v)$$

with

$$(f\circ D_{q,V})(v) = f\left(\underbrace{D_{q,V}(v)}_{=q^n v}\right) = f(q^n v) = q^n f(v) \qquad \text{(since the map } f \text{ is } \mathbf{k}\text{-linear)},$$

we obtain $(D_{q,W}\circ f)(v) = (f\circ D_{q,V})(v)$.

Let us forget that we fixed v. We thus have proved that $(D_{q,W}\circ f)(v) = (f\circ D_{q,V})(v)$ for each homogeneous element v of V. In other words, the two maps $D_{q,W}\circ f$ and $f\circ D_{q,V}$ (from V to W) agree on each homogeneous element of V. Since these two maps are $\mathbf{k}$-linear, we can thus conclude that they must be equal (because if two $\mathbf{k}$-linear maps from V agree on each homogeneous element of V, then these two maps must be equal). In other words, we have $D_{q,W} \circ f = f \circ D_{q,V}$. This proves Proposition 13.8.2. □

Proposition 13.8.3. *Fix $q \in \mathbf{k}$. Let V and W be two graded $\mathbf{k}$-modules. Then, $D_{q,V\otimes W} = D_{q,V} \otimes D_{q,W}$.*

Proof of Proposition 13.8.3. Let us first show the following:

Claim 1: Let $v \in V$ and $w \in W$. Then, $D_{q,V\otimes W}(v \otimes w) = (D_{q,V} \otimes D_{q,W})(v \otimes w)$.

[*Proof of Claim 1:* We must prove the equality $D_{q,V\otimes W}(v \otimes w) = (D_{q,V} \otimes D_{q,W})(v \otimes w)$. This equality is clearly $\mathbf{k}$-linear in v (that is, both of its sides depend $\mathbf{k}$-linearly on v). Thus, we can WLOG assume that v is a homogeneous element of V (since v is always a $\mathbf{k}$-linear combination of homogeneous elements of V (because V is graded)). Assume this.

For the same reason, we can WLOG assume that w is a homogeneous element of W. Assume this.

The element $v \in V$ is homogeneous. Thus, there exists some $i \in \mathbb{N}$ such that $v \in V_i$. Consider this i. From $v \in V_i$, we obtain $D_{q,V}(v) = q^i v$ (by the definition of $D_{q,V}$).

The element $w \in W$ is homogeneous. Thus, there exists some $j \in \mathbb{N}$ such that $w \in W_j$. Consider this j. From $w \in W_j$, we obtain $D_{q,W}(w) = q^j w$ (by the definition of $D_{q,W}$).

The definition of the grading on $V \otimes W$ yields $V_i \otimes W_j \subset (V \otimes W)_{i+j}$. Hence,

$$\underbrace{v}_{\in V_i} \otimes \underbrace{w}_{\in W_j} \in V_i \otimes W_j \subset (V \otimes W)_{i+j}.$$

But the definition of $D_{q,V\otimes W}$ shows that $D_{q,V\otimes W}(a) = q^n a$ for each $n \in \mathbb{N}$ and each $a \in (V \otimes W)_n$. Applying this to $n = i+j$ and $a = v \otimes w$, we obtain

$$\begin{aligned} D_{q,V\otimes W}(v \otimes w) &= \underbrace{q^{i+j}}_{=q^i q^j}(v \otimes w) \qquad \left(\text{since } v \otimes w \in (V \otimes W)_{i+j}\right) \\ &= q^i q^j (v \otimes w) = q^i v \otimes q^j w = (D_{q,V} \otimes D_{q,W})(v \otimes w) \\ &\qquad \left(\text{since } (D_{q,V} \otimes D_{q,W})(v \otimes w) = \underbrace{D_{q,V}(v)}_{=q^i v} \otimes \underbrace{D_{q,W}(w)}_{=q^j w} = q^i v \otimes q^j w\right). \end{aligned}$$

This proves Claim 1.]

Now, recall that the $\mathbf{k}$-module $V \otimes W$ is spanned by the pure tensors. Thus, if two $\mathbf{k}$-linear maps from $V \otimes W$ agree on each pure tensor, then these two maps must be equal.

But Claim 1 shows precisely that the two maps $D_{q,V\otimes W}$ and $D_{q,V} \otimes D_{q,W}$ (from $V \otimes W$ to $V \otimes W$) agree on each pure tensor. Since these two maps are $\mathbf{k}$-linear, we can thus conclude that these two maps must be equal (because if two $\mathbf{k}$-linear maps from $V \otimes W$ agree on each pure tensor, then these two maps must be equal). In other words, $D_{q,V\otimes W} = D_{q,V} \otimes D_{q,W}$. This proves Proposition 13.8.3. □

We are now ready to solve Exercise 1.3.24:

In Definition 13.8.1, we have defined a map $D_{q,V}$ for any graded $\mathbf{k}$-module V. Applying this to $V = A$, we obtain a map $D_{q,A}$. This map $D_{q,A}$ equals our map D_q (because these two maps are defined in the exact same way). In other words, we have $D_{q,A} = D_q$.

Let m, u, Δ, ϵ be the structure maps of the bialgebra A, labelled as usual (so $m : A \otimes A \to A$ is the multiplication map, $u : \mathbf{k} \to A$ is the unit, $\Delta : A \to A \otimes A$ is the comultiplication, and $\epsilon : A \to \mathbf{k}$ is the counit). These structure maps m, u, Δ, ϵ are graded (since A is a graded bialgebra).

The map D_q is a $\mathbf{k}$-algebra homomorphism if and only if it is $\mathbf{k}$-linear and makes the diagrams

$$\begin{tikzcd} A \arrow[r, "D_q"] & A \\ A \otimes A \arrow[u, "m"] \arrow[r, "D_q \otimes D_q"] & A \otimes A \arrow[u, "m"'] \end{tikzcd} \qquad \begin{tikzcd} A \arrow[rr, "D_q"] & & A \\ & \mathbf{k} \arrow[ul, "u"] \arrow[ur, "u"'] & \end{tikzcd}$$

commute (by the definition of a $\mathbf{k}$-algebra homomorphism). In other words, the map D_q is a $\mathbf{k}$-algebra homomorphism if and only if it is $\mathbf{k}$-linear and satisfies $D_q \circ m = m \circ (D_q \otimes D_q)$ and $D_q \circ u = u$.

But Proposition 13.8.3 (applied to $V = A$ and $W = A$) yields $D_{q,A\otimes A} = D_{q,A} \otimes D_{q,A}$.

The map $m : A \otimes A \to A$ is graded and $\mathbf{k}$-linear; hence, Proposition 13.8.2 (applied to $V = A \otimes A$, $W = A$ and $f = m$) yields $D_{q,A} \circ m = m \circ D_{q,A\otimes A}$. In view of $D_{q,A\otimes A} = D_{q,A} \otimes D_{q,A}$, this rewrites as $D_{q,A} \circ m = m \circ (D_{q,A} \otimes D_{q,A})$. In other words, $D_q \circ m = m \circ (D_q \otimes D_q)$ (since $D_{q,A} = D_q$).

Also, the graded $\mathbf{k}$-module $\mathbf{k}$ satisfies $D_{q,\mathbf{k}} = \operatorname{id}_{\mathbf{k}}$ [397].

The map $u : \mathbf{k} \to A$ is graded and $\mathbf{k}$-linear; hence, Proposition 13.8.2 (applied to $V = \mathbf{k}$, $W = A$ and $f = u$) yields $D_{q,A} \circ u = u \circ \underbrace{D_{q,\mathbf{k}}}_{=\operatorname{id}_{\mathbf{k}}} = u$. In other words, $D_q \circ u = u$ (since $D_{q,A} = D_q$).

Recall that the map D_q is a $\mathbf{k}$-algebra homomorphism if and only if it is $\mathbf{k}$-linear and satisfies $D_q \circ m = m \circ (D_q \otimes D_q)$ and $D_q \circ u = u$. Thus, the map D_q is a $\mathbf{k}$-algebra homomorphism (since it is $\mathbf{k}$-linear and satisfies $D_q \circ m = m \circ (D_q \otimes D_q)$ and $D_q \circ u = u$).

The map D_q is a $\mathbf{k}$-coalgebra homomorphism if and only if it is $\mathbf{k}$-linear and makes the diagrams

$$\begin{tikzcd}
A \arrow[r, "D_q"] \arrow[d, "\Delta"'] & A \arrow[d, "\Delta"] \\
A \otimes A \arrow[r, "D_q \otimes D_q"] & A \otimes A
\end{tikzcd}
\qquad
\begin{tikzcd}
A \arrow[rr, "D_q"] \arrow[dr, "\epsilon"'] & & A \arrow[dl, "\epsilon"] \\
& \mathbf{k} &
\end{tikzcd}$$

commute (by the definition of a $\mathbf{k}$-coalgebra homomorphism). In other words, the map D_q is a $\mathbf{k}$-coalgebra homomorphism if and only if it is $\mathbf{k}$-linear and satisfies $(D_q \otimes D_q) \circ \Delta = \Delta \circ D_q$ and $\epsilon \circ D_q = \epsilon$.

The map $\Delta : A \to A \otimes A$ is graded and $\mathbf{k}$-linear; hence, Proposition 13.8.2 (applied to $V = A$, $W = A \otimes A$ and $f = \Delta$) yields $D_{q,A\otimes A} \circ \Delta = \Delta \circ D_{q,A}$. In view of $D_{q,A\otimes A} = D_{q,A} \otimes D_{q,A}$, this rewrites as $(D_{q,A} \otimes D_{q,A}) \circ \Delta = \Delta \circ D_{q,A}$. In other words, $(D_q \otimes D_q) \circ \Delta = \Delta \circ D_q$ (since $D_{q,A} = D_q$).

The map $\epsilon : A \to \mathbf{k}$ is graded and $\mathbf{k}$-linear; hence, Proposition 13.8.2 (applied to $V = A$, $W = \mathbf{k}$ and $f = \epsilon$) yields $D_{q,\mathbf{k}} \circ \epsilon = \epsilon \circ \underbrace{D_{q,A}}_{=D_q} = \epsilon \circ D_q$. Hence, $\epsilon \circ D_q = \underbrace{D_{q,\mathbf{k}}}_{=\operatorname{id}_{\mathbf{k}}} \circ \epsilon = \epsilon$.

Recall that the map D_q is a $\mathbf{k}$-coalgebra homomorphism if and only if it is $\mathbf{k}$-linear and satisfies $(D_q \otimes D_q) \circ \Delta = \Delta \circ D_q$ and $\epsilon \circ D_q = \epsilon$. Hence, the map D_q is a $\mathbf{k}$-coalgebra homomorphism (since it is $\mathbf{k}$-linear and satisfies $(D_q \otimes D_q) \circ \Delta = \Delta \circ D_q$ and $\epsilon \circ D_q = \epsilon$).

Now, we know that the map D_q is a $\mathbf{k}$-algebra homomorphism and a $\mathbf{k}$-coalgebra homomorphism. In other words, D_q is a $\mathbf{k}$-bialgebra homomorphism (by the definition of a $\mathbf{k}$-bialgebra homomorphism). This solves Exercise 1.3.24.

13.9. **Solution to Exercise 1.3.26.** *Solution to Exercise 1.3.26.* (a)

Proof of Proposition 1.3.25. Let θ be the canonical $\mathbf{k}$-module isomorphism $\mathbf{k} \otimes \mathbf{k} \to \mathbf{k}$ (sending each $\lambda \otimes \mu$ to $\lambda\mu$). Thus, θ is a $\mathbf{k}$-algebra isomorphism. Our definition of the $\mathbf{k}$-coalgebra $A \otimes B$ yields

$$\Delta_{A\otimes B} = (\operatorname{id}_A \otimes T \otimes \operatorname{id}_B) \circ (\Delta_A \otimes \Delta_B) \qquad \text{and} \tag{13.9.1}$$

$$\epsilon_{A\otimes B} = \theta \circ (\epsilon_A \otimes \epsilon_B) . \tag{13.9.2}$$

But recall that A is a $\mathbf{k}$-bialgebra. Thus, the maps $\Delta_A : A \to A \otimes A$ and $\epsilon_A : A \to \mathbf{k}$ are $\mathbf{k}$-algebra homomorphisms (by the definition of a bialgebra). The same argument (applied to B instead of A) shows that the maps $\Delta_B : B \to B \otimes B$ and $\epsilon_B : B \to \mathbf{k}$ are $\mathbf{k}$-algebra homomorphisms.

Now, Exercise 1.3.6(a) (applied to $A' = A \otimes A$, $B' = B \otimes B$, $f = \Delta_A$ and $g = \Delta_B$) shows that $\Delta_A \otimes \Delta_B : A \otimes B \to A \otimes A \otimes B \otimes B$ is a $\mathbf{k}$-algebra homomorphism.

Also, it is easy to show that if $\mathfrak{A}$ and $\mathfrak{B}$ are any two $\mathbf{k}$-algebras, then the map $T : \mathfrak{A} \otimes \mathfrak{B} \to \mathfrak{B} \otimes \mathfrak{A}$ is a $\mathbf{k}$-algebra homomorphism. Applying this to $\mathfrak{A} = A$ and $\mathfrak{B} = B$, we conclude that the map $T : A \otimes B \to B \otimes A$ is

[397] *Proof.* Let $\lambda \in \mathbf{k}$. Then, $\lambda \in \mathbf{k} = \mathbf{k}_0$ (by the definition of the grading on $\mathbf{k}$). But the definition of $D_{q,\mathbf{k}}$ yields

$$D_{q,\mathbf{k}}(a) = q^n a \qquad \text{for each } n \in \mathbb{N} \text{ and each } a \in \mathbf{k}_n .$$

Applying this to $n = 0$ and $a = \lambda$, we obtain $D_{q,\mathbf{k}}(\lambda) = q^0 \lambda$ (since $\lambda \in \mathbf{k}_0$). Hence, $D_{q,\mathbf{k}}(\lambda) = \underbrace{q^0}_{=1} \lambda = \lambda = \operatorname{id}_{\mathbf{k}}(\lambda)$.

Forget that we fixed λ. We thus have shown that $D_{q,\mathbf{k}}(\lambda) = \operatorname{id}_{\mathbf{k}}(\lambda)$ for each $\lambda \in \mathbf{k}$. In other words, $D_{q,\mathbf{k}} = \operatorname{id}_{\mathbf{k}}$.

a $\mathbf{k}$-algebra homomorphism. Also, the maps $\operatorname{id}_A : A \to A$ and $\operatorname{id}_B : B \to B$ are $\mathbf{k}$-algebra homomorphisms. Now, Exercise 1.3.6(a) (applied to A, A, $A \otimes B$, $B \otimes A$, id_A and T instead of A, A', B, B', f and g) shows that $\operatorname{id}_A \otimes T : A \otimes A \otimes B \to A \otimes B \otimes A$ is a $\mathbf{k}$-algebra homomorphism. Hence, Exercise 1.3.6(a) (applied to $A \otimes A \otimes B$, $A \otimes B \otimes A$, B, B, $\operatorname{id}_A \otimes T$ and id_B instead of A, A', B, B', f and g) shows that $\operatorname{id}_A \otimes T \otimes \operatorname{id}_B : A \otimes A \otimes B \otimes B \to A \otimes B \otimes A \otimes B$ is a $\mathbf{k}$-algebra homomorphism.

We now know that the two maps $\Delta_A \otimes \Delta_B : A \otimes B \to A \otimes A \otimes B \otimes B$ and $\operatorname{id}_A \otimes T \otimes \operatorname{id}_B : A \otimes A \otimes B \otimes B \to A \otimes B \otimes A \otimes B$ are $\mathbf{k}$-algebra homomorphisms. Hence, their composition $(\operatorname{id}_A \otimes T \otimes \operatorname{id}_B) \circ (\Delta_A \otimes \Delta_B)$ must also be a $\mathbf{k}$-algebra homomorphism. In light of (13.9.1), this rewrites as follows: The map $\Delta_{A \otimes B}$ is a $\mathbf{k}$-algebra homomorphism.

Furthermore, Exercise 1.3.6(a) (applied to $A' = \mathbf{k}$, $B' = \mathbf{k}$, $f = \epsilon_A$ and $g = \epsilon_B$) shows that $\epsilon_A \otimes \epsilon_B : A \otimes B \to \mathbf{k} \otimes \mathbf{k}$ is a $\mathbf{k}$-algebra homomorphism. We now know that the two maps $\epsilon_A \otimes \epsilon_B : A \otimes B \to \mathbf{k} \otimes \mathbf{k}$ and $\theta : \mathbf{k} \otimes \mathbf{k} \to \mathbf{k}$ are $\mathbf{k}$-algebra homomorphisms. Hence, their composition $\theta \circ (\epsilon_A \otimes \epsilon_B)$ must also be a $\mathbf{k}$-algebra homomorphism. In light of (13.9.2), this rewrites as follows: The map $\epsilon_{A \otimes B}$ is a $\mathbf{k}$-algebra homomorphism.

Now, we know that the maps $\Delta_{A \otimes B}$ and $\epsilon_{A \otimes B}$ are $\mathbf{k}$-algebra homomorphisms. In other words, $A \otimes B$ is a $\mathbf{k}$-bialgebra (by the definition of a bialgebra). This proves Proposition 1.3.25. □

Hence, Exercise 1.3.26(a) is solved.

(b) We know that $(t_g)_{g \in G}$ is a basis of the $\mathbf{k}$-module $\mathbf{k}G$, whereas $(t_h)_{h \in H}$ is a basis of the $\mathbf{k}$-module $\mathbf{k}H$. Thus, $(t_g \otimes t_h)_{(g,h) \in G \times H}$ is a basis of the $\mathbf{k}$-module $\mathbf{k}G \otimes \mathbf{k}H$. Hence, we can define a $\mathbf{k}$-linear map $\Phi : \mathbf{k}G \otimes \mathbf{k}H \to \mathbf{k}[G \times H]$ by setting

$$\left(\Phi(t_g \otimes t_h) = t_{(g,h)} \qquad \text{for each } (g,h) \in G \times H\right). \tag{13.9.3}$$

Consider this map Φ.

The family $\left(t_{(g,h)}\right)_{(g,h) \in G \times H}$ is a basis of the $\mathbf{k}$-module $\mathbf{k}[G \times H]$. Now, the map Φ is $\mathbf{k}$-linear, and sends the basis $(t_g \otimes t_h)_{(g,h) \in G \times H}$ of $\mathbf{k}G \otimes \mathbf{k}H$ to the basis $\left(t_{(g,h)}\right)_{(g,h) \in G \times H}$ of $\mathbf{k}[G \times H]$ (by (13.9.3)). Therefore, this map Φ is an isomorphism of $\mathbf{k}$-modules. Thus, in particular, the map Φ is invertible.

It is easy to see that

$$\Phi(ab) = \Phi(a)\Phi(b) \qquad \text{for every } a \in \mathbf{k}G \otimes \mathbf{k}H \text{ and } b \in \mathbf{k}G \otimes \mathbf{k}H \tag{13.9.4}$$

[398]. It is also easy to see that $\Phi(1_{\mathbf{k}G \otimes \mathbf{k}H}) = 1_{\mathbf{k}[G \times H]}$ (since $1_{\mathbf{k}G \otimes \mathbf{k}H} = 1_{\mathbf{k}G} \otimes 1_{\mathbf{k}H} = t_{1_G} \otimes t_{1_H}$ and $1_{\mathbf{k}[G \times H]} = t_{1_{G \times H}} = t_{(1_G, 1_H)}$). Combining this with (13.9.4), we conclude that Φ is a $\mathbf{k}$-algebra homomorphism (since the map Φ is $\mathbf{k}$-linear). Hence, Φ is a $\mathbf{k}$-algebra isomorphism (since Φ is invertible).

Furthermore,

$$\Delta_{\mathbf{k}[G \times H]} \circ \Phi = (\Phi \otimes \Phi) \circ \Delta_{\mathbf{k}G \otimes \mathbf{k}H} \tag{13.9.5}$$

[398] *Proof of (13.9.4):* Let $a \in \mathbf{k}G \otimes \mathbf{k}H$ and $b \in \mathbf{k}G \otimes \mathbf{k}H$ be arbitrary. We must prove the equality $\Phi(ab) = \Phi(a)\Phi(b)$.

Since this equality is $\mathbf{k}$-linear in a, we can WLOG assume that a belongs to the basis $(t_g \otimes t_h)_{(g,h) \in G \times H}$ of the $\mathbf{k}$-module $\mathbf{k}G \otimes \mathbf{k}H$. Assume this. Thus, $a = t_{g_1} \otimes t_{h_1}$ for some $(g_1, h_1) \in G \times H$. Consider this (g_1, h_1).

Since the equality $\Phi(ab) = \Phi(a)\Phi(b)$ is $\mathbf{k}$-linear in b, we can WLOG assume that b belongs to the basis $(t_g \otimes t_h)_{(g,h) \in G \times H}$ of the $\mathbf{k}$-module $\mathbf{k}G \otimes \mathbf{k}H$. Assume this. Thus, $b = t_{g_2} \otimes t_{h_2}$ for some $(g_2, h_2) \in G \times H$. Consider this (g_2, h_2).

Multiplying the equalities $a = t_{g_1} \otimes t_{h_1}$ and $b = t_{g_2} \otimes t_{h_2}$, we obtain

$$ab = (t_{g_1} \otimes t_{h_1})(t_{g_2} \otimes t_{h_2}) = \underbrace{t_{g_1} t_{g_2}}_{= t_{g_1 g_2}} \otimes \underbrace{t_{h_1} t_{h_2}}_{= t_{h_1 h_2}} = t_{g_1 g_2} \otimes t_{h_1 h_2}.$$

Applying the map Φ to both sides of this equality, we find

$$\Phi(ab) = \Phi(t_{g_1 g_2} \otimes t_{h_1 h_2}) = t_{(g_1 g_2, h_1 h_2)} \qquad \text{(by the definition of } \Phi\text{)}.$$

Comparing this with

$$\begin{aligned} \Phi\left(\underbrace{a}_{= t_{g_1} \otimes t_{h_1}}\right) \Phi\left(\underbrace{b}_{= t_{g_2} \otimes t_{h_2}}\right) &= \underbrace{\Phi(t_{g_1} \otimes t_{h_1})}_{\substack{= t_{(g_1, h_1)} \\ \text{(by the definition of } \Phi\text{)}}} \underbrace{\Phi(t_{g_2} \otimes t_{h_2})}_{\substack{= t_{(g_2, h_2)} \\ \text{(by the definition of } \Phi\text{)}}} \\ &= t_{(g_1, h_1)} t_{(g_2, h_2)} = t_{(g_1, h_1)(g_2, h_2)} = t_{(g_1 g_2, h_1 h_2)} \qquad \left(\text{since } (g_1, h_1)(g_2, h_2) = (g_1 g_2, h_1 h_2)\right), \end{aligned}$$

we obtain $\Phi(ab) = \Phi(a)\Phi(b)$. This proves (13.9.4).

[399]. A similar argument shows that

$$\epsilon_{\mathbf{k}[G\times H]} \circ \Phi = \epsilon_{\mathbf{k}G\otimes\mathbf{k}H}.$$

Combining this with (13.9.5), we conclude that Φ is a $\mathbf{k}$-coalgebra homomorphism (since the map Φ is $\mathbf{k}$-linear). Hence, Φ is a $\mathbf{k}$-coalgebra isomorphism (since Φ is invertible).

Thus, the map Φ is both a $\mathbf{k}$-algebra isomorphism and a $\mathbf{k}$-coalgebra isomorphism. Hence, Φ is a $\mathbf{k}$-bialgebra isomorphism. Therefore, the $\mathbf{k}$-bialgebra $\mathbf{k}G \otimes \mathbf{k}H$ is isomorphic to the $\mathbf{k}$-bialgebra $\mathbf{k}[G\times H]$ (through the map Φ). This solves Exercise 1.3.26(b).

13.10. Solution to Exercise 1.4.2.

Solution to Exercise 1.4.2. We have to prove that the binary operation $\star$ on $\operatorname{Hom}(C, A)$ is associative. In other words, we have to prove that any three elements f, g and h of $\operatorname{Hom}(C, A)$ satisfy $f \star (g \star h) = (f \star g) \star h$.

So let f, g and h be three elements of $\operatorname{Hom}(C, A)$. As usual, denote by $m : A \otimes A \to A$ the multiplication of A, and by $\Delta : C \to C \otimes C$ the comultiplication of C.

By the definition of $g \star h$, we have $g \star h = m \circ (g \otimes h) \circ \Delta$, so that

$$\begin{aligned}\underbrace{f}_{=\operatorname{id}_A \circ f \circ \operatorname{id}_C} \otimes \underbrace{(g \star h)}_{=m\circ(g\otimes h)\circ\Delta} &= (\operatorname{id}_A \circ f \circ \operatorname{id}_C) \otimes (m \circ (g \otimes h) \circ \Delta)\\ &= (\operatorname{id}_A \otimes m) \circ (f \otimes (g \otimes h)) \circ (\operatorname{id}_C \otimes \Delta)\\ &= (\operatorname{id}_A \otimes m) \circ (f \otimes g \otimes h) \circ (\operatorname{id}_C \otimes \Delta).\end{aligned}$$

[399] *Proof of (13.9.5):* Let $(g, h) \in G \times H$. Then,

$$(\Delta_{\mathbf{k}[G\times H]} \circ \Phi)(t_g \otimes t_h) = \Delta_{\mathbf{k}[G\times H]}\left(\underbrace{\Phi(t_g \otimes t_h)}_{\substack{=t_{(g,h)}\\ \text{(by the definition of }\Phi)}}\right) = \Delta_{\mathbf{k}[G\times H]}\left(t_{(g,h)}\right) = t_{(g,h)} \otimes t_{(g,h)} \tag{13.9.6}$$

(by the definition of the coalgebra structure on $\mathbf{k}[G\times H]$).

On the other hand, the definition of the coalgebra structure on $\mathbf{k}G$ shows that $\Delta_{\mathbf{k}G}(t_g) = t_g \otimes t_g$. Similarly, $\Delta_{\mathbf{k}H}(t_h) = t_h \otimes t_h$. But the definition of the coalgebra $\mathbf{k}G \otimes \mathbf{k}H$ shows that $\Delta_{\mathbf{k}G\otimes\mathbf{k}H} = (\operatorname{id}_{\mathbf{k}G} \otimes T \otimes \operatorname{id}_{\mathbf{k}H}) \circ (\Delta_{\mathbf{k}G} \otimes \Delta_{\mathbf{k}H})$. Applying both sides of this equality to $t_g \otimes t_h$, we find

$$\begin{aligned}\Delta_{\mathbf{k}G\otimes\mathbf{k}H}(t_g \otimes t_h) &= ((\operatorname{id}_{\mathbf{k}G} \otimes T \otimes \operatorname{id}_{\mathbf{k}H}) \circ (\Delta_{\mathbf{k}G} \otimes \Delta_{\mathbf{k}H}))(t_g \otimes t_h) = (\operatorname{id}_{\mathbf{k}G} \otimes T \otimes \operatorname{id}_{\mathbf{k}H})\left(\underbrace{(\Delta_{\mathbf{k}G} \otimes \Delta_{\mathbf{k}H})(t_g \otimes t_h)}_{=\Delta_{\mathbf{k}G}(t_g)\otimes\Delta_{\mathbf{k}H}(t_h)}\right)\\ &= (\operatorname{id}_{\mathbf{k}G} \otimes T \otimes \operatorname{id}_{\mathbf{k}H})\left(\underbrace{\Delta_{\mathbf{k}G}(t_g)}_{=t_g\otimes t_g} \otimes \underbrace{\Delta_{\mathbf{k}H}(t_h)}_{=t_h\otimes t_h}\right) = (\operatorname{id}_{\mathbf{k}G} \otimes T \otimes \operatorname{id}_{\mathbf{k}H})(t_g \otimes t_g \otimes t_h \otimes t_h)\\ &= \underbrace{\operatorname{id}_{\mathbf{k}G}(t_g)}_{=t_g} \otimes \underbrace{T(t_g \otimes t_h)}_{\substack{=t_h\otimes t_g\\ \text{(by the definition of }T)}} \otimes \underbrace{\operatorname{id}_{\mathbf{k}H}(t_h)}_{=t_h} = t_g \otimes t_h \otimes t_g \otimes t_h.\end{aligned}$$

Applying the map $\Phi \otimes \Phi$ to both sides of this equality, we find

$$\begin{aligned}(\Phi \otimes \Phi)(\Delta_{\mathbf{k}G\otimes\mathbf{k}H}(t_g \otimes t_h)) &= (\Phi \otimes \Phi)(t_g \otimes t_h \otimes t_g \otimes t_h) = \underbrace{\Phi(t_g \otimes t_h)}_{\substack{=t_{(g,h)}\\ \text{(by the definition of }\Phi)}} \otimes \underbrace{\Phi(t_g \otimes t_h)}_{\substack{=t_{(g,h)}\\ \text{(by the definition of }\Phi)}}\\ &= t_{(g,h)} \otimes t_{(g,h)}.\end{aligned}$$

Comparing this with (13.9.6), we obtain

$$\left(\Delta_{\mathbf{k}[G\times H]} \circ \Phi\right)(t_g \otimes t_h) = (\Phi \otimes \Phi)(\Delta_{\mathbf{k}G\otimes\mathbf{k}H}(t_g \otimes t_h)) = ((\Phi \otimes \Phi) \circ \Delta_{\mathbf{k}G\otimes\mathbf{k}H})(t_g \otimes t_h).$$

Now, forget that we fixed (g, h). We thus have shown that $\left(\Delta_{\mathbf{k}[G\times H]} \circ \Phi\right)(t_g \otimes t_h) = ((\Phi \otimes \Phi) \circ \Delta_{\mathbf{k}G\otimes\mathbf{k}H})(t_g \otimes t_h)$ for every $(g, h) \in G \times H$. In other words, the two maps $\Delta_{\mathbf{k}[G\times H]} \circ \Phi$ and $(\Phi \otimes \Phi) \circ \Delta_{\mathbf{k}G\otimes\mathbf{k}H}$ are equal to each other on the basis $(t_g \otimes t_h)_{(g,h)\in G\times H}$ of the $\mathbf{k}$-module $\mathbf{k}G \otimes \mathbf{k}H$. Therefore, these two maps must be identical (because they are $\mathbf{k}$-linear). In other words, $\Delta_{\mathbf{k}[G\times H]} \circ \Phi = (\Phi \otimes \Phi) \circ \Delta_{\mathbf{k}G\otimes\mathbf{k}H}$. This proves (13.9.5).

But the definition of $f \star (g \star h)$ yields

$$f \star (g \star h) = m \circ \underbrace{(f \otimes (g \star h))}_{=(\mathrm{id}_A \otimes m)\circ(f\otimes g\otimes h)\circ(\mathrm{id}_C \otimes\Delta)} \circ\Delta = m \circ (\mathrm{id}_A \otimes m) \circ (f \otimes g \otimes h) \circ (\mathrm{id}_C \otimes\Delta) \circ \Delta. \tag{13.10.1}$$

On the other hand, $f \star g = m \circ (f \otimes g) \circ \Delta$ (by the definition of $f \star g$), so that

$$\begin{aligned}\underbrace{(f \star g)}_{=m\circ(f\otimes g)\circ\Delta} \otimes \underbrace{h}_{=\mathrm{id}_A \circ h\circ\mathrm{id}_C} &= (m \circ (f \otimes g) \circ \Delta) \otimes (\mathrm{id}_C \circ h \circ \mathrm{id}_C) \\ &= (m \otimes \mathrm{id}_A) \circ ((f \otimes g) \otimes h) \circ (\Delta \otimes \mathrm{id}_C) \\ &= (m \otimes \mathrm{id}_A) \circ (f \otimes g \otimes h) \circ (\Delta \otimes \mathrm{id}_C) .\end{aligned}$$

Now, the definition of $(f \star g) \star h$ yields

$$(f \star g) \star h = m \circ \underbrace{((f \star g) \otimes h)}_{=(m\otimes\mathrm{id}_A)\circ(f\otimes g\otimes h)\circ(\Delta\otimes\mathrm{id}_C)} \circ\Delta = m \circ (m \otimes \mathrm{id}_A) \circ (f \otimes g \otimes h) \circ (\Delta \otimes \mathrm{id}_C) \circ \Delta. \tag{13.10.2}$$

Now, recall that A is a **k**-algebra, and hence the diagram (1.1.1) commutes (by the definition of a **k**-algebra). Thus, $m \circ (\mathrm{id}_A \otimes m) = m \circ (m \otimes \mathrm{id}_A)$. Also, C is a **k**-coalgebra, and thus the diagram (1.2.1) commutes (by the definition of a **k**-coalgebra). Hence, $(\mathrm{id}_C \otimes\Delta) \circ \Delta = (\Delta \otimes \mathrm{id}_C) \circ \Delta$. Now, (13.10.1) becomes

$$\begin{aligned}f \star (g \star h) &= \underbrace{m \circ (\mathrm{id}_A \otimes m)}_{=m\circ(m\otimes\mathrm{id}_A)} \circ (f \otimes g \otimes h) \circ \underbrace{(\mathrm{id}_C \otimes\Delta) \circ \Delta}_{=(\Delta\otimes\mathrm{id}_C)\circ\Delta} \\ &= m \circ (m \otimes \mathrm{id}_A) \circ (f \otimes g \otimes h) \circ (\Delta \otimes \mathrm{id}_C) \circ \Delta = (f \star g) \star h\end{aligned}$$

(by (13.10.2)). Thus, $f \star (g \star h) = (f \star g) \star h$ is proven, and the solution of Exercise 1.4.2 is complete.

13.11. **Solution to Exercise 1.4.4.** *Solution to Exercise 1.4.4.* (a) There are two ways to solve an exercise like this: either by explicitly evaluating the two sides on elements of their domain (in this case, it is of course enough to only evaluate them on pure tensors) or by diagram chasing. In the case of this particular exercise, both solutions are very easy, so let us show them both.

First solution: Here is the solution by explicit computations:

We need to prove that $(f \otimes g) \star (f' \otimes g') = (f \star f') \otimes (g \star g')$. For this, it is clearly enough to show that $((f \otimes g) \star (f' \otimes g')) (c \otimes d) = ((f \star f') \otimes (g \star g')) (c \otimes d)$ for every $c \in C$ and $d \in D$. But this can be done directly: Since $\Delta (c \otimes d) = \sum_{(c),(d)} (c_1 \otimes d_1) \otimes (c_2 \otimes d_2)$ (we are using the Sweedler notation here), we have

$$\begin{aligned}((f \otimes g) \star (f' \otimes g')) (c \otimes d) &= \sum_{(c),(d)} \underbrace{(f \otimes g) (c_1 \otimes d_1)}_{=f(c_1)\otimes g(d_1)} \cdot \underbrace{(f' \otimes g') (c_2 \otimes d_2)}_{=f'(c_2)\otimes g'(d_2)} \\ &= \sum_{(c),(d)} (f (c_1) \otimes g (d_1)) (f' (c_2) \otimes g' (d_2)) \\ &= \sum_{(c),(d)} f (c_1) f' (c_2) \otimes g (d_1) g' (d_2) .\end{aligned}$$

Compared to

$$\begin{aligned}((f \star f') \otimes (g \star g')) (c \otimes d) &= \underbrace{(f \star f') (c)}_{=\sum_{(c)} f(c_1)f'(c_2)} \otimes \underbrace{(g \star g') (d)}_{=\sum_{(d)} g(d_1)g'(d_2)} \\ &= \left(\sum_{(c)} f (c_1) f' (c_2)\right) \otimes \left(\sum_{(d)} g (d_1) g' (d_2)\right) \\ &= \sum_{(c),(d)} f (c_1) f' (c_2) \otimes g (d_1) g' (d_2) ,\end{aligned}$$

this yields $((f\otimes g)\star(f'\otimes g'))(c\otimes d) = ((f\star f')\otimes(g\star g'))(c\otimes d)$, which completes our solution of Exercise 1.4.4(a).

Second solution: Now comes the solution by diagram chasing. Actually, we will not see any diagrams here because in this particular case it would be a waste of space to draw them; everything can be done by a short computation. Denote by T the twist map $U\otimes V\to V\otimes U$ for any two $\mathbf{k}$-modules U and V. (We leave U and V out of the notation since these will always be clear from the context.) By the definition of convolution, we have

$$\begin{aligned}
(f\otimes g)\star(f'\otimes g') &= \underbrace{m_{A\otimes B}}_{=(m_A\otimes m_B)\circ(\mathrm{id}\otimes T\otimes\mathrm{id})} \circ\left((f\otimes g)\otimes(f'\otimes g')\right)\circ \underbrace{\Delta_{C\otimes D}}_{=(\mathrm{id}\otimes T\otimes\mathrm{id})\circ(\Delta_C\otimes\Delta_D)}\\
&= (m_A\otimes m_B)\circ\underbrace{(\mathrm{id}\otimes T\otimes\mathrm{id})\circ\left((f\otimes g)\otimes(f'\otimes g')\right)}_{=(f\otimes f'\otimes g\otimes g')\circ(\mathrm{id}\otimes T\otimes\mathrm{id})}\circ(\mathrm{id}\otimes T\otimes\mathrm{id})\circ(\Delta_C\otimes\Delta_D)\\
&= (m_A\otimes m_B)\circ(f\otimes f'\otimes g\otimes g')\circ\underbrace{(\mathrm{id}\otimes T\otimes\mathrm{id})\circ(\mathrm{id}\otimes T\otimes\mathrm{id})}_{\substack{=\mathrm{id}\\ (\text{since } T^2=\mathrm{id})}}\circ(\Delta_C\otimes\Delta_D)\\
&= (m_A\otimes m_B)\circ(f\otimes f'\otimes g\otimes g')\circ(\Delta_C\otimes\Delta_D)\\
&= \underbrace{(m_A\circ(f\otimes f')\circ\Delta_C)}_{\substack{=f\star f'\\ (\text{by the definition of convolution})}}\otimes\underbrace{(m_B\circ(g\otimes g')\circ\Delta_D)}_{\substack{=g\star g'\\ (\text{by the definition of convolution})}}\\
&= (f\star f')\otimes(g\star g').
\end{aligned}$$

This solves Exercise 1.4.4(a) again.

(b) Recall that the products in the $\mathbf{k}$-algebras $(\operatorname{Hom}(C,A),\star)$, $(\operatorname{Hom}(D,B),\star)$ and $(\operatorname{Hom}(C\otimes D,A\otimes B),\star)$ are the convolution products $\star$. For the sake of consistency, we will also denote by $\star$ the product in the $\mathbf{k}$-algebra $(\operatorname{Hom}(C,A),\star)\otimes(\operatorname{Hom}(D,B),\star)$.

We need to prove that R is a $\mathbf{k}$-algebra homomorphism. For this, it is clearly enough to show that R preserves products and sends the unity $(u_A\epsilon_C)\otimes(u_B\epsilon_D)$ of the $\mathbf{k}$-algebra $(\operatorname{Hom}(C,A),\star)\otimes(\operatorname{Hom}(D,B),\star)$ to the unity $u_{A\otimes B}\epsilon_{C\otimes D}$ of the $\mathbf{k}$-algebra $(\operatorname{Hom}(C\otimes D,A\otimes B),\star)$.

Let us check this. First, let us verify that R preserves products. So we need to show that $R(F\star F') = R(F)\star R(F')$ for all F and F' in $(\operatorname{Hom}(C,A),\star)\otimes(\operatorname{Hom}(D,B),\star)$. In proving this, we can WLOG assume that F and F' are pure tensors (since the claim $R(F\star F')=R(F)\star R(F')$ is linear in F and F'). Assume this, and write F and F' as $F=f\otimes g$ and $F'=f'\otimes g'$, respectively. Then,

$$\begin{aligned}
R(F\star F') &= R\left(\underbrace{(f\otimes g)\star(f'\otimes g')}_{\substack{=(f\star f')\otimes(g\star g')\\ (\text{by the definition of the product in}\\ (\operatorname{Hom}(C,A),\star)\otimes(\operatorname{Hom}(D,B),\star))}}\right) = R((f\star f')\otimes(g\star g')) = (f\star f')\otimes(g\star g')\\
&\qquad (\text{by the definition of } R\text{; note that the tensor sign changed its meaning here})\\
&= \underbrace{(f\otimes g)}_{\substack{=R(f\otimes g)\\ (\text{by the definition of } R)}}\star\underbrace{(f'\otimes g')}_{\substack{=R(f'\otimes g')\\ (\text{by the definition of } R)}} \qquad (\text{by Exercise 1.4.4(a)})\\
&= R\left(\underbrace{f\otimes g}_{=F}\right)\star R\left(\underbrace{f'\otimes g'}_{=F'}\right) = R(F)\star R(F'),
\end{aligned}$$

where the meaning of the tensor sign (each time standing either for a tensor or for a tensor product of maps) should be clear from the context. Thus, we are done checking that R preserves products.

It thus remains to verify that R sends $(u_A\epsilon_C) \otimes (u_B\epsilon_D)$ to $u_{A\otimes B}\epsilon_{C\otimes D}$. This is straightforward: If s denotes the canonical isomorphism $\mathbf{k} \to \mathbf{k} \otimes \mathbf{k}$, then

$$R\left((u_A\epsilon_C) \otimes (u_B\epsilon_D)\right) = (u_A\epsilon_C) \otimes (u_B\epsilon_D) = \underbrace{(u_A \otimes u_B)}_{=u_{A\otimes B}\circ s^{-1}} \circ \underbrace{(\epsilon_C \otimes \epsilon_D)}_{=s\circ\epsilon_{C\otimes D}} = u_{A\otimes B} \circ s^{-1} \circ s \circ \epsilon_{C\otimes D} = u_{A\otimes B}\epsilon_{C\otimes D}.$$

The solution of Exercise 1.4.4(b) is thus complete.

13.12. Solution to Exercise 1.4.5.

Solution to Exercise 1.4.5. Let f and g be two elements of $\operatorname{Hom}(C \otimes D, A)$. We are going to show that $\Phi(f) \star \Phi(g) = \Phi(f \star g)$.

Let $c \in C$. Let $d \in D$.

We shall use the Sweedler notation, namely writing $\Delta_C(c) = \sum_{(c)} c_1 \otimes c_2$ and $\Delta_D(d) = \sum_{(d)} d_1 \otimes d_2$. (This is a neat opportunity to practice the use of the Sweedler notation in a particularly simple setting. If you are uncomfortable with the Sweedler notation, you are invited to fix a decomposition $\Delta_C(c) = \sum_{i=1}^n p_i \otimes q_i$ of $\Delta_C(c)$ into a sum of pure tensors, as well as a similar decomposition $\Delta_D(d) = \sum_{j=1}^k x_j \otimes y_j$ for $\Delta_D(d)$, and to replace each appearance of one of the symbols

$$\sum_{(c)}, \quad c_1, \quad c_2, \quad \sum_{(d)}, \quad d_1, \quad d_2$$

by the symbol

$$\sum_{i=1}^n, \quad p_i, \quad q_i, \quad \sum_{j=1}^k, \quad x_j, \quad y_j,$$

respectively. This will translate our argument into a perfectly valid argument that does not use the Sweedler notation.)

The definition of the $\mathbf{k}$-coalgebra $C \otimes D$ yields

$$\Delta_{C\otimes D} = (\operatorname{id}_C \otimes T \otimes \operatorname{id}_D) \circ (\Delta_C \otimes \Delta_D).$$

Applying both sides of this equality to $c \otimes d$, we obtain

$$\begin{aligned}
\Delta_{C\otimes D}(c\otimes d) &= ((\mathrm{id}_C \otimes T \otimes \mathrm{id}_D)\circ(\Delta_C \otimes \Delta_D))(c\otimes d) = (\mathrm{id}_C \otimes T \otimes \mathrm{id}_D)\underbrace{((\Delta_C \otimes \Delta_D)(c\otimes d))}_{=\Delta_C(c)\otimes\Delta_D(d)} \\
&= (\mathrm{id}_C \otimes T \otimes \mathrm{id}_D)\left(\underbrace{\Delta_C(c)}_{=\sum_{(c)} c_1\otimes c_2} \otimes \underbrace{\Delta_D(d)}_{=\sum_{(d)} d_1\otimes d_2}\right) \\
&= (\mathrm{id}_C \otimes T \otimes \mathrm{id}_D)\left(\underbrace{\left(\sum_{(c)} c_1\otimes c_2\right)\otimes\left(\sum_{(d)} d_1\otimes d_2\right)}_{=\sum_{(c)}\sum_{(d)} c_1\otimes c_2\otimes d_1\otimes d_2}\right) \\
&= (\mathrm{id}_C \otimes T \otimes \mathrm{id}_D)\left(\sum_{(c)}\sum_{(d)} c_1\otimes c_2\otimes d_1\otimes d_2\right) \\
&= \sum_{(c)}\sum_{(d)}\underbrace{(\mathrm{id}_C \otimes T \otimes \mathrm{id}_D)(c_1\otimes c_2\otimes d_1\otimes d_2)}_{=\mathrm{id}_C(c_1)\otimes T(c_2\otimes d_1)\otimes \mathrm{id}_D(d_2)} \\
&= \sum_{(c)}\sum_{(d)}\underbrace{\mathrm{id}_C(c_1)}_{=c_1}\otimes\underbrace{T(c_2\otimes d_1)}_{\substack{=d_1\otimes c_2 \\ \text{(by the definition of } T)}}\otimes\underbrace{\mathrm{id}_D(d_2)}_{=d_2} \\
&= \sum_{(c)}\sum_{(d)} c_1\otimes d_1\otimes c_2\otimes d_2. \qquad (13.12.1)
\end{aligned}$$

Now, the definition of convolution yields

$$(\Phi(f)\star\Phi(g))(c) = \sum_{(c)} (\Phi(f))(c_1)\star(\Phi(g))(c_2)$$

(since the multiplication in the $\mathbf{k}$-algebra $(\operatorname{Hom}(D,A),\star)$ is $\star$). Applying both sides of this equality to d, we obtain

$$\begin{aligned}
&((\Phi(f)\star\Phi(g))(c))(d) \\
&= \left(\sum_{(c)} (\Phi(f))(c_1)\star(\Phi(g))(c_2)\right)(d) \\
&= \sum_{(c)} \underbrace{((\Phi(f))(c_1)\star(\Phi(g))(c_2))(d)}_{\substack{=\sum_{(d)}((\Phi(f))(c_1))(d_1)\cdot((\Phi(g))(c_2))(d_2) \\ \text{(by the definition of convolution)}}} \\
&= \sum_{(c)}\sum_{(d)} \underbrace{((\Phi(f))(c_1))(d_1)}_{\substack{=f(c_1\otimes d_1) \\ \text{(by the definition of }\Phi)}}\cdot\underbrace{((\Phi(g))(c_2))(d_2)}_{\substack{=g(c_2\otimes d_2) \\ \text{(by the definition of }\Phi)}} \\
&= \sum_{(c)}\sum_{(d)} f(c_1\otimes d_1)\,g(c_2\otimes d_2). \qquad (13.12.2)
\end{aligned}$$

On the other hand, the definition of Φ yields

$$\begin{aligned}
&\left(\left(\Phi\left(f\star g\right)\right)\left(c\right)\right)\left(d\right)\\
&= \underbrace{\left(f\star g\right)}_{\substack{=m_A\circ(f\otimes g)\circ\Delta_{C\otimes D}\\ \text{(by the definition of convolution)}}}\left(c\otimes d\right) = \left(m_A\circ\left(f\otimes g\right)\circ\Delta_{C\otimes D}\right)\left(c\otimes d\right)\\
&= m_A\left(\left(f\otimes g\right)\left(\underbrace{\Delta_{C\otimes D}\left(c\otimes d\right)}_{\substack{=\sum_{(c)}\sum_{(d)} c_1\otimes d_1\otimes c_2\otimes d_2\\ \text{(by (13.12.1))}}}\right)\right) = m_A\left(\left(f\otimes g\right)\left(\sum_{(c)}\sum_{(d)} c_1\otimes d_1\otimes c_2\otimes d_2\right)\right)\\
&= \sum_{(c)}\sum_{(d)} m_A\left(\underbrace{\left(f\otimes g\right)\left(c_1\otimes d_1\otimes c_2\otimes d_2\right)}_{=f(c_1\otimes d_1)\otimes g(c_2\otimes d_2)}\right) = \sum_{(c)}\sum_{(d)}\underbrace{m_A\left(f\left(c_1\otimes d_1\right)\otimes g\left(c_2\otimes d_2\right)\right)}_{\substack{=f(c_1\otimes d_1)g(c_2\otimes d_2)\\ \text{(by the definition of } m_A\text{)}}}\\
&= \sum_{(c)}\sum_{(d)} f\left(c_1\otimes d_1\right)g\left(c_2\otimes d_2\right).
\end{aligned}$$

Comparing this with (13.12.2), we obtain

$$\left(\left(\Phi\left(f\right)\star\Phi\left(g\right)\right)\left(c\right)\right)\left(d\right) = \left(\left(\Phi\left(f\star g\right)\right)\left(c\right)\right)\left(d\right).$$

Now, forget that we fixed d. We thus have shown that $\left(\left(\Phi\left(f\right)\star\Phi\left(g\right)\right)\left(c\right)\right)\left(d\right) = \left(\left(\Phi\left(f\star g\right)\right)\left(c\right)\right)\left(d\right)$ for each $d\in D$. In other words, we have $\left(\Phi\left(f\right)\star\Phi\left(g\right)\right)\left(c\right) = \left(\Phi\left(f\star g\right)\right)\left(c\right)$.

Now, forget that we fixed c. We thus have shown that $\left(\Phi\left(f\right)\star\Phi\left(g\right)\right)\left(c\right) = \left(\Phi\left(f\star g\right)\right)\left(c\right)$ for each $c\in C$. In other words, we have $\Phi\left(f\right)\star\Phi\left(g\right) = \Phi\left(f\star g\right)$.

Now, forget that we fixed f and g. We thus have proven that every two elements f and g of $\operatorname{Hom}\left(C\otimes D, A\right)$ satisfy

$$\Phi\left(f\right)\star\Phi\left(g\right) = \Phi\left(f\star g\right). \tag{13.12.3}$$

Now, let θ be the canonical $\mathbf{k}$-module isomorphism $\mathbf{k}\otimes\mathbf{k}\to\mathbf{k}$. Then, the definition of the $\mathbf{k}$-coalgebra $C\otimes D$ yields

$$\epsilon_{C\otimes D} = \theta\circ\left(\epsilon_C\otimes\epsilon_D\right).$$

The unity of the $\mathbf{k}$-algebra $\left(\operatorname{Hom}\left(D,A\right),\star\right)$ is $u_A\circ\epsilon_D$. In other words,

$$1_{\left(\operatorname{Hom}(D,A),\star\right)} = u_A\circ\epsilon_D. \tag{13.12.4}$$

Similarly,

$$1_{\left(\operatorname{Hom}(C,(\operatorname{Hom}(D,A),\star)),\star\right)} = u_{\left(\operatorname{Hom}(D,A),\star\right)}\circ\epsilon_C \tag{13.12.5}$$

and

$$1_{\left(\operatorname{Hom}(C\otimes D,A),\star\right)} = u_A\circ\epsilon_{C\otimes D}. \tag{13.12.6}$$

Let $c\in C$. Let $d\in D$. The definition of Φ yields

$$\begin{aligned}
&\left(\left(\Phi\left(u_A\circ\epsilon_{C\otimes D}\right)\right)\left(c\right)\right)\left(d\right)\\
&= \left(u_A\circ\epsilon_{C\otimes D}\right)\left(c\otimes d\right) = u_A\left(\underbrace{\epsilon_{C\otimes D}}_{=\theta\circ(\epsilon_C\otimes\epsilon_D)}\left(c\otimes d\right)\right) = u_A\left(\left(\theta\circ\left(\epsilon_C\otimes\epsilon_D\right)\right)\left(c\otimes d\right)\right)\\
&= u_A\left(\theta\left(\underbrace{\left(\epsilon_C\otimes\epsilon_D\right)\left(c\otimes d\right)}_{=\epsilon_C(c)\otimes\epsilon_D(d)}\right)\right) = u_A\left(\underbrace{\theta\left(\epsilon_C\left(c\right)\otimes\epsilon_D\left(d\right)\right)}_{\substack{=\epsilon_C(c)\epsilon_D(d)\\ \text{(by the definition of } \theta\text{)}}}\right) = u_A\left(\epsilon_C\left(c\right)\epsilon_D\left(d\right)\right)\\
&= \epsilon_C\left(c\right)\epsilon_D\left(d\right)\cdot 1_A \qquad \left(\text{by the definition of } u_A\right).
\end{aligned}$$

Comparing this with

$$\Bigg(\underbrace{\left(u_{(\operatorname{Hom}(D,A),\star)}\circ\epsilon_C\right)(c)}_{=u_{(\operatorname{Hom}(D,A),\star)}(\epsilon_C(c))}\Bigg)(d) = \Bigg(\underbrace{u_{(\operatorname{Hom}(D,A),\star)}\left(\epsilon_C(c)\right)}_{\substack{=\epsilon_C(c)\cdot 1_{(\operatorname{Hom}(D,A),\star)} \\ \text{(by the definition of } u_{(\operatorname{Hom}(D,A),\star)}\text{)}}}\Bigg)(d) = \left(\epsilon_C(c)\cdot 1_{(\operatorname{Hom}(D,A),\star)}\right)(d)$$

$$= \epsilon_C(c)\cdot \underbrace{1_{(\operatorname{Hom}(D,A),\star)}}_{=u_A\circ\epsilon_D}(d) = \epsilon_C(c)\cdot \underbrace{(u_A\circ\epsilon_D)(d)}_{\substack{=u_A(\epsilon_D(d))=\epsilon_D(d)\cdot 1_A \\ \text{(by the definition of } u_A\text{)}}} = \epsilon_C(c)\,\epsilon_D(d)\cdot 1_A,$$

we obtain $\left(\left(\Phi\left(u_A\circ\epsilon_{C\otimes D}\right)\right)(c)\right)(d) = \left(\left(u_{(\operatorname{Hom}(D,A),\star)}\circ\epsilon_C\right)(c)\right)(d)$.

Now, forget that we fixed d. We thus have shown that $\left(\left(\Phi\left(u_A\circ\epsilon_{C\otimes D}\right)\right)(c)\right)(d) = \left(\left(u_{(\operatorname{Hom}(D,A),\star)}\circ\epsilon_C\right)(c)\right)(d)$ for each $d\in D$. In other words, we have $\left(\Phi\left(u_A\circ\epsilon_{C\otimes D}\right)\right)(c) = \left(u_{(\operatorname{Hom}(D,A),\star)}\circ\epsilon_C\right)(c)$.

Now, forget that we fixed c. We thus have shown that $\left(\Phi\left(u_A\circ\epsilon_{C\otimes D}\right)\right)(c) = \left(u_{(\operatorname{Hom}(D,A),\star)}\circ\epsilon_C\right)(c)$ for each $c\in C$. In other words, we have $\Phi\left(u_A\circ\epsilon_{C\otimes D}\right) = u_{(\operatorname{Hom}(D,A),\star)}\circ\epsilon_C$.

Now,

$$\Phi\Bigg(\underbrace{1_{(\operatorname{Hom}(C\otimes D,A),\star)}}_{\substack{=u_A\circ\epsilon_{C\otimes D} \\ \text{(by (13.12.6))}}}\Bigg) = \Phi\left(u_A\circ\epsilon_{C\otimes D}\right) = u_{(\operatorname{Hom}(D,A),\star)}\circ\epsilon_C = 1_{(\operatorname{Hom}(C,(\operatorname{Hom}(D,A),\star)),\star)}$$

(by (13.12.5)).

We now know that the map Φ is $\mathbf{k}$-linear and satisfies (13.12.3) and $\Phi\left(1_{(\operatorname{Hom}(C\otimes D,A),\star)}\right) = 1_{(\operatorname{Hom}(C,(\operatorname{Hom}(D,A),\star)),\star)}$. Thus, the map Φ is a $\mathbf{k}$-algebra homomorphism

$$\left(\operatorname{Hom}\left(C\otimes D,A\right),\star\right)\to\left(\operatorname{Hom}\left(C,\left(\operatorname{Hom}\left(D,A\right),\star\right)\right),\star\right).$$

Since the map Φ is furthermore invertible (because Φ is a $\mathbf{k}$-module isomorphism), we thus conclude that Φ is a $\mathbf{k}$-algebra isomorphism

$$\left(\operatorname{Hom}\left(C\otimes D,A\right),\star\right)\to\left(\operatorname{Hom}\left(C,\left(\operatorname{Hom}\left(D,A\right),\star\right)\right),\star\right).$$

This solves Exercise 1.4.5.

13.13. **Solution to Exercise 1.4.15.** *Solution to Exercise 1.4.15.*

Proof of Proposition 1.4.14. Let S_A and S_B be the antipodes of the Hopf algebras A and B. Recall that the antipode S_A of the Hopf algebra A is the 2-sided inverse under $\star$ for the identity map $\operatorname{id}_A\in\operatorname{Hom}(A,A)$. In other words, S_A is the multiplicative inverse of id_A in the convolution algebra $\left(\operatorname{Hom}(A,A),\star\right)$. Therefore,

$$S_A\star\operatorname{id}_A = 1_{(\operatorname{Hom}(A,A),\star)} \qquad\text{and}\qquad \operatorname{id}_A\star S_A = 1_{(\operatorname{Hom}(A,A),\star)}.$$

The same argument (applied to B instead of A) shows that

$$S_B\star\operatorname{id}_B = 1_{(\operatorname{Hom}(B,B),\star)} \qquad\text{and}\qquad \operatorname{id}_B\star S_B = 1_{(\operatorname{Hom}(B,B),\star)}.$$

Now, it is easy to see that

$$1_{(\operatorname{Hom}(A,A),\star)}\otimes 1_{(\operatorname{Hom}(B,B),\star)} = 1_{(\operatorname{Hom}(A\otimes B,A\otimes B),\star)}$$

(as maps from $A\otimes B$ to $A\otimes B$) [400].

[400] *Proof.* We know that the unity of the convolution algebra $\left(\operatorname{Hom}(A,A),\star\right)$ is $u_A\epsilon_A$. In other words, $1_{(\operatorname{Hom}(A,A),\star)} = u_A\epsilon_A$. Similarly, $1_{(\operatorname{Hom}(B,B),\star)} = u_B\epsilon_B$ and $1_{(\operatorname{Hom}(A\otimes B,A\otimes B),\star)} = u_{A\otimes B}\epsilon_{A\otimes B}$.

Now, let s denote the canonical isomorphism $\mathbf{k}\to\mathbf{k}\otimes\mathbf{k}$. Then, the definition of the $\mathbf{k}$-algebra $A\otimes B$ yields $u_{A\otimes B} = (u_A\otimes u_B)\circ s$. On the other hand, the definition of the $\mathbf{k}$-coalgebra $A\otimes B$ yields $\epsilon_{A\otimes B} = s^{-1}\circ(\epsilon_A\otimes\epsilon_B)$ (since s^{-1} is the

Exercise 1.4.4(a) (applied to $C = A$, $D = B$, $f = S_A$, $f' = \operatorname{id}_A$, $g = S_B$ and $g' = \operatorname{id}_B$) shows that

$$(S_A \otimes S_B) \star (\operatorname{id}_A \otimes \operatorname{id}_B) = \underbrace{(S_A \star \operatorname{id}_A)}_{=1_{(\operatorname{Hom}(A,A),\star)}} \otimes \underbrace{(S_B \star \operatorname{id}_B)}_{=1_{(\operatorname{Hom}(B,B),\star)}}$$
$$= 1_{(\operatorname{Hom}(A,A),\star)} \otimes 1_{(\operatorname{Hom}(B,B),\star)} = 1_{(\operatorname{Hom}(A\otimes B,A\otimes B),\star)} \qquad (13.13.1)$$

in the convolution algebra $\operatorname{Hom}(A \otimes B, A \otimes B)$.

Exercise 1.4.4(a) (applied to $C = A$, $D = B$, $f = \operatorname{id}_A$, $f' = S_A$, $g = \operatorname{id}_B$ and $g' = S_B$) shows that

$$(\operatorname{id}_A \otimes \operatorname{id}_B) \star (S_A \otimes S_B) = \underbrace{(\operatorname{id}_A \star S_A)}_{=1_{(\operatorname{Hom}(A,A),\star)}} \otimes \underbrace{(\operatorname{id}_B \star S_B)}_{=1_{(\operatorname{Hom}(B,B),\star)}}$$
$$= 1_{(\operatorname{Hom}(A,A),\star)} \otimes 1_{(\operatorname{Hom}(B,B),\star)} = 1_{(\operatorname{Hom}(A\otimes B,A\otimes B),\star)}$$

in the convolution algebra $\operatorname{Hom}(A \otimes B, A \otimes B)$. Combining this with (13.13.1), we conclude that the two elements $S_A \otimes S_B$ and $\operatorname{id}_A \otimes \operatorname{id}_B$ of the convolution algebra $(\operatorname{Hom}(A \otimes B, A \otimes B), \star)$ are mutually inverse. In other words, $S_A \otimes S_B$ is a 2-sided inverse for $\operatorname{id}_A \otimes \operatorname{id}_B$ under $\star$. In other words, $S_A \otimes S_B$ is a 2-sided inverse for $\operatorname{id}_{A\otimes B}$ under $\star$ (since $\operatorname{id}_A \otimes \operatorname{id}_B = \operatorname{id}_{A\otimes B}$). Hence, the element $\operatorname{id}_{A\otimes B} \in \operatorname{Hom}(A \otimes B, A \otimes B)$ has a 2-sided inverse under $\star$ (namely, $S_A \otimes S_B$).

We recall that a bialgebra D is a Hopf algebra if and only if the element $\operatorname{id}_D \in \operatorname{Hom}(D, D)$ has a 2-sided inverse under $\star$. Applying this to $D = A \otimes B$, we conclude that the bialgebra $A \otimes B$ is a Hopf algebra if and only if the element $\operatorname{id}_{A\otimes B} \in \operatorname{Hom}(A \otimes B, A \otimes B)$ has a 2-sided inverse under $\star$. Therefore, the bialgebra $A \otimes B$ is a Hopf algebra (since the element $\operatorname{id}_{A\otimes B} \in \operatorname{Hom}(A \otimes B, A \otimes B)$ has a 2-sided inverse under $\star$).

The antipode of any Hopf algebra D is the 2-sided inverse for id_D under $\star$. Applying this to $D = A \otimes B$, we conclude that the antipode of $A\otimes B$ is the 2-sided inverse for $\operatorname{id}_{A\otimes B}$ under $\star$. In other words, the antipode of $A \otimes B$ is the map $S_A \otimes S_B : A \otimes B \to A \otimes B$ (since the 2-sided inverse for $\operatorname{id}_{A\otimes B}$ under $\star$ is the map $S_A \otimes S_B : A \otimes B \to A \otimes B$ (since $S_A \otimes S_B$ is a 2-sided inverse for $\operatorname{id}_{A\otimes B}$ under $\star$)). This completes the proof of Proposition 1.4.14. □

Thus, Exercise 1.4.15 is solved.

13.14. Solution to Exercise 1.4.19.

Solution to Exercise 1.4.19. Let us start with an observation which is irrelevant to our solution of the exercise. Namely, let us notice that

$$m^{(k)} (a_1 \otimes a_2 \otimes ... \otimes a_{k+1}) = a_1 (a_2 (a_3 (... (a_k a_{k+1}) ...)))$$

for any $k \geq 0$ and any $k+1$ elements $a_1, a_2, ..., a_{k+1}$ of A. The statements of Exercise 1.4.19 are nothing but different aspects of what is known as "general associativity"[401] (although they all fall short of defining an "arbitrary bracketing" of a $(k+1)$-fold product), written in an element-free fashion (that is, written without any reference to elements of A, but only in terms of maps). For instance, part (a) of the exercise says that any $k+1$ elements $a_1, a_2, ..., a_{k+1}$ of A satisfy

$$a_1 (a_2 (a_3 (... (a_k a_{k+1}) ...))) = (a_1 (a_2 (a_3 (... (a_i a_{i+1}) ...)))) \cdot (a_{i+2} (a_{i+3} (a_{i+4} (... (a_k a_{k+1}) ...)))) .$$

However, there is virtue in solving Exercise 1.4.19 in an element-free way (i.e., without referring to elements, but only referring to maps), because such a solution will automatically yield a solution of Exercise 1.4.20 by reversing all arrows. So let us show an element-free solution of Exercise 1.4.19.

(a) We will solve Exercise 1.4.19(a) by induction over k.

The induction base ($k = 0$) is vacuously true, since there exists no $0 \leq i \leq k-1$ for $k = 0$. So let us proceed to the induction step. Let K be a positive integer. We want to prove that the claim of Exercise 1.4.19(a)

canonical isomorphism $\mathbf{k} \otimes \mathbf{k} \to \mathbf{k}$). Thus, $\epsilon_A \otimes \epsilon_B = s \circ \epsilon_{A\otimes B}$. Now,

$$\underbrace{1_{(\operatorname{Hom}(A,A),\star)}}_{=u_A\epsilon_A} \otimes \underbrace{1_{(\operatorname{Hom}(B,B),\star)}}_{=u_B\epsilon_B} = (u_A\epsilon_A) \otimes (u_B\epsilon_B) = (u_A \otimes u_B) \circ \underbrace{(\epsilon_A \otimes \epsilon_B)}_{=s\circ\epsilon_{A\otimes B}} = \underbrace{(u_A \otimes u_B) \circ s}_{=u_{A\otimes B}} \circ \epsilon_{A\otimes B} = u_{A\otimes B}\epsilon_{A\otimes B}$$
$$= 1_{(\operatorname{Hom}(A\otimes B,A\otimes B),\star)}.$$

Qed.

[401] that is, the rule stating that the product of several elements of a $\mathbf{k}$-algebra does not depend on the bracketing

holds for $k = K$, assuming (as the induction hypothesis) that the claim of Exercise 1.4.19(a) holds for $k = K - 1$.

We first notice that

$$m^{(K-1)} = m \circ \left(m^{(i)} \otimes m^{((K-1)-1-i)} \right) \tag{13.14.1}$$

for all $0 \leq i \leq (K-1) - 1$. (This is merely a restatement of the induction hypothesis.)

Now fix $0 \leq i \leq K - 1$. We need to show that

$$m^{(K)} = m \circ \left(m^{(i)} \otimes m^{(K-1-i)} \right). \tag{13.14.2}$$

If $i = 0$, then (13.14.2) is obviously true[402]. Hence, we can WLOG assume that we don't have $i = 0$. Assume this. Then, $i > 0$. Hence, the recursive definition of $m^{(i)}$ yields $m^{(i)} = m \circ \left(\mathrm{id}_A \otimes m^{(i-1)} \right)$. Thus,

$$\begin{aligned}
\underbrace{m^{(i)}}_{= m \circ \left(\mathrm{id}_A \otimes m^{(i-1)} \right)} \otimes \underbrace{m^{(K-1-i)}}_{= \mathrm{id}_A \circ m^{(K-1-i)}} &= \left(m \circ \left(\mathrm{id}_A \otimes m^{(i-1)} \right) \right) \otimes \left(\mathrm{id}_A \circ m^{(K-1-i)} \right) \\
&= \left(m \otimes \mathrm{id}_A \right) \circ \left(\left(\mathrm{id}_A \otimes m^{(i-1)} \right) \otimes m^{(K-1-i)} \right) \\
&= \left(m \otimes \mathrm{id}_A \right) \circ \left(\mathrm{id}_A \otimes m^{(i-1)} \otimes m^{(K-1-i)} \right)
\end{aligned} \tag{13.14.3}$$

On the other hand,

$$\begin{aligned}
\mathrm{id}_A \otimes \underbrace{m^{(K-1)}}_{\substack{= m \circ \left(m^{(i-1)} \otimes m^{((K-1)-1-(i-1))} \right) \\ \text{(by (13.14.1), applied to } i-1 \text{ instead of } i)}} &= \mathrm{id}_A \otimes \left(m \circ \left(m^{(i-1)} \otimes m^{((K-1)-1-(i-1))} \right) \right) \\
&= \mathrm{id}_A \otimes \left(m \circ \left(m^{(i-1)} \otimes m^{(K-1-i)} \right) \right) \\
&= \left(\mathrm{id}_A \otimes m \right) \circ \left(\mathrm{id}_A \otimes \left(m^{(i-1)} \otimes m^{(K-1-i)} \right) \right) \\
&= \left(\mathrm{id}_A \otimes m \right) \circ \left(\mathrm{id}_A \otimes m^{(i-1)} \otimes m^{(K-1-i)} \right)
\end{aligned} \tag{13.14.4}$$

Now, the upper left triangle in the diagram

$$\begin{tikzcd}
A^{\otimes(K+1)} \arrow[rrd, bend left, "\mathrm{id}_A \otimes m^{(K-1)}"] \arrow[rd, "\mathrm{id}_A \otimes m^{(i-1)} \otimes m^{(K-1-i)}"] \arrow[rdd, bend right, "m^{(i)} \otimes m^{(K-1-i)}"'] & & \\
& A \otimes A \otimes A \arrow[r, "\mathrm{id}_A \otimes m"'] \arrow[d, "m \otimes \mathrm{id}_A"'] & A \otimes A \arrow[d, "m"] \\
& A \otimes A \arrow[r, "m"'] & A
\end{tikzcd} \tag{13.14.5}$$

is commutative (by (13.14.4)), and so is the lower left triangle (according to (13.14.3)). Since the square in the diagram (13.14.5) is also commutative (by the commutativity of (1.1.1)), we thus conclude that the whole diagram (13.14.5) is commutative. Hence, following the outermost arrows in this diagram, we obtain $m \circ \left(\mathrm{id}_A \otimes m^{(K-1)} \right) = m \circ \left(m^{(i)} \otimes m^{(K-1-i)} \right)$. Now, the recursive definition of $m^{(K)}$ yields $m^{(K)} = m \circ \left(\mathrm{id}_A \otimes m^{(K-1)} \right) = m \circ \left(m^{(i)} \otimes m^{(K-1-i)} \right)$. Hence, (13.14.2) is proven.

We thus have shown that Exercise 1.4.19(a) holds for $k = K$. This completes the induction step, and thus Exercise 1.4.19(a) is solved by induction.

[402] because if $i = 0$, then $m \circ \left(\underbrace{m^{(i)}}_{= m^{(0)} = \mathrm{id}_A} \otimes \underbrace{m^{(K-1-i)}}_{= m^{(K-1-0)} = m^{(K-1)}} \right) = m \circ \left(\mathrm{id}_A \otimes m^{(K-1)} \right) = m^{(K)}$ (by the inductive definition of $m^{(K)}$)

(b) Let $k \geq 1$. Then, Exercise 1.4.19(a) (applied to $i = k-1$) yields $m^{(k)} = m \circ \left(m^{(k-1)} \otimes \underbrace{m^{(k-1-(k-1))}}_{=m^{(0)}=\mathrm{id}_A} \right) = m \circ \left(m^{(k-1)} \otimes \mathrm{id}_A \right)$. Thus, Exercise 1.4.19(b) is solved.

(c) We will solve Exercise 1.4.19(c) by induction over k.

The induction base ($k = 0$) is vacuously true, since there exists no $0 \leq i \leq k-1$ for $k = 0$. So let us proceed to the induction step. Let K be a positive integer. We want to prove that the claim of Exercise 1.4.19(c) holds for $k = K$, assuming (as the induction hypothesis) that the claim of Exercise 1.4.19(c) holds for $k = K - 1$.

So let $0 \leq i \leq K - 1$ be arbitrary. Thus, $K - 1 \geq 0$, so that $K \geq 1$. We are going to prove that

$$m^{(K)} = m^{(K-1)} \circ \left(\mathrm{id}_{A^{\otimes i}} \otimes m \otimes \mathrm{id}_{A^{\otimes (K-1-i)}} \right). \tag{13.14.6}$$

We must be in one of the following three cases:

Case 1: We have $i \neq 0$.

Case 2: We have $i \neq K - 1$.

Case 3: We have neither $i \neq 0$ nor $i \neq K - 1$.

Let us consider Case 1 first. In this case, we have $i \neq 0$. Hence, $i \geq 1$ (because $0 \leq i$), so that $i - 1 \geq 0$. Thus, we can apply Exercise 1.4.19(c) to $K - 1$ and $i - 1$ instead of k and i (because we have assumed that the claim of Exercise 1.4.19(c) holds for $k = K - 1$). As a result, we obtain

$$\begin{aligned} m^{(K-1)} &= \underbrace{m^{((K-1)-1)}}_{=m^{(K-2)}} \circ \left(\mathrm{id}_{A^{\otimes (i-1)}} \otimes m \otimes \underbrace{\mathrm{id}_{A^{\otimes ((K-1)-1-(i-1))}}}_{=\mathrm{id}_{A^{\otimes (K-1-i)}}} \right) \\ &= m^{(K-2)} \circ \left(\mathrm{id}_{A^{\otimes (i-1)}} \otimes m \otimes \mathrm{id}_{A^{\otimes (K-1-i)}} \right). \end{aligned}$$

But $K \geq 1$. Hence, the recursive definition of $m^{(K)}$ yields

$$\begin{aligned} m^{(K)} &= m \circ \left(\mathrm{id}_A \otimes \underbrace{m^{(K-1)}}_{=m^{(K-2)} \circ \left(\mathrm{id}_{A^{\otimes (i-1)}} \otimes m \otimes \mathrm{id}_{A^{\otimes (K-1-i)}} \right)} \right) \\ &= m \circ \underbrace{\left(\mathrm{id}_A \otimes \left(m^{(K-2)} \circ \left(\mathrm{id}_{A^{\otimes (i-1)}} \otimes m \otimes \mathrm{id}_{A^{\otimes (K-1-i)}} \right) \right) \right)}_{=\left(\mathrm{id}_A \otimes m^{(K-2)} \right) \circ \left(\mathrm{id}_A \otimes \left(\mathrm{id}_{A^{\otimes (i-1)}} \otimes m \otimes \mathrm{id}_{A^{\otimes (K-1-i)}} \right) \right)} \\ &= m \circ \left(\mathrm{id}_A \otimes m^{(K-2)} \right) \circ \underbrace{\left(\mathrm{id}_A \otimes \left(\mathrm{id}_{A^{\otimes (i-1)}} \otimes m \otimes \mathrm{id}_{A^{\otimes (K-1-i)}} \right) \right)}_{=\mathrm{id}_A \otimes \mathrm{id}_{A^{\otimes (i-1)}} \otimes m \otimes \mathrm{id}_{A^{\otimes (K-1-i)}}} \\ &= m \circ \left(\mathrm{id}_A \otimes m^{(K-2)} \right) \circ \left(\underbrace{\mathrm{id}_A \otimes \mathrm{id}_{A^{\otimes (i-1)}}}_{=\mathrm{id}_{A^{\otimes i}}} \otimes m \otimes \mathrm{id}_{A^{\otimes (K-1-i)}} \right) \\ &= m \circ \left(\mathrm{id}_A \otimes m^{(K-2)} \right) \circ \left(\mathrm{id}_{A^{\otimes i}} \otimes m \otimes \mathrm{id}_{A^{\otimes (K-1-i)}} \right). \end{aligned} \tag{13.14.7}$$

But $i \geq 1$ and $i \leq K - 1$ together yield $1 \leq i \leq K - 1$, so that $K - 1 \geq 1$. Thus, the recursive definition of $m^{(K-1)}$ yields

$$m^{(K-1)} = m \circ \left(\mathrm{id}_A \otimes \underbrace{m^{((K-1)-1)}}_{=m^{(K-2)}} \right) = m \circ \left(\mathrm{id}_A \otimes m^{(K-2)} \right). \tag{13.14.8}$$

Hence, (13.14.7) becomes

$$\begin{aligned} m^{(K)} &= \underbrace{m \circ \left(\mathrm{id}_A \otimes m^{(K-2)}\right)}_{\substack{=m^{(K-1)} \\ \text{(by (13.14.8))}}} \circ \left(\mathrm{id}_{A^{\otimes i}} \otimes m \otimes \mathrm{id}_{A^{\otimes(K-1-i)}}\right) \\ &= m^{(K-1)} \circ \left(\mathrm{id}_{A^{\otimes i}} \otimes m \otimes \mathrm{id}_{A^{\otimes(K-1-i)}}\right). \end{aligned}$$

Thus, (13.14.6) is proven in Case 1.

Let us next consider Case 2. In this case, we have $i \neq K-1$. Hence, $i < K-1$ (since $0 \leq k-1$), so that $i \leq (K-1)-1$ (since i and $K-1$ are integers). Thus, we can apply Exercise 1.4.19(c) to $K-1$ instead of k (because we have assumed that the claim of Exercise 1.4.19(c) holds for $k = K-1$). As a result, we obtain

$$\begin{aligned} m^{(K-1)} &= \underbrace{m^{((K-1)-1)}}_{=m^{(K-2)}} \circ \left(\mathrm{id}_{A^{\otimes i}} \otimes m \otimes \underbrace{\mathrm{id}_{A^{\otimes((K-1)-1-i)}}}_{=\mathrm{id}_{A^{\otimes(K-1-i-1)}}}\right) \\ &= m^{(K-2)} \circ \left(\mathrm{id}_{A^{\otimes i}} \otimes m \otimes \mathrm{id}_{A^{\otimes(K-1-i-1)}}\right). \end{aligned}$$

But $K \geq 1$. Hence, Exercise 1.4.19(b) (applied to $k = K$) yields

$$\begin{aligned} m^{(K)} &= m \circ \left(\underbrace{m^{(K-1)}}_{=m^{(K-2)} \circ \left(\mathrm{id}_{A^{\otimes i}} \otimes m \otimes \mathrm{id}_{A^{\otimes(K-1-i-1)}}\right)} \otimes \mathrm{id}_A\right) \\ &= m \circ \underbrace{\left(\left(m^{(K-2)} \circ \left(\mathrm{id}_{A^{\otimes i}} \otimes m \otimes \mathrm{id}_{A^{\otimes(K-1-i-1)}}\right)\right) \otimes \mathrm{id}_A\right)}_{=\left(m^{(K-2)} \otimes \mathrm{id}_A\right) \circ \left(\left(\mathrm{id}_{A^{\otimes i}} \otimes m \otimes \mathrm{id}_{A^{\otimes(K-1-i-1)}}\right) \otimes \mathrm{id}_A\right)} \\ &= m \circ \left(m^{(K-2)} \otimes \mathrm{id}_A\right) \circ \underbrace{\left(\left(\mathrm{id}_{A^{\otimes i}} \otimes m \otimes \mathrm{id}_{A^{\otimes(K-1-i-1)}}\right) \otimes \mathrm{id}_A\right)}_{=\mathrm{id}_{A^{\otimes i}} \otimes m \otimes \mathrm{id}_{A^{\otimes(K-1-i-1)}} \otimes \mathrm{id}_A} \\ &= m \circ \left(m^{(K-2)} \otimes \mathrm{id}_A\right) \circ \left(\mathrm{id}_{A^{\otimes i}} \otimes m \otimes \underbrace{\mathrm{id}_{A^{\otimes(K-1-i-1)}} \otimes \mathrm{id}_A}_{=\mathrm{id}_{A^{\otimes(K-1-i)}}}\right) \\ &= m \circ \left(m^{(K-2)} \otimes \mathrm{id}_A\right) \circ \left(\mathrm{id}_{A^{\otimes i}} \otimes m \otimes \mathrm{id}_{A^{\otimes(K-1-i)}}\right). \end{aligned} \tag{13.14.9}$$

But $0 \leq i \leq (K-1)-1$ yields $1 \leq K-1$, so that $K-1 \geq 1$. Thus, Exercise 1.4.19(b) (applied to $k = K-1$) yields

$$m^{(K-1)} = m \circ \left(\underbrace{m^{((K-1)-1)}}_{=m^{(K-2)}} \otimes \mathrm{id}_A\right) = m \circ \left(m^{(K-2)} \otimes \mathrm{id}_A\right). \tag{13.14.10}$$

Hence, (13.14.9) becomes

$$\begin{aligned} m^{(K)} &= \underbrace{m \circ \left(m^{(K-2)} \otimes \mathrm{id}_A\right)}_{\substack{=m^{(K-1)} \\ \text{(by (13.14.10))}}} \circ \left(\mathrm{id}_{A^{\otimes i}} \otimes m \otimes \mathrm{id}_{A^{\otimes(K-1-i)}}\right) \\ &= m^{(K-1)} \circ \left(\mathrm{id}_{A^{\otimes i}} \otimes m \otimes \mathrm{id}_{A^{\otimes(K-1-i)}}\right). \end{aligned}$$

Thus, (13.14.6) is proven in Case 2.

Let us finally consider Case 3. In this case, we have neither $i \neq 0$ nor $i \neq K-1$. Hence, we have both $i = 0$ and $i = K-1$. Thus, $K - 1 = i = 0$, so that $K = 1$. Thus,

$$\begin{aligned} m^{(K)} = m^{(1)} &= m \circ \left(\operatorname{id}_A \otimes \underbrace{m^{(1-1)}}_{= m^{(0)} = \operatorname{id}_A} \right) \qquad \left(\text{by the recursive definition of } m^{(1)} \right) \\ &= m \circ \underbrace{(\operatorname{id}_A \otimes \operatorname{id}_A)}_{= \operatorname{id}_{A \otimes A}} = m. \end{aligned}$$

Compared with

$$\begin{aligned} & m^{(K-1)} \circ \left(\operatorname{id}_{A^{\otimes i}} \otimes m \otimes \operatorname{id}_{A^{\otimes (K-1-i)}} \right) \\ &= \underbrace{m^{(0)}}_{= \operatorname{id}_A} \circ \underbrace{\left(\operatorname{id}_{A^{\otimes 0}} \otimes m \otimes \operatorname{id}_{A^{\otimes 0}} \right)}_{= m} \qquad \left(\text{since } K - 1 = 0 \text{ and } i = 0 \text{ and } \underbrace{K-1}_{=0} - \underbrace{i}_{=0} = 0 - 0 = 0 \right) \\ &= \operatorname{id}_A \circ m = m, \end{aligned}$$

this yields $m^{(K)} = m^{(K-1)} \circ \left(\operatorname{id}_{A^{\otimes i}} \otimes m \otimes \operatorname{id}_{A^{\otimes (K-1-i)}} \right)$. Thus, (13.14.6) is proven in Case 3.

We have now proven (13.14.6) in each of the three Cases 1, 2 and 3. Since these three Cases cover all possibilities, this shows that (13.14.6) always holds (for all $0 \leq i \leq K-1$). In other words, the claim of Exercise 1.4.19(c) holds for $k = K$. This completes the induction step. The induction proof of the claim of Exercise 1.4.19(c) is therefore complete.

(d) Let $k \geq 1$. Applying Exercise 1.4.19(c) to $i = 0$, we obtain

$$m^{(k)} = m^{(k-1)} \circ \underbrace{\left(\operatorname{id}_{A^{\otimes 0}} \otimes m \otimes \operatorname{id}_{A^{\otimes (k-1-0)}} \right)}_{= m \otimes \operatorname{id}_{A^{\otimes (k-1-0)}} = m \otimes \operatorname{id}_{A^{\otimes (k-1)}}} = m^{(k-1)} \circ \left(m \otimes \operatorname{id}_{A^{\otimes (k-1)}} \right).$$

Applying Exercise 1.4.19(c) to $i = k-1$, we obtain

$$\begin{aligned} m^{(k)} &= m^{(k-1)} \circ \left(\operatorname{id}_{A^{\otimes (k-1)}} \otimes m \otimes \underbrace{\operatorname{id}_{A^{\otimes (k-1-(k-1))}}}_{= \operatorname{id}_{A^{\otimes 0}}} \right) = m^{(k-1)} \circ \left(\operatorname{id}_{A^{\otimes (k-1)}} \otimes \underbrace{m \otimes \operatorname{id}_{A^{\otimes 0}}}_{= m} \right) \\ &= m^{(k-1)} \circ \left(\operatorname{id}_{A^{\otimes (k-1)}} \otimes m \right). \end{aligned}$$

This solves Exercise 1.4.19(d).

13.15. Solution to Exercise 1.4.20.

Solution to Exercise 1.4.20. A solution for Exercise 1.4.20 can be obtained by reversing all arrows (and renaming A, m and $m^{(k)}$ by C, Δ and $\Delta^{(k)}$) in the element-free solution of Exercise 1.4.19 that we gave above.

13.16. Solution to Exercise 1.4.22.

Solution to Exercise 1.4.22. (a) We will solve Exercise 1.4.22(a) by induction over k:

The induction base ($k = 0$) requires us to prove that the map $m_H^{(0)} : H^{\otimes (0+1)} \to H$ is a $\mathbf{k}$-coalgebra homomorphism. But this is obvious, because $m_H^{(0)} = \operatorname{id}_H$ (by the definition of $m_H^{(0)}$).

Now, let us proceed to the induction step. Let K be a positive integer. We want to prove that the claim of Exercise 1.4.22(a) holds for $k = K$, assuming (as the induction hypothesis) that the claim of Exercise 1.4.22(a) holds for $k = K - 1$.

By the axioms of a bialgebra, we know that m_H is a $\mathbf{k}$-coalgebra homomorphism (since H is a $\mathbf{k}$-bialgebra).

We have $m_H^{(K)} = m_H \circ \left(\operatorname{id}_H \otimes m_H^{(K-1)} \right)$ (by the recursive definition of $m_H^{(K)}$). We know that $m_H^{(K-1)} : H^{\otimes ((K-1)+1)} \to H$ is a $\mathbf{k}$-coalgebra homomorphism (since the claim of Exercise 1.4.22(a) holds for $k = K-1$). Of course, $\operatorname{id}_H : H \to H$ is also a $\mathbf{k}$-coalgebra homomorphism. Exercise 1.3.6(b) (applied to $C = H$, $C' = H$, $D = H^{\otimes ((K-1)+1)}$, $D' = H$, $f = \operatorname{id}_H$ and $g = m_H^{(K-1)}$) thus yields that the map $\operatorname{id}_H \otimes m_H^{(K-1)}$:

$H \otimes H^{\otimes((K-1)+1)} \to H \otimes H$ is a **k**-coalgebra homomorphism. Since $H \otimes H^{\otimes((K-1)+1)} = H^{\otimes(((K-1)+1)+1)} = H^{\otimes(K+1)}$, this rewrites as follows: The map $\operatorname{id}_H \otimes m_H^{(K-1)} : H^{\otimes(K+1)} \to H \otimes H$ is a **k**-coalgebra homomorphism. Thus, $m_H \circ \left(\operatorname{id}_H \otimes m_H^{(K-1)}\right)$ is the composition of two **k**-coalgebra homomorphisms (namely, of m_H and $\operatorname{id}_H \otimes m_H^{(K-1)}$), hence a **k**-coalgebra homomorphism itself (since the composition of any two **k**-coalgebra homomorphisms is a **k**-coalgebra homomorphism). In other words, $m_H^{(K)}$ is a **k**-coalgebra homomorphism (since $m_H^{(K)} = m_H \circ \left(\operatorname{id}_H \otimes m_H^{(K-1)}\right)$). In other words, the claim of Exercise 1.4.22(a) holds for $k = K$. This completes the induction step. Thus, Exercise 1.4.22(a) is solved by induction.

(b) The solution of Exercise 1.4.22(b) can be obtained from the solution of Exercise 1.4.22(a) by "reversing all arrows". (The details of this are left to the reader.)

(d) Let $\ell \in \mathbb{N}$. Exercise 1.4.22(a) (applied to ℓ instead of k) yields that the map $m_H^{(\ell)} : H^{\otimes(\ell+1)} \to H$ is a **k**-coalgebra homomorphism.

For every **k**-coalgebra C, consider the map $\Delta_C^{(k)} : C \to C^{\otimes(k+1)}$ (this is the map $\Delta^{(k)}$ defined in Exercise 1.4.20). This map $\Delta_C^{(k)}$ is clearly functorial in C. By this we mean that if C and D are any two **k**-coalgebras, and $f : C \to D$ is any **k**-coalgebra homomorphism, then the diagram

$$\begin{tikzcd}
C \arrow[r, "f"] \arrow[d, "\Delta_C^{(k)}"] & D \arrow[d, "\Delta_D^{(k)}"] \\
C^{\otimes(k+1)} \arrow[r, "f^{\otimes(k+1)}"'] & D^{\otimes(k+1)}
\end{tikzcd}$$

commutes.[403] We can apply this to $C = H^{\otimes(\ell+1)}$, $D = H$ and $f = m_H^{(\ell)}$ (since $m_H^{(\ell)}$ is a **k**-coalgebra homomorphism). As a result, we conclude that the diagram

$$\begin{tikzcd}
H^{\otimes(\ell+1)} \arrow[r, "m_H^{(\ell)}"] \arrow[d, "\Delta_{H^{\otimes(\ell+1)}}^{(k)}"] & H \arrow[d, "\Delta_H^{(k)}"] \\
\left(H^{\otimes(\ell+1)}\right)^{\otimes(k+1)} \arrow[r, "\left(m_H^{(\ell)}\right)^{\otimes(k+1)}"'] & H^{\otimes(k+1)}
\end{tikzcd}$$

commutes. In other words, $\left(m_H^{(\ell)}\right)^{\otimes(k+1)} \circ \Delta_{H^{\otimes(\ell+1)}}^{(k)} = \Delta_H^{(k)} \circ m_H^{(\ell)}$. This solves Exercise 1.4.22(d).

(c) The solution of Exercise 1.4.22(c) can be obtained from the solution of Exercise 1.4.22(d) by "reversing all arrows".

13.17. **Solution to Exercise 1.4.23.** *Solution to Exercise 1.4.23.* If $k = 0$, then the statement of Exercise 1.4.23 is clearly true[404]. Hence, for the rest of this solution, we can WLOG assume that we don't have $k = 0$. Assume this.

[403] This can be proven by induction over k in a completely straightforward manner.

[404] *Proof.* Let $k = 0$. Then,

$$\begin{aligned} f_1 \star f_2 \star \cdots \star f_k &= f_1 \star f_2 \star \cdots \star f_0 = \left(\text{empty product in } \left(\operatorname{Hom}(C, A), \star\right)\right) \\ &= 1_{(\operatorname{Hom}(C,A),\star)} = u_A \epsilon_C \end{aligned}$$

and

$$\begin{aligned} m_A^{(k-1)} \circ (f_1 \otimes f_2 \otimes \cdots \otimes f_k) \circ \Delta_C^{(k-1)} &= \underbrace{m_A^{(0-1)}}_{=m_A^{(-1)}=u_A} \circ \underbrace{(f_1 \otimes f_2 \otimes \cdots \otimes f_0)}_{=(\text{empty tensor product})=\operatorname{id}_{\mathbf{k}}} \circ \underbrace{\Delta_C^{(k-1)}}_{=\Delta_C^{(-1)}=\epsilon_C} \\ &= u_A \circ \operatorname{id}_{\mathbf{k}} \circ \epsilon_C = u_A \epsilon_C. \end{aligned}$$

Hence, $f_1 \star f_2 \star \cdots \star f_k = u_A \epsilon_C = m_A^{(k-1)} \circ (f_1 \otimes f_2 \otimes \cdots \otimes f_k) \circ \Delta_C^{(k-1)}$, so that Exercise 1.4.23 is solved in the case when $k = 0$.

We shall show that

$$f_1 \star f_2 \star \cdots \star f_i = m_A^{(i-1)} \circ (f_1 \otimes f_2 \otimes \cdots \otimes f_i) \circ \Delta_C^{(i-1)} \tag{13.17.1}$$

for every $i \in \{1, 2, ..., k\}$.

Proof of (13.17.1)*:* We will prove (13.17.1) by induction over i:

Induction base: If $i = 1$, then (13.17.1) holds.[405] This completes the induction base.

Induction step: Let $I \in \{1, 2, ..., k\}$ be such that $I < k$. Assume that (13.17.1) holds for $i = I$. We must prove that (13.17.1) holds for $i = I + 1$.

Denote the $\mathbf{k}$-linear map $f_1 \otimes f_2 \otimes \cdots \otimes f_I : C^{\otimes I} \to A^{\otimes I}$ by g. Then, $g = f_1 \otimes f_2 \otimes \cdots \otimes f_I$. But we assumed that (13.17.1) holds for $i = I$. In other words,

$$f_1 \star f_2 \star \cdots \star f_I = m_A^{(I-1)} \circ \underbrace{(f_1 \otimes f_2 \otimes \cdots \otimes f_I)}_{=g} \circ \Delta_C^{(I-1)} = m_A^{(I-1)} \circ g \circ \Delta_C^{(I-1)}.$$

Now,

$$\begin{aligned}
f_1 \star f_2 \star \cdots \star f_{I+1} &= (f_1 \star f_2 \star \cdots \star f_I) \star f_{I+1} \\
&= m_A \circ \left(\underbrace{(f_1 \star f_2 \star \cdots \star f_I)}_{=m_A^{(I-1)} \circ g \circ \Delta_C^{(I-1)}} \otimes \underbrace{f_{I+1}}_{=\operatorname{id}_A \circ f_{I+1} \circ \operatorname{id}_C} \right) \circ \Delta_C \\
&\qquad \text{(by the definition of convolution)} \\
&= m_A \circ \underbrace{\left(\left(m_A^{(I-1)} \circ g \circ \Delta_C^{(I-1)} \right) \otimes (\operatorname{id}_A \circ f_{I+1} \circ \operatorname{id}_C) \right)}_{=\left(m_A^{(I-1)} \otimes \operatorname{id}_A \right) \circ (g \otimes f_{I+1}) \circ \left(\Delta_C^{(I-1)} \otimes \operatorname{id}_C \right)} \circ \Delta_C \\
&= \underbrace{m_A \circ \left(m_A^{(I-1)} \otimes \operatorname{id}_A \right)}_{\substack{=m_A^{(I)} \\ \text{(since } m_A^{(I)} = m_A \circ \left(m_A^{(I-1)} \otimes \operatorname{id}_A \right) \\ \text{(by Exercise 1.4.19(b), applied to } k=I\text{))}}} \circ \left(\underbrace{g}_{=f_1 \otimes f_2 \otimes \cdots \otimes f_I} \otimes f_{I+1} \right) \circ \underbrace{\left(\Delta_C^{(I-1)} \otimes \operatorname{id}_C \right) \circ \Delta_C}_{\substack{=\Delta_C^{(I)} \\ \text{(since } \Delta_C^{(I)} = \left(\Delta_C^{(I-1)} \otimes \operatorname{id}_C \right) \circ \Delta_C \\ \text{(by Exercise 1.4.20(b), applied to } k=I\text{))}}} \\
&= \underbrace{m_A^{(I)}}_{=m_A^{((I+1)-1)}} \circ \underbrace{((f_1 \otimes f_2 \otimes \cdots \otimes f_I) \otimes f_{I+1})}_{=f_1 \otimes f_2 \otimes \cdots \otimes f_{I+1}} \circ \underbrace{\Delta_C^{(I)}}_{=\Delta_C^{((I+1)-1)}} \\
&= m_A^{((I+1)-1)} \circ (f_1 \otimes f_2 \otimes \cdots \otimes f_{I+1}) \circ \Delta_C^{((I+1)-1)}.
\end{aligned}$$

In other words, (13.17.1) holds for $i = I + 1$. Thus, the induction step is complete. Therefore, (13.17.1) is proven by induction.

Recall that we don't have $k = 0$. Hence, $k > 0$, so that $k \in \{1, 2, ..., k\}$. Therefore, (13.17.1) (applied to $i = k$) yields

$$f_1 \star f_2 \star \cdots \star f_k = m_A^{(k-1)} \circ (f_1 \otimes f_2 \otimes \cdots \otimes f_k) \circ \Delta_C^{(k-1)}.$$

This solves Exercise 1.4.23.

---

[405] *Proof.* Assume that $i = 1$. Then, since $i = 1$, we have

$$f_1 \star f_2 \star \cdots \star f_i = f_1 \star f_2 \star \cdots \star f_1 = f_1.$$

Also, since $i = 1$, we have

$$m_A^{(i-1)} \circ (f_1 \otimes f_2 \otimes \cdots \otimes f_i) \circ \Delta_C^{(i-1)} = \underbrace{m_A^{(1-1)}}_{=m_A^{(0)} = \operatorname{id}_A} \circ \underbrace{(f_1 \otimes f_2 \otimes \cdots \otimes f_1)}_{=f_1} \circ \underbrace{\Delta_C^{(1-1)}}_{=\Delta_C^{(0)} = \operatorname{id}_C} = \operatorname{id}_A \circ f_1 \circ \operatorname{id}_C = f_1.$$

Thus, $f_1 \star f_2 \star \cdots \star f_i = f_1 = m_A^{(i-1)} \circ (f_1 \otimes f_2 \otimes \cdots \otimes f_i) \circ \Delta_C^{(i-1)}$. Hence, (13.17.1) holds, qed.

13.18. **Solution to Exercise 1.4.28.** *Solution to Exercise 1.4.28.* The solution of this exercise, of course, proceeds by inverting all the arrows in the proof of Proposition 1.4.10, but this is complicated by the fact that said proof first has to be rewritten in terms of arrows (i.e., freed of all uses of elements)[406]. We give such a solution to Exercise 1.4.28 further below – see the solution of Exercise 1.4.29(c).

Alternatively, solutions of Exercise 1.4.28 can be found in many places: e.g., [1, Thm. 2.1.4(iv)], [107, III.(3.5)], [149, Prop. I.7.1 2)].

13.19. **Solution to Exercise 1.4.29.** *Solution to Exercise 1.4.29.* (a) We shall give two solutions of Exercise 1.4.29(a). The first of these solutions will be a (completely straightforward) generalization of the proof of Proposition 1.4.10 (which proposition, as we will see in solving Exercise 1.4.29(c), is a particular case of this exercise). The second solution will be element-free (i.e., it will only manipulate linear maps, but never refer to elements of C or A), but essentially is just a result of rewriting the first solution in element-free terms. One of the advantages of an element-free solution is the ease of applying it to more general contexts (such as tensor categories), but also the possibility of straightforwardly "reversing the arrows" in such a solution and thus obtaining a solution to Exercise 1.4.29(b). (We will elaborate on this when we solve Exercise 1.4.29(b).)

First solution to Exercise 1.4.29(a). Let $\overline{r} = r^{\star(-1)}$. Then, $\overline{r} = r^{\star(-1)}$ is the $\star$-inverse of f. Thus, $\overline{r} \star r = 1_{(\operatorname{Hom}(C,A),\star)} = u_A \circ \epsilon_C$ and similarly $r \star \overline{r} = u_A \circ \epsilon_C$.

Since Δ_C is an algebra morphism, one has $\Delta_C(1_C) = 1_C \otimes 1_C$, and thus

$$(r \star \overline{r})(1_C) = \underbrace{r(1_C)}_{\substack{=1_A \\ \text{(since } r \text{ is a } \mathbf{k}\text{-algebra} \\ \text{homomorphism)}}} \overline{r}(1_C) = \overline{r}(1_C),$$

so that $\overline{r}(1_C) = \underbrace{(r \star \overline{r})}_{=u_A \circ \epsilon_C}(1_C) = (u_A \circ \epsilon_C)(1_C) = 1_A$.

But we need to prove that $r^{\star(-1)}$ is a $\mathbf{k}$-algebra anti-homomorphism. In other words, we need to prove that $\overline{r}$ is a $\mathbf{k}$-algebra anti-homomorphism (because $\overline{r} = r^{\star(-1)}$). In other words, we need to prove that $\overline{r}(1_C) = 1_A$ and that every $a \in C$ and $b \in C$ satisfy $\overline{r}(ab) = \overline{r}(b)\overline{r}(a)$. Since $\overline{r}(1_C) = 1_A$ is already proven, it thus remains to prove that every $a \in C$ and $b \in C$ satisfy $\overline{r}(ab) = \overline{r}(b)\overline{r}(a)$.

For this purpose, we shall **not** fix a and b. Instead, we consider $C \otimes C$ as a $\mathbf{k}$-coalgebra, and A as a $\mathbf{k}$-algebra. Then, $\operatorname{Hom}(C \otimes C, A)$ is an associative algebra with a convolution product $\circledast$ (to be distinguished from the convolution $\star$ on $\operatorname{Hom}(C, A)$), having two-sided identity element $u_A \epsilon_{C \otimes C}$. We define three $\mathbf{k}$-linear maps f, g and h from $C \otimes C$ to A as follows:

$$f(a \otimes b) = r(a) r(b) \qquad \text{for all } a \in C \text{ and } b \in C;$$
$$g(a \otimes b) = \overline{r}(b) \overline{r}(a) \qquad \text{for all } a \in C \text{ and } b \in C;$$
$$h(a \otimes b) = \overline{r}(ab) \qquad \text{for all } a \in C \text{ and } b \in C.$$

(These definitions make sense, since each of $r(a) r(b)$, $\overline{r}(b)\overline{r}(a)$ and $\overline{r}(ab)$ depends $\mathbf{k}$-bilinearly on (a, b).) Thus, f, g and h are three elements of $\operatorname{Hom}(C \otimes C, A)$. We shall now prove that

$$h \circledast f = u_A \epsilon_{C \otimes C} = f \circledast g. \tag{13.19.1}$$

Once this equality is proven, we will then obtain

$$\begin{aligned} h &= h \circledast (u_A \epsilon_{C \otimes C}) && \text{(since } u_A \epsilon_{C \otimes C} \text{ is the identity element of } (\operatorname{Hom}(C \otimes C, A), \circledast)) \\ &= h \circledast (f \circledast g) = (h \circledast f) \circledast g && \text{(by the associativity of the convolution } \circledast) \\ &= (u_A \epsilon_{C \otimes C}) \circledast g = g && \text{(since } u_A \epsilon_{C \otimes C} \text{ is the identity element of } (\operatorname{Hom}(C \otimes C, A), \circledast)). \end{aligned}$$

In order to prove (13.19.1), we evaluate the three maps $h \circledast f$, $u_A \epsilon_{C \otimes C}$ and $f \circledast g$ on pure tensors $a \otimes b \in C \otimes C$. We use the Sweedler notation in the form $\Delta(a) = \sum_{(a)} a_1 \otimes a_2$ and $\Delta(b) = \sum_{(b)} b_1 \otimes b_2$; thus, $\Delta(ab) = \sum_{(a),(b)} a_1 b_1 \otimes a_2 b_2$ (since Δ is a $\mathbf{k}$-algebra homomorphism). We then obtain

$$(u_A \epsilon_{C \otimes C})(a \otimes b) = u_A(\epsilon_C(a) \epsilon_C(b)) = u_A(\epsilon_C(ab)).$$

[406] Rewriting a proof in terms of arrows is usually an exercise in category theory (or, rather, category practice); see the First solution to Exercise 1.2.3 (or the solution to Exercise 1.5.6 further below) for how this is done (in a simple case).

Furthermore,

$$\begin{aligned}
(h \circledast f)(a \otimes b) &= \sum_{(a),(b)} h(a_1 \otimes b_1) f(a_2 \otimes b_2) = \sum_{(a),(b)} \overline{r}(a_1 b_1) \underbrace{r(a_2) r(b_2)}_{\substack{=r(a_2 b_2) \\ \text{(since } r \text{ is a } \mathbf{k}\text{-algebra} \\ \text{homomorphism)}}} \\
&= \sum_{(a),(b)} \overline{r}(a_1 b_1) r(a_2 b_2) = \underbrace{(\overline{r} \star r)}_{=u_A \circ \epsilon_C}(ab) \qquad \left(\text{since } \sum_{(a),(b)} a_1 b_1 \otimes a_2 b_2 = \Delta(ab)\right) \\
&= (u_A \circ \epsilon_C)(ab) = u_A(\epsilon_C(ab)) \qquad (13.19.2)
\end{aligned}$$

and

$$\begin{aligned}
(f \circledast g)(a \otimes b) &= \sum_{(a),(b)} f(a_1 \otimes b_1) g(a_2 \otimes b_2) = \sum_{(a),(b)} r(a_1) r(b_1) \overline{r}(b_2) \overline{r}(a_2) \\
&= \sum_{(a)} r(a_1) \underbrace{\left(\sum_{(b)} r(b_1) \overline{r}(b_2)\right)}_{\substack{=(r \star \overline{r})(b) = (u_A \circ \epsilon_C)(b) \\ \text{(since } r \star \overline{r} = u_A \circ \epsilon_C\text{)}}} \overline{r}(a_2) = \sum_{(a)} r(a_1) \underbrace{(u_A \circ \epsilon_C)(b) \overline{r}(a_2)}_{=\epsilon_C(b) \overline{r}(a_2)} \\
&= \sum_{(a)} r(a_1) \epsilon_C(b) \overline{r}(a_2) = \underbrace{\left(\sum_{(a)} r(a_1) \overline{r}(a_2)\right)}_{\substack{=(r \star \overline{r})(a) = (u_A \circ \epsilon_C)(a) \\ \text{(since } r \star \overline{r} = u_A \circ \epsilon_C\text{)}}} \epsilon_C(b) \\
&= (u_A \circ \epsilon_C)(a) \epsilon_C(b) = (u_A \circ \epsilon_C)(a) (u_A \circ \epsilon_C)(b) = (u_A \circ \epsilon_C)(ab) \\
&\qquad \text{(since } u_A \circ \epsilon_C \text{ is a } \mathbf{k}\text{-algebra homomorphism)} \\
&= u_A(\epsilon_C(ab)). \qquad (13.19.3)
\end{aligned}$$

These three results being all equal, we thus have shown that the maps $h \circledast f$, $u_A \epsilon_{C \otimes C}$ and $f \circledast g$ are equal on each pure tensor. Correspondingly, these maps must be identical (since they are **k**-linear). In other words, (13.19.1) holds. As explained above, this yields $h = g$. Thus, every $a \in C$ and $b \in C$ satisfy

$$\begin{aligned}
\overline{r}(ab) &= \underbrace{h}_{=g}(a \otimes b) \qquad \text{(since } h(a \otimes b) \text{ is defined to be } \overline{r}(ab)\text{)} \\
&= g(a \otimes b) = \overline{r}(b) \overline{r}(a) \qquad \text{(by the definition of } g(a \otimes b)\text{)}.
\end{aligned}$$

As we already know, this completes the solution of Exercise 1.4.29(a).

Second solution to Exercise 1.4.29(a). Let us first recall some linear-algebraic facts. One such fact states that if U, V, U' and V' are four **k**-modules and $x : U \to U'$ and $y : V \to V'$ are two **k**-linear maps, then

$$(y \otimes x) \circ T_{U,V} = T_{U',V'} \circ (x \otimes y). \qquad (13.19.4)$$

For every **k**-module U, let $\operatorname{kan}_{1,U} : U \to U \otimes \mathbf{k}$ and $\operatorname{kan}_{2,U} : U \to \mathbf{k} \otimes U$ be the canonical **k**-module isomorphisms. These two isomorphisms are related to each other via the equalities

$$\begin{aligned}
\operatorname{kan}_{2,U}^{-1} \circ T_{U,\mathbf{k}} &= \operatorname{kan}_{1,U}^{-1}, \qquad (13.19.5) \\
\operatorname{kan}_{1,U}^{-1} \circ T_{\mathbf{k},U} &= \operatorname{kan}_{2,U}^{-1}, \qquad (13.19.6) \\
T_{\mathbf{k},U} \circ \operatorname{kan}_{2,U} &= \operatorname{kan}_{1,U}, \qquad \text{and} \qquad (13.19.7) \\
T_{U,\mathbf{k}} \circ \operatorname{kan}_{1,U} &= \operatorname{kan}_{2,U} \qquad (13.19.8)
\end{aligned}$$

for every $\mathbf{k}$-module U. Moreover, every $\mathbf{k}$-modules U and V satisfy

$$\operatorname{id}_U \otimes \operatorname{kan}_{1,V} = \operatorname{kan}_{1,U\otimes V}, \tag{13.19.9}$$

$$\operatorname{kan}_{2,V} \otimes \operatorname{id}_U = \operatorname{kan}_{2,V\otimes U}, \tag{13.19.10}$$

$$\operatorname{id}_U \otimes \operatorname{kan}_{1,V}^{-1} = \operatorname{kan}_{1,U\otimes V}^{-1}, \tag{13.19.11}$$

$$\operatorname{kan}_{2,V}^{-1} \otimes \operatorname{id}_U = \operatorname{kan}_{2,V\otimes U}^{-1}, \tag{13.19.12}$$

$$\operatorname{id}_U \otimes \operatorname{kan}_{2,V} = \operatorname{kan}_{1,U} \otimes \operatorname{id}_V, \tag{13.19.13}$$

$$\operatorname{id}_U \otimes \operatorname{kan}_{2,V}^{-1} = \operatorname{kan}_{1,U}^{-1} \otimes \operatorname{id}_V . \tag{13.19.14}$$

Furthermore, if U and V are two $\mathbf{k}$-modules and $\alpha : U \to V$ is a $\mathbf{k}$-linear map, then

$$\operatorname{kan}_{1,V} \circ \alpha = (\alpha \otimes \operatorname{id}_{\mathbf{k}}) \circ \operatorname{kan}_{1,U}, \tag{13.19.15}$$

$$\operatorname{kan}_{2,V} \circ \alpha = (\operatorname{id}_{\mathbf{k}} \otimes \alpha) \circ \operatorname{kan}_{2,U}, \tag{13.19.16}$$

$$\operatorname{kan}_{1,V}^{-1} \circ (\alpha \otimes \operatorname{id}_{\mathbf{k}}) = \alpha \circ \operatorname{kan}_{1,U}^{-1}, \qquad \text{and} \tag{13.19.17}$$

$$\operatorname{kan}_{2,V}^{-1} \circ (\operatorname{id}_{\mathbf{k}} \otimes \alpha) = \alpha \circ \operatorname{kan}_{2,U}^{-1} . \tag{13.19.18}$$

Now, let us step to the actual solution of Exercise 1.4.29(a).

Let $\overline{r} = r^{\star(-1)}$. Then, $\overline{r} = r^{\star(-1)}$ is the $\star$-inverse of f. Thus, $\overline{r} \star r = 1_{(\operatorname{Hom}(C,A),\star)} = u_A \circ \epsilon_C$ and similarly $r \star \overline{r} = u_A \circ \epsilon_C$.

Recall that C is a $\mathbf{k}$-coalgebra. By the axioms of a $\mathbf{k}$-coalgebra, this shows that

$$(\Delta_C \otimes \operatorname{id}_C) \circ \Delta_C = (\operatorname{id}_C \otimes \Delta_C) \circ \Delta_C,$$
$$(\operatorname{id}_C \otimes \epsilon_C) \circ \Delta_C = \operatorname{kan}_{1,C};$$
$$(\epsilon_C \otimes \operatorname{id}_C) \circ \Delta_C = \operatorname{kan}_{2,C} .$$

Also, recall that C is a $\mathbf{k}$-algebra. By the axioms of a $\mathbf{k}$-algebra, this shows that

$$m_C \circ (m_C \otimes \operatorname{id}_C) = m_C \circ (\operatorname{id}_C \otimes m_C) ;$$
$$m_C \circ (\operatorname{id}_C \otimes u_C) = \operatorname{kan}_{1,C}^{-1};$$
$$m_C \circ (u_C \otimes \operatorname{id}_C) = \operatorname{kan}_{2,C}^{-1} .$$

Also, recall that C is a $\mathbf{k}$-bialgebra. Due to the axioms of a $\mathbf{k}$-bialgebra, this shows that Δ_C and ϵ_C are $\mathbf{k}$-algebra homomorphisms, and that m_C and u_C are $\mathbf{k}$-coalgebra homomorphisms.

Furthermore, A is a $\mathbf{k}$-algebra. By the axioms of a $\mathbf{k}$-algebra, this shows that

$$m_A \circ (m_A \otimes \operatorname{id}_A) = m_A \circ (\operatorname{id}_A \otimes m_A) ;$$
$$m_A \circ (\operatorname{id}_A \otimes u_A) = \operatorname{kan}_{1,A}^{-1};$$
$$m_A \circ (u_A \otimes \operatorname{id}_A) = \operatorname{kan}_{2,A}^{-1} .$$

Since ϵ_C is a $\mathbf{k}$-algebra homomorphism, we have $\epsilon_C \circ u_C = u_{\mathbf{k}} = \operatorname{id}_{\mathbf{k}}$.

Recall that $\operatorname{kan}_{1,\mathbf{k}}$ is the canonical $\mathbf{k}$-module isomorphism $\mathbf{k} \to \mathbf{k} \otimes \mathbf{k}$. Hence, $u_{C\otimes C} = (u_C \otimes u_C) \circ \operatorname{kan}_{1,\mathbf{k}}$ (by the definition of the $\mathbf{k}$-algebra $C \otimes C$). Also, $\operatorname{kan}_{1,\mathbf{k}} = \operatorname{kan}_{2,\mathbf{k}}$ (since each of $\operatorname{kan}_{1,\mathbf{k}}$ and $\operatorname{kan}_{2,\mathbf{k}}$ is the canonical $\mathbf{k}$-module isomorphism $\mathbf{k} \to \mathbf{k} \otimes \mathbf{k}$).

We know that $r : C \to A$ is a $\mathbf{k}$-algebra homomorphism. In other words, r is a $\mathbf{k}$-linear map satisfying $r \circ m_C = m_A \circ (r \otimes r)$ and $r \circ u_C = u_A$.

Now, we need to prove that $r^{\star(-1)}$ is a $\mathbf{k}$-algebra anti-homomorphism. In other words, we need to prove that $\overline{r}$ is a $\mathbf{k}$-algebra anti-homomorphism (since $\overline{r} = r^{\star(-1)}$). In other words, we need to prove that $\overline{r}$ satisfies the two equations $\overline{r} \circ m_C = m_A \circ (\overline{r} \otimes \overline{r}) \circ T_{C,C}$ and $\overline{r} \circ u_C = u_A$ (because these two equations are what makes $\overline{r}$ a $\mathbf{k}$-algebra anti-homomorphism, according to the definition of a "$\mathbf{k}$-algebra anti-homomorphism").

Let us first prove the equality $\overline{r} \circ u_C = u_A$.

We have $r \star \overline{r} = u_A \circ \epsilon_C$, so that $\underbrace{(r \star \overline{r})}_{=u_A \circ \epsilon_C} \circ u_C = u_A \circ \underbrace{\epsilon_C \circ u_C}_{=\mathrm{id}_{\mathbf{k}}} = u_A$. Hence,

$$\begin{aligned}
u_A &= \underbrace{(r \star \overline{r})}_{\substack{=m_A \circ (r \otimes \overline{r}) \circ \Delta_C \\ \text{(by the definition} \\ \text{of convolution)}}} \circ u_C = m_A \circ (r \otimes \overline{r}) \circ \underbrace{\Delta_C \circ u_C}_{\substack{=u_{C\otimes C} \\ \text{(since Δ_C is a $\mathbf{k}$-algebra} \\ \text{homomorphism)}}} \\
&= m_A \circ (r \otimes \overline{r}) \circ \underbrace{u_{C\otimes C}}_{=(u_C \otimes u_C) \circ \mathrm{kan}_{1,\mathbf{k}}} = m_A \circ \underbrace{(r \otimes \overline{r}) \circ (u_C \otimes u_C)}_{=(r \circ u_C) \otimes (\overline{r} \circ u_C)} \circ \mathrm{kan}_{1,\mathbf{k}} \\
&= m_A \circ \left(\underbrace{(r \circ u_C)}_{=u_A = u_A \circ \mathrm{id}_{\mathbf{k}}} \otimes \underbrace{(\overline{r} \circ u_C)}_{=\mathrm{id}_A \circ (\overline{r} \circ u_C)} \right) \circ \underbrace{\mathrm{kan}_{1,\mathbf{k}}}_{=\mathrm{kan}_{2,\mathbf{k}}} = m_A \circ \underbrace{((u_A \circ \mathrm{id}_{\mathbf{k}}) \otimes (\mathrm{id}_A \circ (\overline{r} \circ u_C)))}_{=(u_A \otimes \mathrm{id}_A) \circ (\mathrm{id}_{\mathbf{k}} \otimes (\overline{r} \circ u_C))} \circ \mathrm{kan}_{2,\mathbf{k}} \\
&= \underbrace{m_A \circ (u_A \otimes \mathrm{id}_A)}_{=\mathrm{kan}_{2,A}^{-1}} \circ \underbrace{(\mathrm{id}_{\mathbf{k}} \otimes (\overline{r} \circ u_C)) \circ \mathrm{kan}_{2,\mathbf{k}}}_{\substack{=\mathrm{kan}_{2,A} \circ (\overline{r} \circ u_C) \\ \text{(since (13.19.16) (applied to $U=\mathbf{k}$,} \\ \text{$V=A$ and $\alpha=\overline{r}\circ u_C$) yields} \\ \mathrm{kan}_{2,A} \circ (\overline{r} \circ u_C) = (\mathrm{id}_{\mathbf{k}} \otimes (\overline{r} \circ u_C)) \circ \mathrm{kan}_{2,\mathbf{k}}\text{)}}} = \underbrace{\mathrm{kan}_{2,A}^{-1} \circ \mathrm{kan}_{2,A}}_{=\mathrm{id}_A} \circ (\overline{r} \circ u_C) \\
&= \overline{r} \circ u_C.
\end{aligned}$$

Hence, $\overline{r} \circ u_C = u_A$.

Now, it remains to prove the equality $\overline{r} \circ m_C = m_A \circ (\overline{r} \otimes \overline{r}) \circ T_{C,C}$. This is harder. Let us first make some preparations.

Recall that $C \otimes C$ is a $\mathbf{k}$-coalgebra (since C is a $\mathbf{k}$-coalgebra), and A is an $\mathbf{k}$-algebra. Thus, $\mathrm{Hom}(C \otimes C, A)$ is an associative algebra with respect to convolution. We shall denote the convolution on $\mathrm{Hom}(C \otimes C, A)$ by $\circledast$ rather than by $\star$ (in order to distinguish it from the convolution $\star$ on $\mathrm{Hom}(C, A)$). The algebra $(\mathrm{Hom}(C \otimes C, A), \star)$ has two-sided identity element $u_A \circ \epsilon_{C\otimes C}$.

We define a $\mathbf{k}$-linear map $f : C \otimes C \to A$ by $f = m_A \circ (r \otimes r)$.

We define a $\mathbf{k}$-linear map $g : C \otimes C \to A$ by $g = m_A \circ T_{A,A} \circ (\overline{r} \otimes \overline{r})$.

We define a $\mathbf{k}$-linear map $h : C \otimes C \to A$ by $h = \overline{r} \circ m_C$.

Clearly, f, g and h are three elements of $\mathrm{Hom}(C \otimes C, A)$. Our next goal is to prove that

$$h \circledast f = u_A \circ \epsilon_{C\otimes C} = f \circledast g. \tag{13.19.19}$$

Once this equality is proven, we will then obtain

$$\begin{aligned}
h &= h \circledast \underbrace{(u_A \circ \epsilon_{C\otimes C})}_{\substack{=f \circledast g \\ \text{(by (13.19.19))}}} \\
&\qquad \text{(since $u_A \circ \epsilon_{C\otimes C}$ is the identity element of $(\mathrm{Hom}(C \otimes C, A), \circledast)$)} \\
&= h \circledast (f \circledast g) = \underbrace{(h \circledast f)}_{\substack{=u_A \circ \epsilon_{C\otimes C} \\ \text{(by (13.19.19))}}} \circledast g \qquad \text{(by the associativity of the convolution $\circledast$)} \\
&= (u_A \circ \epsilon_{C\otimes C}) \circledast g = g \tag{13.19.20} \\
&\qquad \text{(since $u_A \circ \epsilon_{C\otimes C}$ is the identity element of $(\mathrm{Hom}(C \otimes C, A), \circledast)$)}.
\end{aligned}$$

So let us concentrate on proving (13.19.19). Let us first notice that

$$u_A \circ \epsilon_C \circ m_C = u_A \circ \epsilon_{C\otimes C}. \tag{13.19.21}$$

[407]

Next, we make the following observations:

[407] *Proof of (13.19.21):* We know that m_C is a $\mathbf{k}$-coalgebra homomorphism. Hence, $\epsilon_C \circ m_C = \epsilon_{C\otimes C}$. Thus, $u_A \circ \underbrace{\epsilon_C \circ m_C}_{=\epsilon_{C\otimes C}} = u_A \circ \epsilon_{C\otimes C}$. This proves (13.19.21).

- First, we notice that

$$h \circledast f = m_A \circ (h \otimes f) \circ (\mathrm{id}_C \otimes T_{C,C} \otimes \mathrm{id}_C) \circ (\Delta_C \otimes \Delta_C) . \tag{13.19.22}$$

[408]

- Next, we notice that

$$\begin{aligned} & m_A \circ (h \otimes f) \circ (\mathrm{id}_C \otimes T_{C,C} \otimes \mathrm{id}_C) \circ (\Delta_C \otimes \Delta_C) \\ & = m_A \circ ((\overline{r} \circ m_C) \otimes (m_A \circ (r \otimes r))) \circ (\mathrm{id}_C \otimes T_{C,C} \otimes \mathrm{id}_C) \circ (\Delta_C \otimes \Delta_C) . \end{aligned} \tag{13.19.23}$$

[409]

- Next, we have

$$\begin{aligned} & m_A \circ ((\overline{r} \circ m_C) \otimes (m_A \circ (r \otimes r))) \circ (\mathrm{id}_C \otimes T_{C,C} \otimes \mathrm{id}_C) \circ (\Delta_C \otimes \Delta_C) \\ & = m_A \circ ((\overline{r} \circ m_C) \otimes (r \circ m_C)) \circ (\mathrm{id}_C \otimes T_{C,C} \otimes \mathrm{id}_C) \circ (\Delta_C \otimes \Delta_C) \end{aligned} \tag{13.19.24}$$

[410]

- Furthermore, we have

$$\begin{aligned} & m_A \circ ((\overline{r} \circ m_C) \otimes (r \circ m_C)) \circ (\mathrm{id}_C \otimes T_{C,C} \otimes \mathrm{id}_C) \circ (\Delta_C \otimes \Delta_C) \\ & = m_A \circ (\overline{r} \otimes r) \circ \Delta_C \circ m_C \end{aligned} \tag{13.19.25}$$

[411]

- Moreover, we have

$$m_A \circ (\overline{r} \otimes r) \circ \Delta_C \circ m_C = (\overline{r} \star r) \circ m_C . \tag{13.19.26}$$

[408] *Proof of (13.19.22):* By the definition of the convolution $h \circledast f$, we have

$$h \circledast f = m_A \circ (h \otimes f) \circ \underbrace{\Delta_{C \otimes C}}_{\substack{= (\mathrm{id}_C \otimes T_{C,C} \otimes \mathrm{id}_C) \circ (\Delta_C \otimes \Delta_C) \\ \text{(by the definition of the } \mathbf{k}\text{-coalgebra } C \otimes C)}} = m_A \circ (h \otimes f) \circ \left(\mathrm{id}_C \otimes T_{C,C} \otimes \mathrm{id}_C\right) \circ (\Delta_C \otimes \Delta_C) .$$

This proves (13.19.22).

[409] *Proof of (13.19.23):* We have

$$\begin{aligned} & m_A \circ \left(\underbrace{h}_{= \overline{r} \circ m_C} \otimes \underbrace{f}_{= m_A \circ (r \otimes r)} \right) \circ \left(\mathrm{id}_C \otimes T_{C,C} \otimes \mathrm{id}_C\right) \circ (\Delta_C \otimes \Delta_C) \\ & = m_A \circ ((\overline{r} \circ m_C) \otimes (m_A \circ (r \otimes r))) \circ \left(\mathrm{id}_C \otimes T_{C,C} \otimes \mathrm{id}_C\right) \circ (\Delta_C \otimes \Delta_C) . \end{aligned}$$

This proves (13.19.23).

[410] *Proof of (13.19.24):* We have

$$\begin{aligned} & m_A \circ \left((\overline{r} \circ m_C) \otimes \underbrace{(r \circ m_C)}_{= m_A \circ (r \otimes r)} \right) \circ \left(\mathrm{id}_C \otimes T_{C,C} \otimes \mathrm{id}_C\right) \circ (\Delta_C \otimes \Delta_C) \\ & = m_A \circ ((\overline{r} \circ m_C) \otimes (m_A \circ (r \otimes r))) \circ \left(\mathrm{id}_C \otimes T_{C,C} \otimes \mathrm{id}_C\right) \circ (\Delta_C \otimes \Delta_C) . \end{aligned}$$

This proves (13.19.24).

[411] *Proof of (13.19.25):* We have

$$\begin{aligned} & m_A \circ \underbrace{((\overline{r} \circ m_C) \otimes (r \circ m_C))}_{= (\overline{r} \otimes r) \circ (m_C \otimes m_C)} \circ \left(\mathrm{id}_C \otimes T_{C,C} \otimes \mathrm{id}_C\right) \circ (\Delta_C \otimes \Delta_C) \\ & = m_A \circ (\overline{r} \otimes r) \circ \underbrace{(m_C \otimes m_C) \circ \left(\mathrm{id}_C \otimes T_{C,C} \otimes \mathrm{id}_C\right)}_{\substack{= m_{C \otimes C} \\ \text{(since the definition of the } \mathbf{k}\text{-algebra } C \otimes C \\ \text{yields } m_{C \otimes C} = (m_C \otimes m_C) \circ (\mathrm{id}_C \otimes T_{C,C} \otimes \mathrm{id}_C))}} \circ (\Delta_C \otimes \Delta_C) \\ & = m_A \circ (\overline{r} \otimes r) \circ \underbrace{m_{C \otimes C} \circ (\Delta_C \otimes \Delta_C)}_{\substack{= \Delta_C \circ m_C \\ \text{(since } \Delta_C : C \to C \otimes C \text{ is a } \mathbf{k}\text{-algebra homomorphism)}}} = m_A \circ (\overline{r} \otimes r) \circ \Delta_C \circ m_C . \end{aligned}$$

This proves (13.19.25).

[412]

- Finally, we have

$$(\overline{r}\star r)\circ m_C = u_A\circ \epsilon_C\circ m_C. \tag{13.19.27}$$

[413]

Now, we have

$$\begin{aligned} h\circledast f &= m_A\circ (h\otimes f)\circ (\mathrm{id}_C \otimes T_{C,C}\otimes \mathrm{id}_C)\circ (\Delta_C\otimes \Delta_C) \qquad (\text{by } (13.19.22))\\ &= m_A\circ ((\overline{r}\circ m_C)\otimes (m_A\circ (r\otimes r)))\circ (\mathrm{id}_C \otimes T_{C,C}\otimes \mathrm{id}_C)\circ (\Delta_C\otimes \Delta_C) \qquad (\text{by } (13.19.23))\\ &= m_A\circ ((\overline{r}\circ m_C)\otimes (r\circ m_C))\circ (\mathrm{id}_C \otimes T_{C,C}\otimes \mathrm{id}_C)\circ (\Delta_C\otimes \Delta_C) \qquad (\text{by } (13.19.24))\\ &= m_A\circ (\overline{r}\otimes r)\circ \Delta_C\circ m_C \qquad (\text{by } (13.19.25))\\ &= (\overline{r}\star r)\circ m_C \qquad (\text{by } (13.19.26))\\ &= u_A\circ \epsilon_C\circ m_C \qquad (\text{by } (13.19.27))\\ &= u_A\circ \epsilon_{C\otimes C} \qquad (\text{by } (13.19.21)). \end{aligned}$$

The first equality in (13.19.19) is thus proven.

We shall next show the second equality in (13.19.19).

For every $k\in\mathbb{N}$, let us define the map $m^{(k)}: A^{\otimes(k+1)}\to A$ as in Exercise 1.4.19. We recall that these maps $m^{(k)}$ are defined by induction over k, with the induction base $m^{(0)} = \mathrm{id}_A$, and with the induction step

$$m^{(k)} = m_A\circ \left(\mathrm{id}_A \otimes m^{(k-1)}\right) \qquad \text{for every } k\geq 1. \tag{13.19.28}$$

[414] Out of these maps, we shall only need $m^{(0)}$, $m^{(1)}$, $m^{(2)}$ and $m^{(3)}$. These satisfy the following formulae:

$$m^{(0)} = \mathrm{id}_A; \tag{13.19.29}$$

$$m^{(1)} = m_A; \tag{13.19.30}$$

$$m^{(2)} = m_A\circ (m_A\otimes \mathrm{id}_A); \tag{13.19.31}$$

$$m^{(2)} = m_A\circ (\mathrm{id}_A \otimes m_A); \tag{13.19.32}$$

$$m^{(3)} = m_A\circ (m_A\otimes m_A); \tag{13.19.33}$$

$$m^{(3)} = m^{(2)}\circ (\mathrm{id}_A \otimes m_A\otimes \mathrm{id}_A). \tag{13.19.34}$$

[415]

Now, we make the following observations:

[412] *Proof of (13.19.26):* The definition of the convolution $\overline{r}\star r$ yields $\overline{r}\star r = m_A\circ (\overline{r}\otimes r)\circ \Delta_C$. Hence, $\underbrace{(\overline{r}\star r)}_{=m_A\circ (\overline{r}\otimes r)\circ \Delta_C}\circ m_C = m_A\circ (\overline{r}\otimes r)\circ \Delta_C\circ m_C$. This proves (13.19.26).

[413] *Proof of (13.19.27):* We have $\underbrace{(\overline{r}\star r)}_{=u_A\circ \epsilon_C}\circ m_C = u_A\circ \epsilon_C\circ m_C$. Thus, (13.19.27) is proven.

[414] This is precisely the definition of these maps given in Exercise 1.4.19, with the only difference that m_A was denoted by m in that exercise.

[415] Here are proofs for these formulae:

Proof of (13.19.29): The formula (13.19.29) follows immediately from the definition of $m^{(0)}$.

Proof of (13.19.30): Applying (13.19.28) to $k=1$, we obtain $m^{(1)} = m_A\circ \left(\mathrm{id}_A\otimes \underbrace{m^{(1-1)}}_{=m^{(0)}=\mathrm{id}_A}\right) = m_A\circ \underbrace{(\mathrm{id}_A\otimes \mathrm{id}_A)}_{=\mathrm{id}_{A\otimes A}} = m_A$. This proves (13.19.30).

Proof of (13.19.32): Applying (13.19.28) to $k=2$, we obtain $m^{(2)} = m_A\circ \left(\mathrm{id}_A\otimes \underbrace{m^{(2-1)}}_{\substack{=m^{(1)}=m_A\\ (\text{by } (13.19.30))}}\right) = m_A\circ (\mathrm{id}_A\otimes m_A)$. This proves (13.19.32).

Proof of (13.19.31): From (13.19.32), we obtain $m^{(2)} = m_A\circ (\mathrm{id}_A\otimes m_A) = m_A\circ (m_A\otimes \mathrm{id}_A)$ (since $m_A\circ (m_A\otimes \mathrm{id}_A) = m_A\circ (\mathrm{id}_A\otimes m_A)$). This proves (13.19.31).

- We have

$$f \circledast g = m_A \circ (f \otimes g) \circ (\mathrm{id}_C \otimes T_{C,C} \otimes \mathrm{id}_C) \circ (\Delta_C \otimes \Delta_C) . \tag{13.19.35}$$

[416]

- We have

$$\begin{aligned} & m_A \circ (f \otimes g) \circ (\mathrm{id}_C \otimes T_{C,C} \otimes \mathrm{id}_C) \circ (\Delta_C \otimes \Delta_C) \\ & = m^{(3)} \circ (r \otimes r \otimes \overline{r} \otimes \overline{r}) \circ (\mathrm{id}_C \otimes \mathrm{id}_C \otimes T_{C,C}) \circ (\mathrm{id}_C \otimes T_{C,C} \otimes \mathrm{id}_C) \circ (\Delta_C \otimes \Delta_C) . \end{aligned} \tag{13.19.36}$$

[417]

Proof of (13.19.33): Applying (13.19.28) to $k = 3$, we obtain

$$\begin{aligned} m^{(3)} &= m_A \circ \left(\underbrace{\mathrm{id}_A}_{= \mathrm{id}_A \circ \mathrm{id}_A} \otimes \underbrace{m^{(3-1)}}_{\substack{= m^{(2)} = m_A \circ (\mathrm{id}_A \otimes m_A) \\ \text{(by (13.19.32))}}} \right) = m_A \circ \underbrace{((\mathrm{id}_A \circ \mathrm{id}_A) \otimes (m_A \circ (\mathrm{id}_A \otimes m_A)))}_{= (\mathrm{id}_A \otimes m_A) \circ (\mathrm{id}_A \otimes (\mathrm{id}_A \otimes m_A))} \\ &= \underbrace{m_A \circ (\mathrm{id}_A \otimes m_A)}_{\substack{= m_A \circ (m_A \otimes \mathrm{id}_A) \\ \text{(since } m_A \circ (m_A \otimes \mathrm{id}_A) = m_A \circ (\mathrm{id}_A \otimes m_A))}} \circ \underbrace{(\mathrm{id}_A \otimes (\mathrm{id}_A \otimes m_A))}_{\substack{= \mathrm{id}_A \otimes \mathrm{id}_A \otimes m_A \\ = \mathrm{id}_{A \otimes A} \otimes m_A}} = m_A \circ \underbrace{(m_A \otimes \mathrm{id}_A) \circ (\mathrm{id}_{A \otimes A} \otimes m_A)}_{= \left(m_A \circ \mathrm{id}_{A \otimes A} \right) \otimes (\mathrm{id}_A \circ m_A)} \\ &= m_A \circ \left(\underbrace{(m_A \circ \mathrm{id}_{A \otimes A})}_{= m_A} \otimes \underbrace{(\mathrm{id}_A \circ m_A)}_{= m_A} \right) = m_A \circ (m_A \otimes m_A) . \end{aligned}$$

This proves (13.19.33).

Proof of (13.19.34): Applying (13.19.28) to $k = 3$, we obtain

$$\begin{aligned} m^{(3)} &= m_A \circ \left(\underbrace{\mathrm{id}_A}_{= \mathrm{id}_A \circ \mathrm{id}_A} \otimes \underbrace{m^{(3-1)}}_{\substack{= m^{(2)} = m_A \circ (m_A \otimes \mathrm{id}_A) \\ \text{(by (13.19.31))}}} \right) = m_A \circ \underbrace{((\mathrm{id}_A \circ \mathrm{id}_A) \otimes (m_A \circ (m_A \otimes \mathrm{id}_A)))}_{= (\mathrm{id}_A \otimes m_A) \circ (\mathrm{id}_A \otimes (m_A \otimes \mathrm{id}_A))} \\ &= \underbrace{m_A \circ (\mathrm{id}_A \otimes m_A)}_{\substack{= m^{(2)} \\ \text{(by (13.19.32))}}} \circ \underbrace{(\mathrm{id}_A \otimes (m_A \otimes \mathrm{id}_A))}_{= \mathrm{id}_A \otimes m_A \otimes \mathrm{id}_A} = m^{(2)} \circ (\mathrm{id}_A \otimes m_A \otimes \mathrm{id}_A) . \end{aligned}$$

This proves (13.19.34).

[416] *Proof of (13.19.35):* By the definition of the convolution $f \circledast g$, we have

$$f \circledast g = m_A \circ (f \otimes g) \circ \underbrace{\Delta_{C \otimes C}}_{\substack{= \left(\mathrm{id}_C \otimes T_{C,C} \otimes \mathrm{id}_C \right) \circ (\Delta_C \otimes \Delta_C) \\ \text{(by the definition of the } \mathbf{k}\text{-coalgebra } C \otimes C)}} = m_A \circ (f \otimes g) \circ \left(\mathrm{id}_C \otimes T_{C,C} \otimes \mathrm{id}_C \right) \circ (\Delta_C \otimes \Delta_C) .$$

This proves (13.19.35).

[417] *Proof of (13.19.36):* The equality (13.19.4) (applied to $U = C$, $V = C$, $U' = A$, $V' = A$, $x = \overline{r}$ and $y = \overline{r}$) yields $(\overline{r} \otimes \overline{r}) \circ T_{C,C} = T_{A,A} \circ (\overline{r} \otimes \overline{r})$. Hence, $g = m_A \circ \underbrace{T_{A,A} \circ (\overline{r} \otimes \overline{r})}_{= (\overline{r} \otimes \overline{r}) \circ T_{C,C}} = m_A \circ \left((\overline{r} \otimes \overline{r}) \circ T_{C,C} \right)$. Now,

$$\begin{aligned} & m_A \circ \left(\underbrace{f}_{= m_A \circ (r \otimes r)} \otimes \underbrace{g}_{= m_A \circ \left((\overline{r} \otimes \overline{r}) \circ T_{C,C} \right)} \right) \\ & = m_A \circ \underbrace{\left((m_A \circ (r \otimes r)) \otimes \left(m_A \circ \left((\overline{r} \otimes \overline{r}) \circ T_{C,C} \right) \right) \right)}_{= (m_A \otimes m_A) \circ \left((r \otimes r) \otimes \left((\overline{r} \otimes \overline{r}) \circ T_{C,C} \right) \right)} = \underbrace{m_A \circ (m_A \otimes m_A)}_{\substack{= m^{(3)} \\ \text{(by (13.19.33))}}} \circ \left(\underbrace{(r \otimes r)}_{= (r \otimes r) \circ \mathrm{id}_{C \otimes C}} \otimes \left((\overline{r} \otimes \overline{r}) \circ T_{C,C} \right) \right) \\ & = m^{(3)} \circ \underbrace{\left(((r \otimes r) \circ \mathrm{id}_{C \otimes C}) \otimes \left((\overline{r} \otimes \overline{r}) \circ T_{C,C} \right) \right)}_{= ((r \otimes r) \otimes (\overline{r} \otimes \overline{r})) \circ \left(\mathrm{id}_{C \otimes C} \otimes T_{C,C} \right)} = m^{(3)} \circ \underbrace{((r \otimes r) \otimes (\overline{r} \otimes \overline{r}))}_{= r \otimes r \otimes \overline{r} \otimes \overline{r}} \circ \left(\underbrace{\mathrm{id}_{C \otimes C}}_{= \mathrm{id}_C \otimes \mathrm{id}_C} \otimes T_{C,C} \right) \\ & = m^{(3)} \circ (r \otimes r \otimes \overline{r} \otimes \overline{r}) \circ \left(\mathrm{id}_C \otimes \mathrm{id}_C \otimes T_{C,C} \right) . \end{aligned}$$

- We have

$$m^{(3)} \circ (r \otimes r \otimes \overline{r} \otimes \overline{r}) \circ (\mathrm{id}_C \otimes \mathrm{id}_C \otimes T_{C,C}) \circ (\mathrm{id}_C \otimes T_{C,C} \otimes \mathrm{id}_C) \circ (\Delta_C \otimes \Delta_C)$$
$$= m^{(2)} \circ (r \otimes (m_A \circ (r \otimes \overline{r}) \circ \Delta_C) \otimes \overline{r}) \circ (\mathrm{id}_C \otimes T_{C,C}) \circ (\Delta_C \otimes \mathrm{id}_C) . \tag{13.19.37}$$

[418]

- We have

$$m^{(2)} \circ (r \otimes (m_A \circ (r \otimes \overline{r}) \circ \Delta_C) \otimes \overline{r}) \circ (\mathrm{id}_C \otimes T_{C,C}) \circ (\Delta_C \otimes \mathrm{id}_C)$$
$$= m^{(2)} \circ (r \otimes (u_A \circ \epsilon_C) \otimes \overline{r}) \circ (\mathrm{id}_C \otimes T_{C,C}) \circ (\Delta_C \otimes \mathrm{id}_C) . \tag{13.19.39}$$

Hence,

$$\underbrace{m_A \circ (f \otimes g)}_{= m^{(3)} \circ (r \otimes r \otimes \overline{r} \otimes \overline{r}) \circ \left(\mathrm{id}_C \otimes \mathrm{id}_C \otimes T_{C,C}\right)} \circ \left(\mathrm{id}_C \otimes T_{C,C} \otimes \mathrm{id}_C\right) \circ (\Delta_C \otimes \Delta_C)$$
$$= m^{(3)} \circ (r \otimes r \otimes \overline{r} \otimes \overline{r}) \circ \left(\mathrm{id}_C \otimes \mathrm{id}_C \otimes T_{C,C}\right) \circ \left(\mathrm{id}_C \otimes T_{C,C} \otimes \mathrm{id}_C\right) \circ (\Delta_C \otimes \Delta_C) .$$

This proves (13.19.36).

[418] *Proof of (13.19.37):* We first notice that

$$\left(\mathrm{id}_C \otimes \mathrm{id}_C \otimes T_{C,C}\right) \circ \left(\mathrm{id}_C \otimes T_{C,C} \otimes \mathrm{id}_C\right) = \mathrm{id}_C \otimes T_{C,C\otimes C}. \tag{13.19.38}$$

In fact, the two maps $\left(\mathrm{id}_C \otimes \mathrm{id}_C \otimes T_{C,C}\right) \circ \left(\mathrm{id}_C \otimes T_{C,C} \otimes \mathrm{id}_C\right)$ and $\mathrm{id}_C \otimes T_{C,C\otimes C}$ are $\mathbf{k}$-linear, and are equal to each other on pure tensors (in fact, each of them sends every pure tensor $a \otimes b \otimes c \otimes d \in C \otimes C \otimes C \otimes C$ to $a \otimes c \otimes d \otimes b$); therefore, they must be identical, so that (13.19.38) holds.

Now,

$$\underbrace{m^{(3)}}_{\substack{= m^{(2)} \circ (\mathrm{id}_A \otimes m_A \otimes \mathrm{id}_A) \\ \text{(by (13.19.34))}}} \circ (r \otimes r \otimes \overline{r} \otimes \overline{r}) \circ \underbrace{\left(\mathrm{id}_C \otimes \mathrm{id}_C \otimes T_{C,C}\right) \circ \left(\mathrm{id}_C \otimes T_{C,C} \otimes \mathrm{id}_C\right)}_{\substack{= \mathrm{id}_C \otimes T_{C,C\otimes C} \\ \text{(by (13.19.38))}}} \circ \left(\underbrace{\Delta_C}_{= \mathrm{id}_{C\otimes C} \circ \Delta_C} \otimes \underbrace{\Delta_C}_{= \Delta_C \circ \mathrm{id}_C} \right)$$
$$= m^{(2)} \circ (\mathrm{id}_A \otimes m_A \otimes \mathrm{id}_A) \circ (r \otimes r \otimes \overline{r} \otimes \overline{r}) \circ \left(\mathrm{id}_C \otimes T_{C,C\otimes C}\right) \circ \underbrace{\left((\mathrm{id}_{C\otimes C} \circ \Delta_C) \otimes (\Delta_C \circ \mathrm{id}_C)\right)}_{= \left(\mathrm{id}_{C\otimes C} \otimes \Delta_C\right) \circ (\Delta_C \otimes \mathrm{id}_C)}$$
$$= m^{(2)} \circ (\mathrm{id}_A \otimes m_A \otimes \mathrm{id}_A) \circ (r \otimes r \otimes \overline{r} \otimes \overline{r}) \circ \left(\mathrm{id}_C \otimes T_{C,C\otimes C}\right) \circ (\mathrm{id}_{C\otimes C} \otimes \Delta_C) \circ (\Delta_C \otimes \mathrm{id}_C) .$$

Since

$$(\mathrm{id}_A \otimes m_A \otimes \mathrm{id}_A) \circ \underbrace{(r \otimes r \otimes \overline{r} \otimes \overline{r})}_{= r \otimes (r \otimes \overline{r}) \otimes \overline{r}}$$
$$= (\mathrm{id}_A \otimes m_A \otimes \mathrm{id}_A) \circ (r \otimes (r \otimes \overline{r}) \otimes \overline{r}) = \underbrace{(\mathrm{id}_A \circ r)}_{= r} \otimes (m_A \circ (r \otimes \overline{r})) \otimes \underbrace{(\mathrm{id}_A \circ \overline{r})}_{= \overline{r}} = r \otimes (m_A \circ (r \otimes \overline{r})) \otimes \overline{r}$$

and

$$\left(\mathrm{id}_C \otimes T_{C,C\otimes C}\right) \circ \left(\underbrace{\mathrm{id}_{C\otimes C}}_{= \mathrm{id}_C \otimes \mathrm{id}_C} \otimes \Delta_C \right)$$
$$= \left(\mathrm{id}_C \otimes T_{C,C\otimes C}\right) \circ (\mathrm{id}_C \otimes \mathrm{id}_C \otimes \Delta_C) = \left(\mathrm{id}_C \otimes T_{C,C\otimes C}\right) \circ (\mathrm{id}_C \otimes (\mathrm{id}_C \otimes \Delta_C))$$
$$= (\mathrm{id}_C \circ \mathrm{id}_C) \otimes \underbrace{\left(T_{C,C\otimes C} \circ (\mathrm{id}_C \otimes \Delta_C)\right)}_{\substack{= (\Delta_C \otimes \mathrm{id}_C) \circ T_{C,C} \\ \text{(since } (\Delta_C \otimes \mathrm{id}_C) \circ T_{C,C} = T_{C,C\otimes C} \circ (\mathrm{id}_C \otimes \Delta_C) \\ \text{(by (13.19.4), applied to } U = C,\ V = C, \\ U' = C,\ V' = C\otimes C,\ x = \mathrm{id}_C \text{ and } y = \Delta_C \text{))}}}$$
$$= (\mathrm{id}_C \circ \mathrm{id}_C) \otimes \left((\Delta_C \otimes \mathrm{id}_C) \circ T_{C,C}\right) = \underbrace{(\mathrm{id}_C \otimes (\Delta_C \otimes \mathrm{id}_C))}_{= \mathrm{id}_C \otimes \Delta_C \otimes \mathrm{id}_C} \circ \left(\mathrm{id}_C \otimes T_{C,C}\right) = (\mathrm{id}_C \otimes \Delta_C \otimes \mathrm{id}_C) \circ \left(\mathrm{id}_C \otimes T_{C,C}\right) ,$$

[419]

- We have

$$m^{(2)} \circ (r \otimes (u_A \circ \epsilon_C) \otimes \overline{r}) \circ \left(\operatorname{id}_C \otimes T_{C,C}\right) \circ (\Delta_C \otimes \operatorname{id}_C)$$
$$= m_A \circ \left(\operatorname{id}_A \otimes \operatorname{kan}_{1,A}^{-1}\right) \circ (r \otimes \overline{r} \otimes \epsilon_C) \circ (\Delta_C \otimes \operatorname{id}_C) . \qquad (13.19.40)$$

[420]

this becomes

$$m^{(3)} \circ (r \otimes r \otimes \overline{r} \otimes \overline{r}) \circ \left(\operatorname{id}_C \otimes \operatorname{id}_C \otimes T_{C,C}\right) \circ \left(\operatorname{id}_C \otimes T_{C,C} \otimes \operatorname{id}_C\right) \circ (\Delta_C \otimes \Delta_C)$$
$$= m^{(2)} \circ \underbrace{(\operatorname{id}_A \otimes m_A \otimes \operatorname{id}_A) \circ (r \otimes r \otimes \overline{r} \otimes \overline{r})}_{= r \otimes (m_A \circ (r \otimes \overline{r})) \otimes \overline{r}} \circ \underbrace{\left(\operatorname{id}_C \otimes T_{C,C \otimes C}\right) \circ \left(\operatorname{id}_{C \otimes C} \otimes \Delta_C\right)}_{= (\operatorname{id}_C \otimes \Delta_C \otimes \operatorname{id}_C) \circ \left(\operatorname{id}_C \otimes T_{C,C}\right)} \circ (\Delta_C \otimes \operatorname{id}_C)$$
$$= m^{(2)} \circ \underbrace{(r \otimes (m_A \circ (r \otimes \overline{r})) \otimes \overline{r}) \circ (\operatorname{id}_C \otimes \Delta_C \otimes \operatorname{id}_C)}_{= (r \circ \operatorname{id}_C) \otimes ((m_A \circ (r \otimes \overline{r})) \circ \Delta_C) \otimes (\overline{r} \circ \operatorname{id}_C)} \circ \left(\operatorname{id}_C \otimes T_{C,C}\right) \circ (\Delta_C \otimes \operatorname{id}_C)$$
$$= m^{(2)} \circ \left(\underbrace{(r \circ \operatorname{id}_C)}_{= r} \otimes \underbrace{((m_A \circ (r \otimes \overline{r})) \circ \Delta_C)}_{= m_A \circ (r \otimes \overline{r}) \circ \Delta_C} \otimes \underbrace{(\overline{r} \circ \operatorname{id}_C)}_{= \overline{r}} \right) \circ \left(\operatorname{id}_C \otimes T_{C,C}\right) \circ (\Delta_C \otimes \operatorname{id}_C)$$
$$= m^{(2)} \circ (r \otimes (m_A \circ (r \otimes \overline{r}) \circ \Delta_C) \otimes \overline{r}) \circ \left(\operatorname{id}_C \otimes T_{C,C}\right) \circ (\Delta_C \otimes \operatorname{id}_C) .$$

This proves (13.19.37).

[419] *Proof of (13.19.39):* We have $r \star \overline{r} = m_A \circ (r \otimes \overline{r}) \circ \Delta_C$ (according to the definition of convolution). Compared with $r \star \overline{r} = u_A \circ \epsilon_C$, this yields $m_A \circ (r \otimes \overline{r}) \circ \Delta_C = u_A \circ \epsilon_C$. Thus,

$$m^{(2)} \circ \left(r \otimes \underbrace{(m_A \circ (r \otimes \overline{r}) \circ \Delta_C)}_{= u_A \circ \epsilon_C} \otimes \overline{r} \right) \circ \left(\operatorname{id}_C \otimes T_{C,C}\right) \circ (\Delta_C \otimes \operatorname{id}_C)$$
$$= m^{(2)} \circ (r \otimes (u_A \circ \epsilon_C) \otimes \overline{r}) \circ \left(\operatorname{id}_C \otimes T_{C,C}\right) \circ (\Delta_C \otimes \operatorname{id}_C) .$$

This proves (13.19.39).

[420] *Proof of (13.19.40):* We have

$$\underbrace{r}_{= \operatorname{id}_A \circ r} \otimes (u_A \circ \epsilon_C) \otimes \underbrace{\overline{r}}_{= \operatorname{id}_A \circ \overline{r}} = (\operatorname{id}_A \circ r) \otimes (u_A \circ \epsilon_C) \otimes (\operatorname{id}_A \circ \overline{r}) = (\operatorname{id}_A \otimes u_A \otimes \operatorname{id}_A) \circ (r \otimes \epsilon_C \otimes \overline{r}) ,$$

so that

$$\underbrace{m^{(2)}}_{\substack{= m_A \circ (\operatorname{id}_A \otimes m_A) \\ \text{(by (13.19.32))}}} \circ \underbrace{(r \otimes (u_A \circ \epsilon_C) \otimes \overline{r})}_{= (\operatorname{id}_A \otimes u_A \otimes \operatorname{id}_A) \circ (r \otimes \epsilon_C \otimes \overline{r})}$$
$$= m_A \circ (\operatorname{id}_A \otimes m_A) \circ \underbrace{(\operatorname{id}_A \otimes u_A \otimes \operatorname{id}_A)}_{= \operatorname{id}_A \otimes (u_A \otimes \operatorname{id}_A)} \circ (r \otimes \epsilon_C \otimes \overline{r}) = m_A \circ \underbrace{(\operatorname{id}_A \otimes m_A) \circ (\operatorname{id}_A \otimes (u_A \otimes \operatorname{id}_A))}_{= (\operatorname{id}_A \circ \operatorname{id}_A) \otimes (m_A \circ (u_A \otimes \operatorname{id}_A))} \circ (r \otimes \epsilon_C \otimes \overline{r})$$
$$= m_A \circ \left(\underbrace{(\operatorname{id}_A \circ \operatorname{id}_A)}_{= \operatorname{id}_A} \otimes \underbrace{(m_A \circ (u_A \otimes \operatorname{id}_A))}_{= \operatorname{kan}_{2,A}^{-1}} \right) \circ (r \otimes \epsilon_C \otimes \overline{r}) = m_A \circ \left(\operatorname{id}_A \otimes \operatorname{kan}_{2,A}^{-1}\right) \circ (r \otimes \epsilon_C \otimes \overline{r}) .$$

Hence,

$$\underbrace{m^{(2)} \circ (r \otimes (u_A \circ \epsilon_C) \otimes \overline{r})}_{= m_A \circ \left(\operatorname{id}_A \otimes \operatorname{kan}_{2,A}^{-1}\right) \circ (r \otimes \epsilon_C \otimes \overline{r})} \circ \left(\operatorname{id}_C \otimes T_{C,C}\right) \circ (\Delta_C \otimes \operatorname{id}_C)$$
$$= m_A \circ \left(\operatorname{id}_A \otimes \operatorname{kan}_{2,A}^{-1}\right) \circ (r \otimes \epsilon_C \otimes \overline{r}) \circ \left(\operatorname{id}_C \otimes T_{C,C}\right) \circ (\Delta_C \otimes \operatorname{id}_C) .$$

Since

$$\underbrace{(r \otimes \epsilon_C \otimes \overline{r})}_{= r \otimes (\epsilon_C \otimes \overline{r})} \circ \left(\operatorname{id}_C \otimes T_{C,C}\right) = (r \otimes (\epsilon_C \otimes \overline{r})) \circ \left(\operatorname{id}_C \otimes T_{C,C}\right) = \underbrace{(r \circ \operatorname{id}_C)}_{= r = \operatorname{id}_A \circ r} \otimes \underbrace{\left((\epsilon_C \otimes \overline{r}) \circ T_{C,C}\right)}_{\substack{= T_{A,\mathbf{k}} \circ (\overline{r} \otimes \epsilon_C) \\ \text{(by (13.19.4), applied to } U = C,\ V = C,\ U' = A,\ V' = \mathbf{k}, \\ x = \overline{r} \text{ and } y = \epsilon_C)}}$$
$$= (\operatorname{id}_A \circ r) \otimes \left(T_{A,\mathbf{k}} \circ (\overline{r} \otimes \epsilon_C)\right) = \left(\operatorname{id}_A \otimes T_{A,\mathbf{k}}\right) \circ \underbrace{(r \otimes (\overline{r} \otimes \epsilon_C))}_{= r \otimes \overline{r} \otimes \epsilon_C} = \left(\operatorname{id}_A \otimes T_{A,\mathbf{k}}\right) \circ (r \otimes \overline{r} \otimes \epsilon_C) ,$$

- We have
$$\begin{aligned} & m_A \circ \left(\mathrm{id}_A \otimes \mathrm{kan}_{1,A}^{-1}\right) \circ (r \otimes \overline{r} \otimes \epsilon_C) \circ (\Delta_C \otimes \mathrm{id}_C) \\ &= \mathrm{kan}_{1,A}^{-1} \circ ((m_A \circ (r \otimes \overline{r}) \circ \Delta_C) \otimes \epsilon_C) . \end{aligned} \tag{13.19.41}$$
[421]
- We have
$$\mathrm{kan}_{1,A}^{-1} \circ ((m_A \circ (r \otimes \overline{r}) \circ \Delta_C) \otimes \epsilon_C) = \mathrm{kan}_{1,A}^{-1} \circ ((u_A \circ \epsilon_C) \otimes \epsilon_C) . \tag{13.19.42}$$
[422]
- We have
$$\mathrm{kan}_{1,A}^{-1} \circ ((u_A \circ \epsilon_C) \otimes \epsilon_C) = m_A \circ ((u_A \circ \epsilon_C) \otimes (u_A \circ \epsilon_C)) \tag{13.19.43}$$
[423]

---

this becomes
$$\begin{aligned} & m^{(2)} \circ (r \otimes (u_A \circ \epsilon_C) \otimes \overline{r}) \circ \left(\mathrm{id}_C \otimes T_{C,C}\right) \circ (\Delta_C \otimes \mathrm{id}_C) \\ &= m_A \circ \left(\mathrm{id}_A \otimes \mathrm{kan}_{2,A}^{-1}\right) \circ \underbrace{(r \otimes \epsilon_C \otimes \overline{r}) \circ \left(\mathrm{id}_C \otimes T_{C,C}\right)}_{=(\mathrm{id}_A \otimes T_{A,\mathbf{k}}) \circ (r \otimes \overline{r} \otimes \epsilon_C)} \circ (\Delta_C \otimes \mathrm{id}_C) \\ &= m_A \circ \left(\mathrm{id}_A \otimes \mathrm{kan}_{2,A}^{-1}\right) \circ \left(\mathrm{id}_A \otimes T_{A,\mathbf{k}}\right) \circ (r \otimes \overline{r} \otimes \epsilon_C) \circ (\Delta_C \otimes \mathrm{id}_C) . \end{aligned}$$
Since
$$\left(\mathrm{id}_A \otimes \mathrm{kan}_{2,A}^{-1}\right) \circ \left(\mathrm{id}_A \otimes T_{A,\mathbf{k}}\right) = \underbrace{(\mathrm{id}_A \circ \mathrm{id}_A)}_{=\mathrm{id}_A} \otimes \underbrace{\left(\mathrm{kan}_{2,A}^{-1} \circ T_{A,\mathbf{k}}\right)}_{\substack{=\mathrm{kan}_{1,A}^{-1} \\ \text{(by (13.19.5), applied to } U=A)}} = \mathrm{id}_A \otimes \mathrm{kan}_{1,A}^{-1},$$
this becomes
$$\begin{aligned} m^{(2)} \circ (r \otimes (u_A \circ \epsilon_C) \otimes \overline{r}) \circ \left(\mathrm{id}_C \otimes T_{C,C}\right) \circ (\Delta_C \otimes \mathrm{id}_C) &= m_A \circ \underbrace{\left(\mathrm{id}_A \otimes \mathrm{kan}_{2,A}^{-1}\right) \circ \left(\mathrm{id}_A \otimes T_{A,\mathbf{k}}\right)}_{=\mathrm{id}_A \otimes \mathrm{kan}_{1,A}^{-1}} \circ (r \otimes \overline{r} \otimes \epsilon_C) \circ (\Delta_C \otimes \mathrm{id}_C) \\ &= m_A \circ \left(\mathrm{id}_A \otimes \mathrm{kan}_{1,A}^{-1}\right) \circ (r \otimes \overline{r} \otimes \epsilon_C) \circ (\Delta_C \otimes \mathrm{id}_C) . \end{aligned}$$
This proves (13.19.40).

[421] *Proof of (13.19.41):* The equality (13.19.17) (applied to $U = A \otimes A$, $V = A$ and $\alpha = m_A$) yields $\mathrm{kan}_{1,A}^{-1} \circ (m_A \otimes \mathrm{id}_{\mathbf{k}}) = m_A \circ \mathrm{kan}_{1,A\otimes A}^{-1}$. But $\mathrm{id}_A \otimes \mathrm{kan}_{1,A}^{-1} = \mathrm{kan}_{1,A\otimes A}^{-1}$ (according to (13.19.11), applied to $U = A$ and $V = A$), and thus $m_A \circ \underbrace{\left(\mathrm{id}_A \otimes \mathrm{kan}_{1,A}^{-1}\right)}_{=\mathrm{kan}_{1,A\otimes A}^{-1}} = m_A \circ \mathrm{kan}_{1,A\otimes A}^{-1} = \mathrm{kan}_{1,A}^{-1} \circ (m_A \otimes \mathrm{id}_{\mathbf{k}})$. Hence,
$$\begin{aligned} & \underbrace{m_A \circ \left(\mathrm{id}_A \otimes \mathrm{kan}_{1,A}^{-1}\right)}_{=\mathrm{kan}_{1,A}^{-1} \circ (m_A \otimes \mathrm{id}_{\mathbf{k}})} \circ (r \otimes \overline{r} \otimes \epsilon_C) \circ (\Delta_C \otimes \mathrm{id}_C) \\ &= \mathrm{kan}_{1,A}^{-1} \circ (m_A \otimes \mathrm{id}_{\mathbf{k}}) \circ \underbrace{(r \otimes \overline{r} \otimes \epsilon_C)}_{=(r \otimes \overline{r}) \otimes \epsilon_C} \circ (\Delta_C \otimes \mathrm{id}_C) = \mathrm{kan}_{1,A}^{-1} \circ \underbrace{(m_A \otimes \mathrm{id}_{\mathbf{k}}) \circ ((r \otimes \overline{r}) \otimes \epsilon_C) \circ (\Delta_C \otimes \mathrm{id}_C)}_{=(m_A \circ (r \otimes \overline{r}) \circ \Delta_C) \otimes (\mathrm{id}_{\mathbf{k}} \circ \epsilon_C \circ \mathrm{id}_C)} \\ &= \mathrm{kan}_{1,A}^{-1} \circ \left((m_A \circ (r \otimes \overline{r}) \circ \Delta_C) \otimes \underbrace{(\mathrm{id}_{\mathbf{k}} \circ \epsilon_C \circ \mathrm{id}_C)}_{=\epsilon_C} \right) = \mathrm{kan}_{1,A}^{-1} \circ ((m_A \circ (r \otimes \overline{r}) \circ \Delta_C) \otimes \epsilon_C) . \end{aligned}$$
This proves (13.19.41).

[422] *Proof of (13.19.42):* The definition of convolution yields $r \star \overline{r} = m_A \circ (r \otimes \overline{r}) \circ \Delta_C$. Compared with $r \star \overline{r} = u_A \circ \epsilon_C$, this yields $m_A \circ (r \otimes \overline{r}) \circ \Delta_C = u_A \circ \epsilon_C$. Hence, $\mathrm{kan}_{1,A}^{-1} \circ \left(\underbrace{(m_A \circ (r \otimes \overline{r}) \circ \Delta_C)}_{=u_A \circ \epsilon_C} \otimes \epsilon_C \right) = \mathrm{kan}_{1,A}^{-1} \circ ((u_A \circ \epsilon_C) \otimes \epsilon_C)$. This proves (13.19.42).

[423] *Proof of (13.19.43):* We have
$$(u_A \circ \epsilon_C) \otimes \underbrace{\epsilon_C}_{=\mathrm{id}_{\mathbf{k}} \circ \epsilon_C} = (u_A \circ \epsilon_C) \otimes (\mathrm{id}_{\mathbf{k}} \circ \epsilon_C) = (u_A \otimes \mathrm{id}_{\mathbf{k}}) \circ (\epsilon_C \otimes \epsilon_C) ,$$

- We have

$$m_A \circ ((u_A \circ \epsilon_C) \otimes (u_A \circ \epsilon_C)) = u_A \circ \epsilon_C \circ m_C. \tag{13.19.44}$$

[424]

Now,

$$\begin{aligned}
f \circledast g &= m_A \circ (f \otimes g) \circ (\mathrm{id}_C \otimes T_{C,C} \otimes \mathrm{id}_C) \circ (\Delta_C \otimes \Delta_C) \qquad (\text{by (13.19.35)}) \\
&= m^{(3)} \circ (r \otimes r \otimes \overline{r} \otimes \overline{r}) \circ (\mathrm{id}_C \otimes \mathrm{id}_C \otimes T_{C,C}) \circ (\mathrm{id}_C \otimes T_{C,C} \otimes \mathrm{id}_C) \circ (\Delta_C \otimes \Delta_C) \qquad (\text{by (13.19.36)}) \\
&= m^{(2)} \circ (r \otimes (m_A \circ (r \otimes \overline{r}) \circ \Delta_C) \otimes \overline{r}) \circ (\mathrm{id}_C \otimes T_{C,C}) \circ (\Delta_C \otimes \mathrm{id}_C) \qquad (\text{by (13.19.37)}) \\
&= m^{(2)} \circ (r \otimes (u_A \circ \epsilon_C) \otimes \overline{r}) \circ (\mathrm{id}_C \otimes T_{C,C}) \circ (\Delta_C \otimes \mathrm{id}_C) \qquad (\text{by (13.19.39)}) \\
&= m_A \circ \left(\mathrm{id}_A \otimes \mathrm{kan}_{1,A}^{-1}\right) \circ (r \otimes \overline{r} \otimes \epsilon_C) \circ (\Delta_C \otimes \mathrm{id}_C) \qquad (\text{by (13.19.40)}) \\
&= \mathrm{kan}_{1,A}^{-1} \circ ((m_A \circ (r \otimes \overline{r}) \circ \Delta_C) \otimes \epsilon_C) \qquad (\text{by (13.19.41)}) \\
&= \mathrm{kan}_{1,A}^{-1} \circ ((u_A \circ \epsilon_C) \otimes \epsilon_C) \qquad (\text{by (13.19.42)}) \\
&= m_A \circ ((u_A \circ \epsilon_C) \otimes (u_A \circ \epsilon_C)) \qquad (\text{by (13.19.43)}) \\
&= u_A \circ \epsilon_C \circ m_C \qquad (\text{by (13.19.44)}) \\
&= u_A \circ \epsilon_{C \otimes C} \qquad (\text{by (13.19.21)}) .
\end{aligned}$$

The second equality in (13.19.19) is thus proven.

Thus, both equalities in (13.19.19) are proven. Hence, (13.19.19) is proven. As we have already seen, this yields (13.19.20). In other words, $h = g$. But $h = \overline{r} \circ m_C$, so that $\overline{r} \circ m_C = h = g = m_A \circ T_{A,A} \circ (\overline{r} \otimes \overline{r})$.

Finally, (13.19.4) (applied to $U = C$, $V = C$, $U' = A$, $V' = A$, $x = \overline{r}$ and $y = \overline{r}$) yields $(\overline{r} \otimes \overline{r}) \circ T_{C,C} = T_{A,A} \circ (\overline{r} \otimes \overline{r})$. Thus, $T_{A,A} \circ (\overline{r} \otimes \overline{r}) = (\overline{r} \otimes \overline{r}) \circ T_{C,C}$, so that $\overline{r} \circ m_C = m_A \circ \underbrace{T_{A,A} \circ (\overline{r} \otimes \overline{r})}_{= (\overline{r} \otimes \overline{r}) \circ T_{C,C}} = m_A \circ (\overline{r} \otimes \overline{r}) \circ T_{C,C}$. As we know, this completes the solution of Exercise 1.4.29(a).

[*Remark:* The second solution of Exercise 1.4.29(a) has been obtained more or less straightforwardly from the first solution by rewriting it in an element-free fashion. First of all, the maps f, g and h introduced in the second solution are precisely the maps f, g and h introduced in the first solution, just rewritten in an element-free way. Also, for example, the equalities (13.19.22), (13.19.23), (13.19.24), (13.19.25), (13.19.26) and (13.19.27) have been found by rewriting the six equality signs in the computation (13.19.2) in an element-free way: For instance, the second equality sign in (13.19.2) stands for the equality

$$\sum_{(a),(b)} h(a_1 \otimes b_1) f(a_2 \otimes b_2) = \sum_{(a),(b)} \overline{r}(a_1 b_1) r(a_2) r(b_2) \qquad \text{for all } a \in C \text{ and } b \in C,$$

so that

$$\mathrm{kan}_{1,A}^{-1} \circ \underbrace{((u_A \circ \epsilon_C) \otimes \epsilon_C)}_{=(u_A \otimes \mathrm{id}_{\mathbf{k}}) \circ (\epsilon_C \otimes \epsilon_C)} = \mathrm{kan}_{1,A}^{-1} \circ (u_A \otimes \mathrm{id}_{\mathbf{k}}) \circ (\epsilon_C \otimes \epsilon_C) .$$

Compared with

$$\begin{aligned}
m_A \circ \left(\left(\underbrace{u_A}_{= \mathrm{id}_A \circ u_A} \circ \epsilon_C \right) \otimes \left(\underbrace{u_A}_{= u_A \circ \mathrm{id}_{\mathbf{k}}} \circ \epsilon_C \right) \right) &= m_A \circ \underbrace{((\mathrm{id}_A \circ u_A \circ \epsilon_C) \otimes (u_A \circ \mathrm{id}_{\mathbf{k}} \circ \epsilon_C))}_{=(\mathrm{id}_A \otimes u_A) \circ (u_A \otimes \mathrm{id}_{\mathbf{k}}) \circ (\epsilon_C \otimes \epsilon_C)} \\
&= \underbrace{m_A \circ (\mathrm{id}_A \otimes u_A)}_{= \mathrm{kan}_{1,A}^{-1}} \circ (u_A \otimes \mathrm{id}_{\mathbf{k}}) \circ (\epsilon_C \otimes \epsilon_C) = \mathrm{kan}_{1,A}^{-1} \circ (u_A \otimes \mathrm{id}_{\mathbf{k}}) \circ (\epsilon_C \otimes \epsilon_C) ,
\end{aligned}$$

this yields $\mathrm{kan}_{1,A}^{-1} \circ ((u_A \circ \epsilon_C) \otimes \epsilon_C) = m_A \circ ((u_A \circ \epsilon_C) \otimes (u_A \circ \epsilon_C))$. Thus, (13.19.43) is solved.

[424] *Proof of (13.19.44):* We know that u_A is a $\mathbf{k}$-algebra homomorphism (indeed, this is an easy fact that holds whenever A is a $\mathbf{k}$-algebra). Since the two maps u_A and ϵ_C are $\mathbf{k}$-algebra homomorphisms, their composition $u_A \circ \epsilon_C$ is a $\mathbf{k}$-algebra homomorphism. Consequently, $(u_A \circ \epsilon_C) \circ m_C = m_A \circ ((u_A \circ \epsilon_C) \otimes (u_A \circ \epsilon_C))$ (in fact, this is one of the axioms a $\mathbf{k}$-algebra homomorphism has to satisfy), so that $m_A \circ ((u_A \circ \epsilon_C) \otimes (u_A \circ \epsilon_C)) = (u_A \circ \epsilon_C) \circ m_C = u_A \circ \epsilon_C \circ m_C$. This proves (13.19.44).

i.e., for the following equality of maps:

$$\left(\text{the } \mathbf{k}\text{-linear map } C \otimes C \to A \text{ sending every } a \otimes b \text{ to } \sum_{(a),(b)} h(a_1 \otimes b_1) f(a_2 \otimes b_2)\right)$$
$$= \left(\text{the } \mathbf{k}\text{-linear map } C \otimes C \to A \text{ sending every } a \otimes b \text{ to } \sum_{(a),(b)} \overline{r}(a_1 b_1) r(a_2) r(b_2)\right).$$

But upon rewriting these two maps without referring to elements, this takes the form

$$m_A \circ (h \otimes f) \circ (\operatorname{id}_C \otimes T_{C,C} \otimes \operatorname{id}_C) \circ (\Delta_C \otimes \Delta_C)$$
$$= m_A \circ ((\overline{r} \circ m_C) \otimes (m_A \circ (r \otimes r))) \circ (\operatorname{id}_C \otimes T_{C,C} \otimes \operatorname{id}_C) \circ (\Delta_C \otimes \Delta_C)$$

(because the $\mathbf{k}$-linear map $C \otimes C \to A$ sending every $a \otimes b$ to $\sum_{(a),(b)} h(a_1 \otimes b_1) f(a_2 \otimes b_2)$ is $m_A \circ (h \otimes f) \circ (\operatorname{id}_C \otimes T_{C,C} \otimes \operatorname{id}_C) \circ (\Delta_C \otimes \Delta_C)$, and the $\mathbf{k}$-linear map $C \otimes C \to A$ sending every $a \otimes b$ to $\sum_{(a),(b)} \overline{r}(a_1 b_1) r(a_2) r(b_2)$ is $m_A \circ ((\overline{r} \circ m_C) \otimes (m_A \circ (r \otimes r))) \circ (\operatorname{id}_C \otimes T_{C,C} \otimes \operatorname{id}_C) \circ (\Delta_C \otimes \Delta_C)$); and this is precisely the equality (13.19.23). The process of rewriting a map in an element-free way is not completely deterministic (e.g., we could also have rewritten the $\mathbf{k}$-linear map $C \otimes C \to A$ sending every $a \otimes b$ to $\sum_{(a),(b)} \overline{r}(a_1 b_1) r(a_2) r(b_2)$ as $m^{(3)} \circ ((\overline{r} \circ m_C) \otimes r \otimes r) \circ (\operatorname{id}_C \otimes T_{C,C} \otimes \operatorname{id}_C) \circ (\Delta_C \otimes \Delta_C)$, since $m^{(3)}$ is the $\mathbf{k}$-linear map $A \otimes A \otimes A \to A$ sending every $a \otimes b \otimes c$ to abc), but all possible results (for a given map) can be reduced to each other using purely linear-algebraic formulae[425] and various forms of the general associativity law (such as (13.19.33) and (13.19.34))[426]. The equalities (13.19.22), (13.19.23), (13.19.24), (13.19.25), (13.19.26) and (13.19.27) are somewhat more complicated to check than the corresponding equality signs in (13.19.2), because of the fact that one and the same map can be written in different forms whose equivalence needs to be proven. However, this additional complexity is straightforward to surmount: each equality, when rewritten in element-free terms, can be proven by the same arguments as the corresponding equality that uses elements, combined with purely linear-algebraic formulae like (13.19.4) and (13.19.5) and various forms of the general associativity law (such as (13.19.33) and (13.19.34)). (There is also a way how to do element-free computations without such added difficulty, using *string diagrams*[427].)

See the First solution of Exercise 1.2.3 as well as the solution of Exercise 1.5.6 for other examples of how a proof that uses elements can be rewritten in an element-free fashion.]

(b) We have solved Exercise 1.4.29(a) in an element-free fashion (in the Second solution to Exercise 1.4.29(a) given above)[428]. Thus, by "reversing all arrows" in this solution of Exercise 1.4.29(a), we can obtain a solution to the dual of Exercise 1.4.29(a). Consequently, the dual of Exercise 1.4.29(a) holds.

But it is easy to see that the notion of a $\mathbf{k}$-coalgebra anti-homomorphism is dual to the notion of a $\mathbf{k}$-algebra anti-homomorphism (i.e., is obtained from the latter notion by "reversing all arrows"), and the notion of convolution is dual to itself. Hence, the dual of Exercise 1.4.29(a) is the following exercise:

> *Exercise A:* If C is a $\mathbf{k}$-bialgebra, if A is a $\mathbf{k}$-coalgebra, and if $r : A \to C$ is a $\star$-invertible $\mathbf{k}$-coalgebra homomorphism, then prove that the $\star$-inverse $r^{\star(-1)}$ of r is a $\mathbf{k}$-coalgebra anti-homomorphism.

[425] By "purely linear-algebraic formulae", I mean formulae such as (13.19.4) and (13.19.5) and the identity $(\beta \circ \alpha) \otimes (\beta' \circ \alpha') = (\beta \otimes \beta') \circ (\alpha \otimes \alpha')$ for tensor products of compositions of maps. This kind of formulae hold in any tensor category (at least if we follow the abuse of notation that allows us to treat the tensor product as associative).

[426] The reason why we need to use the general associativity law is that, in the first solution, we used unparenthesized products of more than one element of A (for example, the expression $\sum_{(a),(b)} \overline{r}(a_1 b_1) r(a_2) r(b_2)$ contains the unparenthesized product $\overline{r}(a_1 b_1) r(a_2) r(b_2)$ of three factors). If we would parenthesize all such products in such a way that no more than two factors are ever multiplied at the same time (e.g., we could replace $\overline{r}(a_1 b_1) r(a_2) r(b_2)$ by $\overline{r}(a_1 b_1)(r(a_2) r(b_2))$), and if we would explicitly use the (non-general) associativity law $(xy) z = x (yz)$ to switch between these parenthesizations, then we would not have to use general associativity any more when rewriting the proof in an element-free fashion (but, of course, the proof would be longer).

[427] See http://ncatlab.org/nlab/show/string+diagram and the references therein.

[428] To be fully honest, this second solution was not entirely element-free, since we proved the auxiliary equality (13.19.38) using elements. However, for the purposes of what we are going to do (reversing arrows), this is not problematic, since this auxiliary equality (13.19.38) can be easily shown to hold with all arrows reversed (the proof is more or less the same as for the original (13.19.38)).

Exercise 1.4.29(b) immediately follows from this Exercise A (applied to A and C instead of C and A). Exercise 1.4.29(b) is thus solved.

(c) *Alternative proof of Proposition 1.4.10 using Exercise 1.4.29(a).* Let A be a Hopf algebra. The antipode S of this Hopf algebra A is defined as the $\star$-inverse of the identity map id_A; thus, $S = \operatorname{id}_A^{\star(-1)}$. Hence, the $\mathbf{k}$-algebra homomorphism $\operatorname{id}_A : A \to A$ is $\star$-invertible. Therefore, Exercise 1.4.29(a) (applied to $C = A$ and $r = \operatorname{id}_A$) yields that the $\star$-inverse $\operatorname{id}_A^{\star(-1)}$ of id_A is a $\mathbf{k}$-algebra anti-homomorphism. In other words, S is a $\mathbf{k}$-algebra anti-homomorphism (since S is the $\star$-inverse $\operatorname{id}_A^{\star(-1)}$ of id_A). In other words, S is a $\mathbf{k}$-algebra anti-endomorphism of A. This proves Proposition 1.4.10.

Alternative solution of Exercise 1.4.28 using Exercise 1.4.29(b). This is analogous to the proof of Proposition 1.4.10 using Exercise 1.4.29(a) just shown, but instead of Exercise 1.4.29(a) we now need to use Exercise 1.4.29(b).

(d) *Alternative proof of Corollary 1.4.12 using Proposition 1.4.26.* Let A be a commutative Hopf algebra. Then, a $\mathbf{k}$-algebra anti-homomorphism $A \to A$ is the same thing as a $\mathbf{k}$-algebra homomorphism $A \to A$ (since A is commutative). (This follows from Exercise 1.5.8(a), but it is also nearly trivial to check by hand.)

The $\mathbf{k}$-linear map $\operatorname{id}_A : A \to A$ is $\star$-invertible (since A is a Hopf algebra), and its $\star$-inverse $\operatorname{id}_A^{\star(-1)}$ is the antipode S of A. That is, $\operatorname{id}_A^{\star(-1)} = S$. Applying Exercise 1.4.29(a) to $C = A$ and $r = \operatorname{id}_A$, we now conclude that $\operatorname{id}_A^{\star(-1)}$ is a $\mathbf{k}$-algebra anti-homomorphism $A \to A$ (since id_A is a $\mathbf{k}$-algebra homomorphism $A \to A$). In other words, $\operatorname{id}_A^{\star(-1)}$ is a $\mathbf{k}$-algebra homomorphism $A \to A$ (since a $\mathbf{k}$-algebra anti-homomorphism $A \to A$ is the same thing as a $\mathbf{k}$-algebra homomorphism $A \to A$). In other words, S is a $\mathbf{k}$-algebra homomorphism $A \to A$ (since $\operatorname{id}_A^{\star(-1)} = S$). Now, Proposition 1.4.26(a) (applied to $H = A$ and $\alpha = S$) yields $S \circ S = S^{\star(-1)}$. But $S^{\star(-1)} = \operatorname{id}_A$ (since $S = \operatorname{id}_A^{\star(-1)}$). Hence, $S^2 = S \circ S = S^{\star(-1)} = \operatorname{id}_A$. This proves Corollary 1.4.12. Exercise 1.4.29(d) is solved.

(e) Let us first prepare some general properties of maps between graded $\mathbf{k}$-modules. We begin with two definitions:

Definition 13.19.1. Let $V = \bigoplus_{n\in\mathbb{N}} V_n$ be a graded $\mathbf{k}$-module (where the V_n are the homogeneous components of V). Let $n \in \mathbb{N}$. Then:

(a) We shall let $\pi_{n,V} : V \to V_n$ denote the canonical projection from V to its n-th graded component V_n. (This has already been defined in Definition 13.5.1; we are just repeating it here for convenience.)

(b) We shall let $\iota_{n,V} : V_n \to V$ denote the inclusion map from V_n to V.

Definition 13.19.2. Let $V = \bigoplus_{n\in\mathbb{N}} V_n$ and $W = \bigoplus_{n\in\mathbb{N}} W_n$ be two graded $\mathbf{k}$-modules (where the V_n and the W_n are the homogeneous components of V and W, respectively). Let $f : V \to W$ be a $\mathbf{k}$-linear map. For each $n \in \mathbb{N}$, let $\overline{f}_n : V_n \to W_n$ be the $\mathbf{k}$-linear map $\pi_{n,W} \circ f \circ \iota_{n,V} : V_n \to W_n$. We let $\overline{f}$ be the direct sum $\bigoplus_{n\in\mathbb{N}} \overline{f}_n : \bigoplus_{n\in\mathbb{N}} V_n \to \bigoplus_{n\in\mathbb{N}} W_n$ of these $\mathbf{k}$-linear maps $\overline{f}_n : V_n \to W_n$ over all $n \in \mathbb{N}$; thus, $\overline{f}$ is a $\mathbf{k}$-linear map from V to W (since $V = \bigoplus_{n\in\mathbb{N}} V_n$ and $W = \bigoplus_{n\in\mathbb{N}} W_n$).

Let us now state a litany of basic properties of this map $\overline{f}$:

Proposition 13.19.3. *Let V and W be two graded $\mathbf{k}$-modules. Write V as $V = \bigoplus_{n\in\mathbb{N}} V_n$ (where the V_n are the homogeneous components of V). Let $f : V \to W$ be a $\mathbf{k}$-linear map. Then,*

$$\overline{f}(v) = \pi_{n,W}(f(v)) \qquad \text{for each } n \in \mathbb{N} \text{ and each } v \in V_n.$$

Proof of Proposition 13.19.3. Write W in the form $W = \bigoplus_{n\in\mathbb{N}} W_n$ (where the W_n are the homogeneous components of W). Consider the $\mathbf{k}$-linear maps $\overline{f}_n$ defined in Definition 13.19.2; thus, $\overline{f}_n = \pi_{n,W} \circ f \circ \iota_{n,V}$ for each $n \in \mathbb{N}$. The map $\overline{f}$ was defined to be the direct sum $\bigoplus_{n\in\mathbb{N}} \overline{f}_n$. Thus, we have

$$\overline{f}(v) = \overline{f}_n(v) \qquad \text{for each } n \in \mathbb{N} \text{ and each } v \in V_n. \tag{13.19.45}$$

Now, let $n \in \mathbb{N}$ and $v \in V_n$. Then, $\iota_{n,V}(v) = v$ (since $\iota_{n,V}$ is the inclusion map from V_n to V). Now, (13.19.45) yields

$$\overline{f}(v) = \underbrace{\overline{f}_n}_{\substack{=\pi_{n,W}\circ f\circ\iota_{n,V} \\ \text{(by the definition of } \overline{f}_n\text{)}}}(v) = (\pi_{n,W} \circ f \circ \iota_{n,V})(v) = \pi_{n,W}\left(f\left(\underbrace{\iota_{n,V}(v)}_{=v}\right)\right) = \pi_{n,W}(f(v)).$$

This proves Proposition 13.19.3. □

Proposition 13.19.4. *Let V and W be two graded* **k***-modules. Let $f : V \to W$ be a graded* **k***-linear map. Then, $\overline{f} = f$.*

Proof of Proposition 13.19.4. Write V as $V = \bigoplus_{n \in \mathbb{N}} V_n$ (where the V_n are the homogeneous components of V). Write W as $W = \bigoplus_{n \in \mathbb{N}} W_n$ (where the W_n are the homogeneous components of W).

Let v be a homogeneous element of V. Thus, there exists some $n \in \mathbb{N}$ such that $v \in V_n$. Consider this n. We have $f(V_n) \subset W_n$ (since the map f is graded). Hence, $f\left(\underbrace{v}_{\in V_n}\right) \in f(V_n) \subset W_n$.

But $\pi_{n,W}(w) = w$ for each $w \in W_n$ (since $\pi_{n,W}$ is the canonical projection from W to its n-th graded component W_n). Applying this to $w = f(v)$, we obtain $\pi_{n,W}(f(v)) = f(v)$ (since $f(v) \in W_n$). Now, recall that $v \in V_n$; hence, Proposition 13.19.3 yields

$$\overline{f}(v) = \pi_{n,W}(f(v)) = f(v).$$

Forget that we fixed v. We thus have shown that $\overline{f}(v) = f(v)$ for each homogeneous element v of V. In other words, the two **k**-linear maps $\overline{f}$ and f (from V to W) agree on each homogeneous element of V.

But V is a graded **k**-module. Hence, each element of V is a **k**-linear combination of homogeneous elements of V. Therefore, if two **k**-linear maps from V agree on each homogeneous element of V, then these two maps must be equal. Hence, the two **k**-linear maps $\overline{f}$ and f (from V to W) must be equal (because we have shown that they agree on each homogeneous element of V). In other words, $\overline{f} = f$. This proves Proposition 13.19.4. □

Proposition 13.19.5. *Let V and W be two graded* **k***-modules. Let $f : V \to W$ be a* **k***-linear map. Then, the* **k***-linear map $\overline{f} : V \to W$ is graded.*

Proof of Proposition 13.19.5. Write V and W in the form $V = \bigoplus_{n \in \mathbb{N}} V_n$ and $W = \bigoplus_{n \in \mathbb{N}} W_n$ (where the V_n and the W_n are the homogeneous components of V and W, respectively). Consider the **k**-linear maps $\overline{f}_n$ defined in Definition 13.19.2; thus, $\overline{f}_n = \pi_{n,W} \circ f \circ \iota_{n,V}$ for each $n \in \mathbb{N}$. The map $\overline{f}$ was defined to be the direct sum $\bigoplus_{n \in \mathbb{N}} \overline{f}_n$. Thus, we have

$$\overline{f}(v) = \overline{f}_n(v) \qquad \text{for each } n \in \mathbb{N} \text{ and each } v \in V_n. \tag{13.19.46}$$

Let $n \in \mathbb{N}$. We have $\pi_{n,W}(W) \subset W_n$ (since $\pi_{n,W}$ is the canonical projection from W to its n-th graded component W_n). Now,

$$\overline{f}(V_n) = \left\{ \underbrace{\overline{f}(v)}_{\substack{=\overline{f}_n(v) \\ \text{(by (13.19.46))}}} \mid v \in V_n \right\} = \left\{ \overline{f}_n(v) \mid v \in V_n \right\} = \underbrace{\overline{f}_n}_{\substack{=\pi_{n,W} \circ f \circ \iota_{n,V} \\ \text{(by the definition of } \overline{f}_n\text{)}}} (V_n)$$

$$= (\pi_{n,W} \circ f \circ \iota_{n,V})(V_n) = \pi_{n,W}\left(\underbrace{f(\iota_{n,V}(V_n))}_{\subset W}\right) \subset \pi_{n,W}(W) \subset W_n.$$

Now, forget that we fixed n. We thus have shown that $\overline{f}(V_n) \subset W_n$ for every $n \in \mathbb{N}$. In other words, the **k**-linear map $\overline{f} : V \to W$ is graded. This proves Proposition 13.19.5. □

Proposition 13.19.6. *Let U, V and W be three graded* **k***-modules. Let $f : V \to W$ and $g : U \to V$ be two* **k***-linear maps such that f is graded. Then, $\overline{f \circ g} = f \circ \overline{g}$.*

Proof of Proposition 13.19.6. Write U in the form $U = \bigoplus_{n \in \mathbb{N}} U_n$ (where the U_n are the homogeneous components of U).

Every element of U is a **k**-linear combination of homogeneous elements of U (since U is graded). Thus, if two **k**-linear maps from U agree on each homogeneous element of U, then these two maps must be equal.

Let u be a homogeneous element of U. Thus, there exists some $n \in \mathbb{N}$ such that $u \in U_n$. Consider this n. Then, Proposition 13.19.3 (applied to U, V, U_m, g and u instead of V, W, V_m, f and v) yields

$$\overline{g}(u) = \pi_{n,V}(g(u)). \tag{13.19.47}$$

Also, $f \circ g : U \to W$ is a $\mathbf{k}$-linear map; thus, Proposition 13.19.3 (applied to U, U_m, $f \circ g$ and u instead of V, V_m, f and v) yields $\overline{f \circ g}(u) = \pi_{n,W}((f \circ g)(u))$. Hence,

$$\begin{aligned}\overline{f \circ g}(u) = \pi_{n,W}((f \circ g)(u)) = \pi_{n,W}(f(g(u))) &= f\left(\underbrace{\pi_{n,V}(g(u))}_{\substack{=\overline{g}(u) \\ (\text{by } (13.19.47))}}\right) \\ &\qquad (\text{by Proposition 13.5.3, applied to } v = g(u) \text{ and } m = n) \\ &= f(\overline{g}(u)) = (f \circ \overline{g})(u).\end{aligned}$$

Now, forget that we fixed u. We thus have proved that $\overline{f \circ g}(u) = (f \circ \overline{g})(u)$ for each homogeneous element u of U. In other words, the two maps $\overline{f \circ g}$ and $f \circ \overline{g}$ (from U to W) agree on each homogeneous element of U. Since these two maps are $\mathbf{k}$-linear, we can thus conclude that they must be equal (because if two $\mathbf{k}$-linear maps from U agree on each homogeneous element of U, then these two maps must be equal). In other words, we have $\overline{f \circ g} = f \circ \overline{g}$. This proves Proposition 13.19.6. □

Proposition 13.19.7. *Let U, V and W be three graded $\mathbf{k}$-modules. Let $f : V \to W$ and $g : U \to V$ be two $\mathbf{k}$-linear maps such that g is graded. Then, $\overline{f \circ g} = \overline{f} \circ g$.*

Proof of Proposition 13.19.7. Write U and V in the forms $U = \bigoplus_{n \in \mathbb{N}} U_n$ and $V = \bigoplus_{n \in \mathbb{N}} V_n$ (where the U_n and V_n are the homogeneous components of U and V, respectively).

Every element of U is a $\mathbf{k}$-linear combination of homogeneous elements of U (since U is graded). Thus, if two $\mathbf{k}$-linear maps from U agree on each homogeneous element of U, then these two maps must be equal.

Let u be a homogeneous element of U. Thus, there exists some $n \in \mathbb{N}$ such that $u \in U_n$. Consider this n. From $u \in U_n$, we obtain $g(u) \in g(U_n) \subset V_n$ (since g is a graded map). Hence, Proposition 13.19.3 (applied to $v = g(u)$) yields

$$\overline{f}(g(u)) = \pi_{n,W}(f(g(u))). \tag{13.19.48}$$

Also, $f \circ g : U \to W$ is a $\mathbf{k}$-linear map; thus, Proposition 13.19.3 (applied to U, U_m, $f \circ g$ and u instead of V, V_m, f and v) yields $\overline{f \circ g}(u) = \pi_{n,W}((f \circ g)(u)) = \pi_{n,W}(f(g(u)))$. Comparing this with (13.19.48), we find $\overline{f \circ g}(u) = \overline{f}(g(u)) = (\overline{f} \circ g)(u)$.

Now, forget that we fixed u. We thus have proved that $\overline{f \circ g}(u) = (\overline{f} \circ g)(u)$ for each homogeneous element u of U. In other words, the two maps $\overline{f \circ g}$ and $\overline{f} \circ g$ (from U to W) agree on each homogeneous element of U. Since these two maps are $\mathbf{k}$-linear, we can thus conclude that they must be equal (because if two $\mathbf{k}$-linear maps from U agree on each homogeneous element of U, then these two maps must be equal). In other words, we have $\overline{f \circ g} = \overline{f} \circ g$. This proves Proposition 13.19.7. □

Remark 13.19.8. Proposition 13.19.6 and Proposition 13.19.7 cannot be generalized to "$\overline{f \circ g} = \overline{f} \circ \overline{g}$ for any two (not necessarily graded) $\mathbf{k}$-linear maps $f : V \to W$ and $g : U \to V$". Counterexamples are easy to find (e.g., take $U = \mathbf{k}[x]$, $V = \mathbf{k}[x]$, $W = \mathbf{k}[x]$, $f(x^i) = x^{i+1}$, $g(x^j) = \begin{cases} x^{j-1}, & \text{if } j > 0; \\ 0, & \text{if } j = 0 \end{cases}$).

Proposition 13.19.9. *Let U, V, X and Y be four graded $\mathbf{k}$-modules. Let $f : U \to V$ and $g : X \to Y$ be two $\mathbf{k}$-linear maps such that f is graded. Then, $\overline{g \otimes f} = \overline{g} \otimes f$.*

Proof of Proposition 13.19.9. Write U, V, X and Y in the forms $U = \bigoplus_{n \in \mathbb{N}} U_n$, $V = \bigoplus_{n \in \mathbb{N}} V_n$, $X = \bigoplus_{n \in \mathbb{N}} X_n$ and $Y = \bigoplus_{n \in \mathbb{N}} Y_n$ (where the U_n, V_n, X_n and Y_n are the homogeneous components of U, V, X and Y, respectively). Write $X \otimes U$ and $Y \otimes V$ in the forms $X \otimes U = \bigoplus_{n \in \mathbb{N}} (X \otimes U)_n$ and $Y \otimes V = \bigoplus_{n \in \mathbb{N}} (Y \otimes V)_n$ (where the $(X \otimes U)_n$ and $(Y \otimes V)_n$ are the homogeneous components of $X \otimes U$ and $Y \otimes V$, respectively).

Recall that the grading on $X \otimes U$ is defined in such a way that

$$(X \otimes U)_n = \bigoplus_{\substack{(i,j)\in\mathbb{N}^2;\\ i+j=n}} X_i \otimes U_j \tag{13.19.49}$$

for each $n \in \mathbb{N}$. Likewise, the grading on $Y \otimes V$ is defined in such a way that

$$(Y \otimes V)_n = \bigoplus_{\substack{(i,j)\in\mathbb{N}^2;\\ i+j=n}} Y_i \otimes V_j \tag{13.19.50}$$

for each $n \in \mathbb{N}$.

If $n \in \mathbb{N}$, then

$$\pi_{n,X}(X_m) = 0 \qquad \text{for every } m \in \mathbb{N} \text{ satisfying } m \neq n \tag{13.19.51}$$

(since the map $\pi_{n,X}$ is the canonical projection from X to its n-th graded component X_n, and thus annihilates all graded components of X other than X_n). The same reasoning (applied to $X \otimes U$ and $(X \otimes U)_m$ instead of X and X_m) shows that if $n \in \mathbb{N}$, then

$$\pi_{n,X\otimes U}((X \otimes U)_m) = 0 \qquad \text{for every } m \in \mathbb{N} \text{ satisfying } m \neq n. \tag{13.19.52}$$

If $n \in \mathbb{N}$, then

$$\pi_{n,Y}(Y_m) = 0 \qquad \text{for every } m \in \mathbb{N} \text{ satisfying } m \neq n \tag{13.19.53}$$

(since the map $\pi_{n,Y}$ is the canonical projection from Y to its n-th graded component Y_n, and thus annihilates all graded components of Y other than Y_n). The same reasoning (applied to $Y \otimes V$ and $(Y \otimes V)_m$ instead of Y and Y_m) shows that if $n \in \mathbb{N}$, then

$$\pi_{n,Y\otimes V}((Y \otimes V)_m) = 0 \qquad \text{for every } m \in \mathbb{N} \text{ satisfying } m \neq n. \tag{13.19.54}$$

We shall now prove the following:

Claim 1: Let $q \in \mathbb{N}$. Let $v \in V_q$. Let $y \in Y$. Let $p \in \mathbb{N}$. Then,

$$\pi_{p+q,Y\otimes V}(y \otimes v) = \pi_{p,Y}(y) \otimes v.$$

[*Proof of Claim 1:* We have $y \in Y = \bigoplus_{n\in\mathbb{N}} Y_n$. Hence, we can write y in the form $y = \sum_{n\in\mathbb{N}} y_n$ for some family $(y_n)_{n\in\mathbb{N}} \in \prod_{n\in\mathbb{N}} Y_n$ with the property that all but finitely many $n \in \mathbb{N}$ satisfy $y_n = 0$. Consider this family $(y_n)_{n\in\mathbb{N}}$. Note that

$$y_n \in Y_n \qquad \text{for each } n \in \mathbb{N} \tag{13.19.55}$$

(since $(y_n)_{n\in\mathbb{N}} \in \prod_{n\in\mathbb{N}} Y_n$).

It is easy to see that

$$y_n \otimes v \in (Y \otimes V)_{n+q} \qquad \text{for each } n \in \mathbb{N} \tag{13.19.56}$$

[429].

It is easy to see that

$$\pi_{p,Y}(y) = y_p \tag{13.19.58}$$

[429] *Proof of* (13.19.56)*:* Let $n \in \mathbb{N}$. Applying (13.19.50) to $n+q$ instead of n, we find

$$(Y \otimes V)_{n+q} = \bigoplus_{\substack{(i,j)\in\mathbb{N}^2;\\ i+j=n+q}} Y_i \otimes V_j. \tag{13.19.57}$$

But $Y_n \otimes V_q$ is an addend of the direct sum $\bigoplus_{\substack{(i,j)\in\mathbb{N}^2;\\ i+j=n+q}} Y_i \otimes V_j$ (namely, the addend for $(i,j) = (n,q)$), since $(n,q) \in \mathbb{N}^2$ and $n+q = n+q$. Hence, $Y_n \otimes V_q \subset \bigoplus_{\substack{(i,j)\in\mathbb{N}^2;\\ i+j=n+q}} Y_i \otimes V_j$. In view of (13.19.57), this rewrites as $Y_n \otimes V_q \subset (Y \otimes V)_{n+q}$. But (13.19.55) yields $y_n \in Y_n$. Combining this with $v \in V_q$, we find $y_n \otimes v \in Y_n \otimes V_q \subset (Y \otimes V)_{n+q}$. This proves (13.19.56).

[430]. We can also see that

$$\pi_{p+q,Y\otimes V}\left(y\otimes v\right)=y_p\otimes v$$

[431]. Hence,

$$\pi_{p+q,Y\otimes V}\left(y\otimes v\right)=\underbrace{y_p}_{\substack{=\pi_{p,Y}(y)\\ \text{(by (13.19.58))}}}\otimes v=\pi_{p,Y}\left(y\right)\otimes v.$$

This proves Claim 1.]

[430] *Proof.* Applying the map $\pi_{p,Y}$ to both sides of the equality $y=\sum_{n\in\mathbb{N}}y_n$, we obtain

$$\begin{align}
\pi_{p,Y}\left(y\right) &= \pi_{p,Y}\left(\sum_{n\in\mathbb{N}}y_n\right)=\sum_{n\in\mathbb{N}}\pi_{p,Y}\left(y_n\right) \qquad \left(\text{since the map } \pi_{p,Y} \text{ is } \mathbf{k}\text{-linear}\right) \notag\\
&= \pi_{p,Y}\left(y_p\right)+\sum_{\substack{n\in\mathbb{N};\\ n\neq p}}\pi_{p,Y}\left(y_n\right) \tag{13.19.59}
\end{align}$$

(here, we have split off the addend for $n=p$ from the sum).

But (13.19.55) (applied to $n=p$) yields $y_p\in Y_p$. Recall that $\pi_{p,Y}$ is the canonical projection from Y to its p-th graded component Y_p. Hence, $\pi_{p,Y}$ fixes every element of Y_p. In other words, we have $\pi_{p,Y}\left(a\right)=a$ for each $a\in Y_p$. Applying this to $a=y_p$, we obtain $\pi_{p,Y}\left(y_p\right)=y_p$.

The map $\pi_{p,Y}$ is the canonical projection from Y to its p-th graded component Y_p. Hence, $\pi_{p,Y}$ annihilates all graded components of Y other than Y_p. In other words,

$$\pi_{p,Y}\left(Y_n\right)=0 \qquad \text{for every } n\in\mathbb{N} \text{ satisfying } n\neq p. \tag{13.19.60}$$

Now, each $n\in\mathbb{N}$ satisfying $n\neq p$ satisfies

$$\pi_{p,Y}\left(\underbrace{y_n}_{\substack{\in Y_n\\ \text{(by (13.19.55))}}}\right)\in\pi_{p,Y}\left(Y_n\right)=0 \qquad \text{(by (13.19.60))}$$

and thus

$$\pi_{p,Y}\left(y_n\right)=0. \tag{13.19.61}$$

Now, (13.19.59) becomes

$$\pi_{p,Y}\left(y\right)=\underbrace{\pi_{p,Y}\left(y_p\right)}_{=y_p}+\sum_{\substack{n\in\mathbb{N};\\ n\neq p}}\underbrace{\pi_{p,Y}\left(y_n\right)}_{\substack{=0\\ \text{(by (13.19.61))}}}=y_p+\underbrace{\sum_{\substack{n\in\mathbb{N};\\ n\neq p}}0}_{=0}=y_p.$$

[431] *Proof.* We have

$$\underbrace{y}_{=\sum_{n\in\mathbb{N}}y_n}\otimes v=\left(\sum_{n\in\mathbb{N}}y_n\right)\otimes v=\sum_{n\in\mathbb{N}}y_n\otimes v.$$

Applying the map $\pi_{p+q,Y\otimes V}$ to both sides of this equality, we obtain

$$\begin{align}
\pi_{p+q,Y\otimes V}\left(y\otimes v\right) &= \pi_{p+q,Y\otimes V}\left(\sum_{n\in\mathbb{N}}y_n\otimes v\right)=\sum_{n\in\mathbb{N}}\pi_{p+q,Y\otimes V}\left(y_n\otimes v\right) \qquad \left(\text{since the map } \pi_{p+q,Y\otimes V} \text{ is } \mathbf{k}\text{-linear}\right) \notag\\
&= \pi_{p+q,Y\otimes V}\left(y_p\otimes v\right)+\sum_{\substack{n\in\mathbb{N};\\ n\neq p}}\pi_{p+q,Y\otimes V}\left(y_n\otimes v\right) \tag{13.19.62}
\end{align}$$

(here, we have split off the addend for $n=p$ from the sum).

But (13.19.56) (applied to $n=p$) yields $y_p\otimes v\in\left(Y\otimes V\right)_{p+q}$. Recall that $\pi_{p+q,Y\otimes V}$ is the canonical projection from $Y\otimes V$ to its $(p+q)$-th graded component $\left(Y\otimes V\right)_{p+q}$. Hence, $\pi_{p+q,Y\otimes V}$ fixes every element of $\left(Y\otimes V\right)_{p+q}$. In other words, we have $\pi_{p+q,Y\otimes V}\left(a\right)=a$ for each $a\in\left(Y\otimes V\right)_{p+q}$. Applying this to $a=y_p\otimes v$, we obtain $\pi_{p+q,Y\otimes V}\left(y_p\otimes v\right)=y_p\otimes v$.

The map $\pi_{p+q,Y\otimes V}$ is the canonical projection from $Y\otimes V$ to its $(p+q)$-th graded component $\left(Y\otimes V\right)_{p+q}$. Hence, $\pi_{p+q,Y\otimes V}$ annihilates all graded components of $Y\otimes V$ other than $\left(Y\otimes V\right)_{p+q}$. In other words,

$$\pi_{p+q,Y\otimes V}\left(\left(Y\otimes V\right)_m\right)=0 \qquad \text{for every } m\in\mathbb{N} \text{ satisfying } m\neq p+q. \tag{13.19.63}$$

Now, if $n\in\mathbb{N}$ satisfies $n\neq p$, then $n+q\neq p+q$ (since $n\neq p$) and therefore

$$\pi_{p+q,Y\otimes V}\left(\left(Y\otimes V\right)_{n+q}\right)=0 \tag{13.19.64}$$

(by (13.19.63), applied to $m=n+q$).

Claim 2: Let $p \in \mathbb{N}$ and $q \in \mathbb{N}$. Let $x \in X_p$ and $u \in U_q$. Then,

$$\overline{g \otimes f}\left(x \otimes u\right) = \left(\overline{g} \otimes f\right)\left(x \otimes u\right).$$

[*Proof of Claim 2:* We have $u \in U_q$, thus $f\left(u\right) \in f\left(U_q\right) \subset V_q$ (since f is graded).

We have $x \in X_p$. Hence, Proposition 13.19.3 (applied to X, Y, X_m, g, p and x instead of V, W, V_m, f, n and v) yields $\overline{g}\left(x\right) = \pi_{p,Y}\left(g\left(x\right)\right)$. Hence, $\pi_{p,Y}\left(g\left(x\right)\right) = \overline{g}\left(x\right)$.

Applying (13.19.49) to $n = p + q$, we find

$$\left(X \otimes U\right)_{p+q} = \bigoplus_{\substack{(i,j)\in\mathbb{N}^2;\\ i+j=p+q}} X_i \otimes U_j. \tag{13.19.66}$$

But $X_p \otimes U_q$ is an addend of the direct sum $\bigoplus_{\substack{(i,j)\in\mathbb{N}^2;\\ i+j=p+q}} X_i \otimes U_j$ (namely, the addend for $(i,j) = (p,q)$), since $(p,q) \in \mathbb{N}^2$ and $p + q = p + q$. Hence, $X_p \otimes U_q \subset \bigoplus_{\substack{(i,j)\in\mathbb{N}^2;\\ i+j=p+q}} X_i \otimes U_j$. In view of (13.19.66), this rewrites as $X_p \otimes U_q \subset \left(X \otimes U\right)_{p+q}$.

From $x \in X_p$ and $u \in U_q$, we obtain $x \otimes u \in X_p \otimes U_q \subset \left(X \otimes U\right)_{p+q}$. Hence, Proposition 13.19.3 (applied to $X \otimes U$, $Y \otimes V$, $\left(X \otimes U\right)_m$, $g \otimes f$, $p + q$ and $x \otimes u$ instead of V, W, V_m, f, n and v) yields

$$\begin{aligned}
\overline{g \otimes f}\left(x \otimes u\right) &= \pi_{p+q,Y\otimes V}\left(\underbrace{\left(g \otimes f\right)\left(x \otimes u\right)}_{=g(x)\otimes f(u)}\right) = \pi_{p+q,Y\otimes V}\left(g\left(x\right) \otimes f\left(u\right)\right) \\
&= \underbrace{\pi_{p,Y}\left(g\left(x\right)\right)}_{=\overline{g}(x)} \otimes f\left(u\right) \qquad \left(\text{by Claim 1, applied to } y = g\left(x\right) \text{ and } v = f\left(u\right)\right) \\
&= \overline{g}\left(x\right) \otimes f\left(u\right) = \left(\overline{g} \otimes f\right)\left(x \otimes u\right) \qquad \left(\text{since } \left(\overline{g} \otimes f\right)\left(x \otimes u\right) = \overline{g}\left(x\right) \otimes f\left(u\right)\right).
\end{aligned}$$

This proves Claim 2.]

Claim 3: Let $p \in \mathbb{N}$ and $q \in \mathbb{N}$. Let $t \in X_p \otimes U_q$. Then,

$$\overline{g \otimes f}\left(t\right) = \left(\overline{g} \otimes f\right)\left(t\right).$$

[*Proof of Claim 3:* We have $t \in X_p \otimes U_q$. Hence, t is a $\mathbf{k}$-linear combination of pure tensors in $X_p \otimes U_q$. In other words, we can write t in the form

$$t = \sum_{i=1}^{k} x_i \otimes u_i \tag{13.19.67}$$

for some $k \in \mathbb{N}$, some elements $x_1, x_2, \ldots, x_k$ of X_p and some elements $u_1, u_2, \ldots, u_k$ of U_q. Consider this k, these $x_1, x_2, \ldots, x_k$ and these $u_1, u_2, \ldots, u_k$. For each $i \in \{1, 2, \ldots, k\}$, we have

$$\overline{g \otimes f}\left(x_i \otimes u_i\right) = \left(\overline{g} \otimes f\right)\left(x_i \otimes u_i\right) \tag{13.19.68}$$

(by Claim 2, applied to $x = x_i$ and $u = u_i$), since $x_i \in X_p$ and $u_i \in U_q$.

Now, each $n \in \mathbb{N}$ satisfying $n \neq p$ satisfies

$$\pi_{p+q,Y\otimes V}\left(\underbrace{y_n \otimes v}_{\substack{\in(Y\otimes V)_{n+q}\\ (\text{by } (13.19.56))}}\right) \in \pi_{p+q,Y\otimes V}\left(\left(Y \otimes V\right)_{n+q}\right) = 0 \qquad (\text{by } (13.19.64))$$

and thus

$$\pi_{p+q,Y\otimes V}\left(y_n \otimes v\right) = 0. \tag{13.19.65}$$

Now, (13.19.62) becomes

$$\pi_{p+q,Y\otimes V}\left(y \otimes v\right) = \underbrace{\pi_{p+q,Y\otimes V}\left(y_p \otimes v\right)}_{=y_p\otimes v} + \sum_{\substack{n\in\mathbb{N};\\ n\neq p}} \underbrace{\pi_{p+q,Y\otimes V}\left(y_n \otimes v\right)}_{\substack{=0\\ (\text{by } (13.19.65))}} = y_p \otimes v + \underbrace{\sum_{\substack{n\in\mathbb{N};\\ n\neq p}} 0}_{=0} = y_p \otimes v.$$

Applying the map $\overline{g\otimes f}$ to both sides of the equality (13.19.67), we find

$$\begin{aligned}\overline{g\otimes f}\,(t) &= \overline{g\otimes f}\left(\sum_{i=1}^{k} x_i\otimes u_i\right) = \sum_{i=1}^{k} \underbrace{\overline{g\otimes f}\,(x_i\otimes u_i)}_{\substack{=(\overline{g}\otimes f)(x_i\otimes u_i)\\ \text{(by (13.19.68))}}} && \left(\text{since the map } \overline{g\otimes f} \text{ is } \mathbf{k}\text{-linear}\right)\\ &= \sum_{i=1}^{k} (\overline{g}\otimes f)\,(x_i\otimes u_i) = (\overline{g}\otimes f)\left(\underbrace{\sum_{i=1}^{k} x_i\otimes u_i}_{\substack{=t\\ \text{(by (13.19.67))}}}\right) && (\text{since the map } \overline{g}\otimes f \text{ is } \mathbf{k}\text{-linear})\\ &= (\overline{g}\otimes f)\,(t)\,.\end{aligned}$$

This proves Claim 3.]

Claim 4: Let $t\in X\otimes U$. Then,

$$\overline{g\otimes f}\,(t) = (\overline{g}\otimes f)\,(t)\,.$$

[*Proof of Claim 4:* Let $t\in X\otimes U$. Then,

$$\begin{aligned} t \in \underbrace{X}_{\substack{=\bigoplus_{n\in\mathbb{N}} X_n=\sum_{n\in\mathbb{N}} X_n=\sum_{p\in\mathbb{N}} X_p\\ \text{(here, we have renamed the}\\ \text{summation index } n \text{ as } p)}} \otimes \underbrace{U}_{\substack{=\bigoplus_{n\in\mathbb{N}} U_n=\sum_{n\in\mathbb{N}} U_n=\sum_{q\in\mathbb{N}} U_q\\ \text{(here, we have renamed the}\\ \text{summation index } n \text{ as } q)}} &= \left(\sum_{p\in\mathbb{N}} X_p\right)\otimes\left(\sum_{q\in\mathbb{N}} U_q\right)\\ = \sum_{p\in\mathbb{N}}\sum_{q\in\mathbb{N}} X_p\otimes U_q = \sum_{(p,q)\in\mathbb{N}^2} X_p\otimes U_q. \end{aligned}$$

Hence, we can write t in the form

$$t = \sum_{(p,q)\in\mathbb{N}^2} t_{(p,q)} \tag{13.19.69}$$

for some family $\left(t_{(p,q)}\right)_{(p,q)\in\mathbb{N}^2} \in \prod_{(p,q)\in\mathbb{N}^2} X_p\otimes U_q$ such that all but finitely many $(p,q)\in\mathbb{N}^2$ satisfy $t_{(p,q)}=0$. Consider this family $\left(t_{(p,q)}\right)_{(p,q)\in\mathbb{N}^2}$.

Now, we have $\left(t_{(p,q)}\right)_{(p,q)\in\mathbb{N}^2} \in \prod_{(p,q)\in\mathbb{N}^2} X_p\otimes U_q$. In other words, $t_{(p,q)}\in X_p\otimes U_q$ for each $(p,q)\in\mathbb{N}^2$. Hence, for each $(p,q)\in\mathbb{N}^2$, we have

$$\overline{g\otimes f}\left(t_{(p,q)}\right) = (\overline{g}\otimes f)\left(t_{(p,q)}\right) \tag{13.19.70}$$

(by Claim 3, applied to $t=t_{(p,q)}$). Now, applying the map $\overline{g\otimes f}$ to both sides of the equality (13.19.69), we obtain

$$\begin{aligned}\overline{g\otimes f}\,(t) &= \overline{g\otimes f}\left(\sum_{(p,q)\in\mathbb{N}^2} t_{(p,q)}\right) = \sum_{(p,q)\in\mathbb{N}^2} \underbrace{\overline{g\otimes f}\left(t_{(p,q)}\right)}_{=(\overline{g}\otimes f)\left(t_{(p,q)}\right)} && \left(\text{since the map } \overline{g\otimes f} \text{ is } \mathbf{k}\text{-linear}\right)\\ &= \sum_{(p,q)\in\mathbb{N}^2} (\overline{g}\otimes f)\left(t_{(p,q)}\right) = (\overline{g}\otimes f)\left(\underbrace{\sum_{(p,q)\in\mathbb{N}^2} t_{(p,q)}}_{\substack{=t\\ \text{(by (13.19.69))}}}\right) && (\text{since the map } \overline{g}\otimes f \text{ is } \mathbf{k}\text{-linear})\\ &= (\overline{g}\otimes f)\,(t)\,.\end{aligned}$$

This proves Claim 4.]

Now, Claim 4 says that $\overline{g\otimes f}\,(t) = (\overline{g}\otimes f)\,(t)$ for every $t\in X\otimes U$. In other words, $\overline{g\otimes f} = \overline{g}\otimes f$. This proves Proposition 13.19.9. □

Proposition 13.19.10. *Let U, V, X and Y be four graded* $\mathbf{k}$*-modules. Let $f : U \to V$ and $g : X \to Y$ be two* $\mathbf{k}$*-linear maps such that f is graded. Then, $\overline{f \otimes g} = f \otimes \overline{g}$.*

Proof of Proposition 13.19.10. This proof is analogous to the proof of Proposition 13.19.9. (The only difference is that the order of the tensorands has been swapped.) □

Remark 13.19.11. Again, Proposition 13.19.10 and Proposition 13.19.9 cannot be generalized to the case when neither f nor g is graded.

We can now step to the solution of Exercise 1.4.29(e). Let C be a graded $\mathbf{k}$-coalgebra; let A be a graded $\mathbf{k}$-algebra; let $r : C \to A$ be a $\star$-invertible $\mathbf{k}$-linear map that is graded. We must prove that the $\star$-inverse $r^{\star(-1)}$ of r is also graded.

The $\star$-inverse $r^{\star(-1)}$ of r exists (since r is $\star$-invertible), and is a $\mathbf{k}$-linear map from C to A. Let $s = \overline{r^{\star(-1)}}$ (defined according to Definition 13.19.2). Thus, $\overline{r^{\star(-1)}} = s$.

The $\mathbf{k}$-linear map $\overline{r^{\star(-1)}} : C \to A$ is graded (by Proposition 13.19.5, applied to $V = C$, $W = A$ and $f = r^{\star(-1)}$). In other words, the $\mathbf{k}$-linear map $s : C \to A$ is graded (since $s = \overline{r^{\star(-1)}}$).

We know that $r^{\star(-1)}$ is the $\star$-inverse of r. Thus, $r^{\star(-1)} \star r = r \star r^{\star(-1)} = u_A \circ \epsilon_C$ (since $u_A \circ \epsilon_C$ is the unity of the convolution algebra $\operatorname{Hom}(C, A)$). Thus,

$$u_A \circ \epsilon_C = r^{\star(-1)} \star r = m_A \circ \left(r^{\star(-1)} \otimes r\right) \circ \Delta_C \tag{13.19.71}$$

(by the definition of convolution).

Proposition 13.19.9 (applied to $U = C$, $V = A$, $X = C$, $Y = A$, $f = r$ and $g = r^{\star(-1)}$) yields $\overline{r^{\star(-1)} \otimes r} = \underbrace{\overline{r^{\star(-1)}}}_{=s} \otimes r = s \otimes r$.

The map $m_A : A \otimes A \to A$ is graded (since A is a graded $\mathbf{k}$-algebra). Hence, Proposition 13.19.6 (applied to $U = C \otimes C$, $V = A \otimes A$, $W = A$, $f = m_A$ and $g = r^{\star(-1)} \otimes r$) yields $\overline{m_A \circ \left(r^{\star(-1)} \otimes r\right)} = m_A \circ \underbrace{\overline{r^{\star(-1)} \otimes r}}_{=s \otimes r} = m_A \circ (s \otimes r)$.

The map $\Delta_C : C \to C \otimes C$ is graded (since C is a graded $\mathbf{k}$-coalgebra). Thus, Proposition 13.19.7 (applied to $U = C$, $V = C \otimes C$, $W = A$, $f = m_A \circ \left(r^{\star(-1)} \otimes r\right)$ and $g = \Delta_C$) yields

$$\overline{m_A \circ \left(r^{\star(-1)} \otimes r\right) \circ \Delta_C} = \underbrace{\overline{m_A \circ \left(r^{\star(-1)} \otimes r\right)}}_{=m_A \circ (s \otimes r)} \circ \Delta_C = m_A \circ (s \otimes r) \circ \Delta_C = s \star r$$

(since the definition of convolution yields $s \star r = m_A \circ (s \otimes r) \circ \Delta_C$). In view of (13.19.71), we can rewrite this as

$$\overline{u_A \circ \epsilon_C} = s \star r. \tag{13.19.72}$$

But the map $\epsilon_C : C \to \mathbf{k}$ is graded (since C is a graded $\mathbf{k}$-coalgebra), and the map $u_A : \mathbf{k} \to A$ is graded as well (since A is a graded $\mathbf{k}$-algebra). Hence, the composition $u_A \circ \epsilon_C : C \to A$ of these two maps is also graded. Thus, Proposition 13.19.4 (applied to $V = C$, $W = A$ and $f = u_A \circ \epsilon_C$) yields $\overline{u_A \circ \epsilon_C} = u_A \circ \epsilon_C$. Comparing this with (13.19.72), we find

$$s \star r = u_A \circ \epsilon_C = (\text{the unity of the convolution algebra } \operatorname{Hom}(C, A)).$$

This yields that s is the $\star$-inverse of r (since we know that r is $\star$-invertible). In other words, $s = r^{\star(-1)}$. Since s is graded, we thus conclude that $r^{\star(-1)}$ is graded. This solves Exercise 1.4.29(e).

13.20. **Solution to Exercise 1.4.30.** *Solution to Exercise 1.4.30.* (a) Consider a map P satisfying the given assumption. Consider also the antipode S of A. Let T be the twist map $T_{A,A} : A \otimes A \to A \otimes A$, $c \otimes d \mapsto d \otimes c$. Then, $T \circ T = \operatorname{id}$. Moreover, if $f : A \to A$ and $g : A \to A$ are any two $\mathbf{k}$-linear maps, then

$$T \circ (f \otimes g) = (g \otimes f) \circ T. \tag{13.20.1}$$

(Indeed, this is a basic property of the twist map[432], and can easily be checked.)

[432] actually, a particular case of (13.19.4)

We know from Exercise 1.4.28 that S is a coalgebra anti-endomorphism; thus, $\Delta \circ S = T \circ (S \otimes S) \circ \Delta$ and $\epsilon \circ S = \epsilon$. Hence,

$$T \circ \underbrace{\Delta \circ S}_{=T\circ(S\otimes S)\circ\Delta} = \underbrace{T \circ T}_{=\mathrm{id}} \circ (S \otimes S) \circ \Delta = (S \otimes S) \circ \Delta,$$

so that $(S \otimes S) \circ \Delta = T \circ \Delta \circ S$. We also know from Proposition 1.4.10 that S is an algebra anti-endomorphism; this can be rewritten as $S \circ m = m \circ (S \otimes S) \circ T$ and $S \circ u = u$. Hence,

$$\underbrace{S \circ m}_{=m\circ(S\otimes S)\circ T} \circ T = m \circ (S \otimes S) \circ \underbrace{T \circ T}_{=\mathrm{id}} = m \circ (S \otimes S),$$

so that $m \circ (S \otimes S) = S \circ m \circ T$.

On the other hand, we know that $m \circ (P \otimes \mathrm{id}) \circ T \circ \Delta = u \circ \epsilon$ (indeed, this is just an element-free rewriting of the assumption that every $a \in A$ satisfies $\sum_{(a)} P(a_2) \cdot a_1 = u(\epsilon(a))$). Now, the definition of convolution shows that

$$\begin{aligned}(P \circ S) \star S &= m \circ \underbrace{((P \circ S) \otimes S)}_{=(P\otimes\mathrm{id})\circ(S\otimes S)} \circ \Delta = m \circ (P \otimes \mathrm{id}) \circ \underbrace{(S \otimes S) \circ \Delta}_{=T\circ\Delta\circ S} \\ &= \underbrace{m \circ (P \otimes \mathrm{id}) \circ T \circ \Delta}_{=u\circ\epsilon} \circ S = u \circ \underbrace{\epsilon \circ S}_{=\epsilon} = u \circ \epsilon,\end{aligned}$$

so that $P \circ S$ is a left $\star$-inverse to S. Since S is $\star$-invertible with $\star$-inverse id, this yields that $P \circ S = \mathrm{id}$. Furthermore, the definition of convolution shows that

$$\begin{aligned}S \star (S \circ P) &= m \circ \underbrace{(S \otimes (S \circ P))}_{=(S\otimes S)\circ(\mathrm{id}\otimes P)} \circ \Delta = \underbrace{m \circ (S \otimes S)}_{=S\circ m\circ T} \circ (\mathrm{id} \otimes P) \circ \Delta \\ &= S \circ m \circ \underbrace{T \circ (\mathrm{id} \otimes P)}_{\substack{=(P\otimes\mathrm{id})\circ T \\ (\text{by } (13.20.1))}} \circ \Delta = S \circ \underbrace{m \circ (P \otimes \mathrm{id}) \circ T \circ \Delta}_{=u\circ\epsilon} = \underbrace{S \circ u}_{=u} \circ \epsilon = u \circ \epsilon,\end{aligned}$$

whence $S \circ P$ is a right $\star$-inverse to S. Since S is $\star$-invertible with $\star$-inverse id, this yields that $S \circ P = \mathrm{id}$. Combined with $P \circ S = \mathrm{id}$, this yields that S is invertible and its inverse is P.

(b) Similar to the solution for (a), the details being left to the reader.

(c) Let A be a connected graded Hopf algebra. Just as a left $\star$-inverse S to id_A has been constructed in the proof of Proposition 1.4.16, we could construct a $\mathbf{k}$-linear map $P : A \to A$ such that every $a \in A$ satisfies

$$\sum_{(a)} P(a_2) \cdot a_1 = u(\epsilon(a)).$$

By part (a), this yields that the antipode of A is invertible.

13.21. **Solution to Exercise 1.4.32.** *Solution to Exercise 1.4.32.* We will be maximally explicit in this solution; in particular, we will not regard inclusions as identities even in cases where we would usually do that. We will only use the notations Δ and ϵ to denote the maps Δ_C and ϵ_C (not the maps Δ_D and ϵ_D, which we will introduce later).

We know that D is a direct summand of C as a $\mathbf{k}$-module. In other words, there exists a $\mathbf{k}$-submodule E of C such that $C = D \oplus E$. Fix such an E.

Let $i : D \to C$ be the canonical inclusion map. Let $p : D \oplus E \to D$ be the canonical projection from the direct sum $D \oplus E$ onto its summand D. Notice that p is a $\mathbf{k}$-linear map from $D \oplus E = C$ to D, and satisfies $p \circ i = \mathrm{id}_D$. Hence, $i \circ \underbrace{p \circ i}_{=\mathrm{id}_D} = i$.

In the statement of the exercise, we have assumed that $\Delta(D) \subset C \otimes D$. Since we don't want to abuse notation, we have to rewrite this as $\Delta(D) \subset (\mathrm{id}_C \otimes i)(C \otimes D)$ (because the $\mathbf{k}$-submodule of $C \otimes C$ spanned by tensors of the form $c \otimes d$ with $c \in C$ and $d \in D$ is precisely $(\mathrm{id}_C \otimes i)(C \otimes D)$). Similarly, $\Delta(D) \subset (i \otimes \mathrm{id}_C)(D \otimes C)$.

Now,

$$(i\otimes i)\left((p\otimes p)(x)\right)=x \qquad \text{for every } x\in\Delta(D). \tag{13.21.1}$$

Proof of (13.21.1)*:* Let $x\in\Delta(D)$. Then, $x\in\Delta(D)\subset(\mathrm{id}_C\otimes i)(C\otimes D)$. Hence, there exists some $y\in C\otimes D$ such that $x=(\mathrm{id}_C\otimes i)(y)$. Consider this y. We have

$$\begin{aligned}(i\otimes i)\left((p\otimes p)\left(\underbrace{x}_{=(\mathrm{id}_C\otimes i)(y)}\right)\right)&=(i\otimes i)\left((p\otimes p)\left((\mathrm{id}_C\otimes i)(y)\right)\right)=\underbrace{\left((i\otimes i)\circ(p\otimes p)\circ(\mathrm{id}_C\otimes i)\right)}_{=(i\circ p\circ\mathrm{id}_C)\otimes(i\circ p\circ i)}(y)\\&=\left(\underbrace{(i\circ p\circ\mathrm{id}_C)}_{=((i\circ p)\circ\mathrm{id}_C)}\otimes\underbrace{(i\circ p\circ i)}_{=i=\mathrm{id}_C\circ i}\right)(y)=\underbrace{\left(((i\circ p)\circ\mathrm{id}_C)\otimes(\mathrm{id}_C\circ i)\right)}_{=((i\circ p)\otimes\mathrm{id}_C)\circ(\mathrm{id}_C\otimes i)}(y)\\&=\left(((i\circ p)\otimes\mathrm{id}_C)\circ(\mathrm{id}_C\otimes i)\right)(y)=((i\circ p)\otimes\mathrm{id}_C)\left(\underbrace{(\mathrm{id}_C\otimes i)(y)}_{=x}\right)\\&=((i\circ p)\otimes\mathrm{id}_C)(x).\end{aligned}$$

On the other hand, $x\in\Delta(D)\subset(i\otimes\mathrm{id}_C)(D\otimes C)$. Thus, there exists some $z\in D\otimes C$ such that $x=(i\otimes\mathrm{id}_C)(z)$. Consider this z. We have

$$\begin{aligned}(i\otimes i)\left((p\otimes p)(x)\right)&=((i\circ p)\otimes\mathrm{id}_C)\left(\underbrace{x}_{=(i\otimes\mathrm{id}_C)(z)}\right)=((i\circ p)\otimes\mathrm{id}_C)\left((i\otimes\mathrm{id}_C)(z)\right)\\&=\underbrace{\left(((i\circ p)\otimes\mathrm{id}_C)\circ(i\otimes\mathrm{id}_C)\right)}_{=((i\circ p)\circ i)\otimes(\mathrm{id}_C\circ\mathrm{id}_C)}(z)=\left(\underbrace{((i\circ p)\circ i)}_{=i\circ p\circ i=i}\otimes\underbrace{(\mathrm{id}_C\circ\mathrm{id}_C)}_{=\mathrm{id}_C}\right)(z)\\&=(i\otimes\mathrm{id}_C)(z)=x.\end{aligned}$$

This proves (13.21.1).

Next, let us define a $\mathbf{k}$-linear map $\Delta_D:D\to D\otimes D$ by

$$\Delta_D=(p\otimes p)\circ\Delta\circ i.$$

Let us also define a $\mathbf{k}$-linear map $\epsilon_D:D\to\mathbf{k}$ by

$$\epsilon_D=\epsilon\circ i.$$

We will show that (D,Δ_D,ϵ_D) is a $\mathbf{k}$-coalgebra.[433] But first, let us see that

$$(i\otimes i)\circ\Delta_D=\Delta\circ i. \tag{13.21.2}$$

Proof of (13.21.2)*:* Let $d\in D$. Then, $i(d)=d$ (since i is just an inclusion map), so that $\Delta\left(\underbrace{i(d)}_{=d\in D}\right)\in\Delta(D)$. Hence, $(i\otimes i)\left((p\otimes p)(\Delta(i(d)))\right)=\Delta(i(d))$ (by (13.21.1), applied to $x=\Delta(i(d))$). Now,

$$\begin{aligned}\left((i\otimes i)\circ\underbrace{\Delta_D}_{=(p\otimes p)\circ\Delta\circ i}\right)(d)&=\left((i\otimes i)\circ(p\otimes p)\circ\Delta\circ i\right)(d)\\&=(i\otimes i)\left((p\otimes p)(\Delta(i(d)))\right)=\Delta(i(d))=(\Delta\circ i)(d).\end{aligned}$$

Now, forget that we fixed d. We thus have shown that every $d\in D$ satisfies $((i\otimes i)\circ\Delta_D)(d)=(\Delta\circ i)(d)$. In other words, $(i\otimes i)\circ\Delta_D=\Delta\circ i$. This proves (13.21.2).

[433] Recall that we are not going to abbreviate Δ_D and ϵ_D by Δ and ϵ; thus, Δ and ϵ still mean the maps Δ_C and ϵ_C.

Now, let us check that $(D, \Delta_D, \epsilon_D)$ is a $\mathbf{k}$-coalgebra. In order to do so, we must check that the diagrams

(13.21.3)
$$\begin{tikzcd}
& D\otimes D\otimes D & \\
D\otimes D \arrow[ur, "\Delta_D\otimes\mathrm{id}_D"] & & D\otimes D \arrow[ul, "\mathrm{id}_D\otimes\Delta_D"'] \\
& D \arrow[ul, "\Delta_D"'] \arrow[ur, "\Delta_D"] &
\end{tikzcd}$$

and

(13.21.4)
$$\begin{tikzcd}
D\otimes\mathbf{k} \arrow[r, "\cong"] & D & \mathbf{k}\otimes D \arrow[l, "\cong"'] \\
D\otimes D \arrow[u, "\mathrm{id}_D\otimes\epsilon_D"] & D \arrow[l, "\Delta_D"] \arrow[r, "\Delta_D"'] \arrow[u, "\mathrm{id}_D"] & D\otimes D \arrow[u, "\epsilon_D\otimes\mathrm{id}_D"']
\end{tikzcd}$$

are commutative (where the canonical isomorphisms $\mathbf{k}\otimes D\to D$ and $D\otimes\mathbf{k}\to D$ are used in (13.21.4)). Let us start with the first diagram.

Since the diagram (1.2.1) for C is commutative (as C is a $\mathbf{k}$-coalgebra), we have $(\Delta\otimes\mathrm{id}_C)\circ\Delta = (\mathrm{id}_C\otimes\Delta)\circ\Delta$. But

$$\underbrace{(i\otimes i\otimes i)\circ(\Delta_D\otimes\mathrm{id}_D)}_{\substack{=((i\otimes i)\otimes i)\circ(\Delta_D\otimes\mathrm{id}_D)\\ =((i\otimes i)\circ\Delta_D)\otimes(i\circ\mathrm{id}_D)}}\circ\Delta_D = \left(\underbrace{((i\otimes i)\circ\Delta_D)}_{\substack{=\Delta\circ i\\ \text{(by (13.21.2))}}}\otimes\underbrace{(i\circ\mathrm{id}_D)}_{=i=\mathrm{id}_C\circ i}\right)\circ\Delta_D = \underbrace{((\Delta\circ i)\otimes(\mathrm{id}_C\circ i))}_{=(\Delta\otimes\mathrm{id}_C)\circ(i\otimes i)}\circ\Delta_D$$
$$= (\Delta\otimes\mathrm{id}_C)\circ\underbrace{(i\otimes i)\circ\Delta_D}_{\substack{=\Delta\circ i\\ \text{(by (13.21.2))}}} = (\Delta\otimes\mathrm{id}_C)\circ\Delta\circ i.$$

Using this and the analogously provable identity $(i\otimes i\otimes i)\circ(\mathrm{id}_D\otimes\Delta_D)\circ\Delta_D = (\mathrm{id}_C\otimes\Delta)\circ\Delta\circ i$, we obtain

$$(i\otimes i\otimes i)\circ(\Delta_D\otimes\mathrm{id}_D)\circ\Delta_D = \underbrace{(\Delta\otimes\mathrm{id}_C)\circ\Delta}_{=(\mathrm{id}_C\otimes\Delta)\circ\Delta}\circ i = (\mathrm{id}_C\otimes\Delta)\circ\Delta\circ i = (i\otimes i\otimes i)\circ(\mathrm{id}_D\otimes\Delta_D)\circ\Delta_D.$$

We can cancel the left $i\otimes i\otimes i$ factor from this equation[434], and thus obtain $(\Delta_D\otimes\mathrm{id}_D)\circ\Delta_D = (\mathrm{id}_D\otimes\Delta_D)\circ\Delta_D$. In other words, the diagram (13.21.3) is commutative.

Let us now check that the diagram (13.21.4) is commutative. We will only prove this for its left square, leaving the (completely analogous) right square to the reader. For every $\mathbf{k}$-module V, we let can_V denote the canonical $\mathbf{k}$-module isomorphism $V\otimes\mathbf{k}\to V$. The upper left horizontal arrow on diagram (13.21.4) is precisely can_D. Notice that $\mathrm{can}_C\circ(\mathrm{id}_C\otimes\epsilon)\circ\Delta = \mathrm{id}_C$ due to the commutativity of the diagram (1.2.2) (which, of course, commutes since C is a coalgebra).

Now, any two $\mathbf{k}$-modules V and W and any $\mathbf{k}$-linear map $f : V\to W$ satisfy

$$f\circ\mathrm{can}_V = \mathrm{can}_W\circ(f\otimes\mathrm{id}_{\mathbf{k}})$$

[434] because $i\otimes i\otimes i$ is left-invertible:

$$(p\otimes p\otimes p)\circ(i\otimes i\otimes i) = \underbrace{(p\circ i)}_{=\mathrm{id}_D}\otimes\underbrace{(p\circ i)}_{=\mathrm{id}_D}\otimes\underbrace{(p\circ i)}_{=\mathrm{id}_D} = \mathrm{id}_D\otimes\mathrm{id}_D\otimes\mathrm{id}_D = \mathrm{id}_{D\otimes D\otimes D}.$$

(this is just trivial linear algebra). This (applied to $V = D$, $W = C$ and $f = i$) yields $i \circ \operatorname{can}_D = \operatorname{can}_C \circ (i \otimes \operatorname{id}_{\mathbf{k}})$. Hence,

$$\begin{aligned} \underbrace{i \circ \operatorname{can}_D}_{=\operatorname{can}_C \circ (i \otimes \operatorname{id}_{\mathbf{k}})} \circ (\operatorname{id}_D \otimes \epsilon_D) \circ \Delta_D &= \operatorname{can}_C \circ \underbrace{(i \otimes \operatorname{id}_{\mathbf{k}}) \circ (\operatorname{id}_D \otimes \epsilon_D)}_{=(i\circ \operatorname{id}_D)\otimes(\operatorname{id}_{\mathbf{k}} \circ \epsilon_D)} \circ \Delta_D = \operatorname{can}_C \circ \left(\underbrace{(i\circ \operatorname{id}_D)}_{=i=\operatorname{id}_C \circ i} \otimes \underbrace{(\operatorname{id}_{\mathbf{k}} \circ \epsilon_D)}_{=\epsilon_D = \epsilon\circ i} \right) \circ \Delta_D \\ &= \operatorname{can}_C \circ \underbrace{((\operatorname{id}_C \circ i) \otimes (\epsilon \circ i))}_{=(\operatorname{id}_C \otimes \epsilon)\circ(i\otimes i)} \circ \Delta_D = \operatorname{can}_C \circ (\operatorname{id}_C \otimes \epsilon) \circ \underbrace{((i\otimes i)\circ \Delta_D)}_{\substack{=\Delta\circ i \\ \text{(by (13.21.2))}}} \\ &= \underbrace{\operatorname{can}_C \circ (\operatorname{id}_C \otimes \epsilon) \circ \Delta}_{=\operatorname{id}_C} \circ i = \operatorname{id}_C \circ i = i. \end{aligned}$$

We can cancel the i factors from this equation (because i is left-invertible: $p \circ i = \operatorname{id}_D$), and thus are left with $\operatorname{can}_D \circ (\operatorname{id}_D \otimes \epsilon_D) \circ \Delta_D = \operatorname{id}_D$. This means precisely that the left square of (13.21.4) is commutative. As we said, this completes the verification of the fact that $(D, \Delta_D, \epsilon_D)$ must be a $\mathbf{k}$-coalgebra. We will denote this $\mathbf{k}$-coalgebra simply by D.

Now, the diagrams

$$\begin{tikzcd} D \arrow[r, "i"] \arrow[d, "\Delta_D"'] & C \arrow[d, "\Delta"] \\ D \otimes D \arrow[r, "i \otimes i"] & C \otimes C \end{tikzcd} \qquad \text{and} \qquad \begin{tikzcd} D \arrow[rr, "i"] \arrow[dr, "\epsilon_D"'] & & C \arrow[dl, "\epsilon"] \\ & \mathbf{k} & \end{tikzcd}$$

are commutative[435]. Hence, i is a $\mathbf{k}$-coalgebra homomorphism (by the definition of a $\mathbf{k}$-coalgebra homomorphism). In other words, the canonical inclusion map $D \to C$ is a $\mathbf{k}$-coalgebra homomorphism (since i is the canonical inclusion map $D \to C$).

So we know that D is a $\mathbf{k}$-coalgebra such that $D \subset C$ and such that the canonical inclusion map $D \to C$ is a $\mathbf{k}$-coalgebra homomorphism. In other words, D is a $\mathbf{k}$-subcoalgebra of C (by the definition of a $\mathbf{k}$-subcoalgebra). This solves the exercise.

13.22. **Solution to Exercise 1.4.33.** *Solution to Exercise 1.4.33.* We will identify the tensor products $K \otimes K$, $C \otimes K$ and $K \otimes C$ with the corresponding $\mathbf{k}$-submodules of the tensor product $C \otimes C$. (We can afford to do this since $\mathbf{k}$ is a field.)

We will only use the notations Δ and ϵ to denote the maps Δ_C and ϵ_C.

Notice that C is a free $\mathbf{k}$-module. Hence, tensoring with C is an exact functor, thus a left-exact functor.

By the recursive definition of $\Delta^{(1)}$, we have $\Delta^{(1)} = \left(\operatorname{id}_C \otimes \underbrace{\Delta^{(1-1)}}_{=\Delta^{(0)}=\operatorname{id}_C} \right) \circ \Delta = \underbrace{(\operatorname{id}_C \otimes \operatorname{id}_C)}_{=\operatorname{id}_{C\otimes C}} \circ \Delta = \Delta$.

By the recursive definition of $\Delta^{(2)}$, we have $\Delta^{(2)} = \left(\operatorname{id}_C \otimes \underbrace{\Delta^{(1)}}_{=\Delta} \right) \circ \Delta = (\operatorname{id}_C \otimes \Delta) \circ \Delta = (\Delta \otimes \operatorname{id}_C) \circ \Delta$ (by the axioms of a coalgebra).

(a) Applying Exercise 1.4.20(b) to $k = 3$, we obtain $\Delta^{(3)} = \left(\Delta^{(2)} \otimes \operatorname{id}_C\right) \circ \Delta$, so that

$$\left(\Delta^{(2)} \otimes \operatorname{id}_C\right) \circ \Delta = \Delta^{(3)} = \left(\operatorname{id}_C \otimes \Delta^{(2)}\right) \circ \Delta \tag{13.22.1}$$

(by the recursive definition of $\Delta^{(3)}$).

[435] In fact, the commutativity of the first of these diagrams follows from (13.21.2), whereas the commutativity of the second diagram follows from $\epsilon_D = \epsilon \circ i$.

Let $\widetilde{f} = (\mathrm{id}_C \otimes f \otimes \mathrm{id}_C) \circ \Delta^{(2)} : C \to C \otimes U \otimes C$. Then, $\ker \widetilde{f} = \ker \left((\mathrm{id}_C \otimes f \otimes \mathrm{id}_C) \circ \Delta^{(2)} \right) = K$. But

$$\begin{aligned}
&\underbrace{\widetilde{f}}_{=(\mathrm{id}_C \otimes f \otimes \mathrm{id}_C) \circ \Delta^{(2)}} \otimes \mathrm{id}_C \\
&= \left((\mathrm{id}_C \otimes f \otimes \mathrm{id}_C) \circ \Delta^{(2)} \right) \otimes \mathrm{id}_C = \underbrace{((\mathrm{id}_C \otimes f \otimes \mathrm{id}_C) \otimes \mathrm{id}_C)}_{\substack{=\mathrm{id}_C \otimes f \otimes \mathrm{id}_C \otimes \mathrm{id}_C \\ =\mathrm{id}_C \otimes f \otimes \mathrm{id}_{C \otimes C}}} \circ \left(\Delta^{(2)} \otimes \mathrm{id}_C \right) \\
&= (\mathrm{id}_C \otimes f \otimes \mathrm{id}_{C \otimes C}) \circ \left(\Delta^{(2)} \otimes \mathrm{id}_C \right),
\end{aligned}$$

so that

$$\begin{aligned}
&\underbrace{\left(\widetilde{f} \otimes \mathrm{id}_C \right)}_{=(\mathrm{id}_C \otimes f \otimes \mathrm{id}_{C \otimes C}) \circ \left(\Delta^{(2)} \otimes \mathrm{id}_C \right)} \circ \Delta \\
&= (\mathrm{id}_C \otimes f \otimes \mathrm{id}_{C \otimes C}) \circ \underbrace{\left(\Delta^{(2)} \otimes \mathrm{id}_C \right) \circ \Delta}_{\substack{=\left(\mathrm{id}_C \otimes \Delta^{(2)} \right) \circ \Delta \\ \text{(by (13.22.1))}}} \\
&= \underbrace{(\mathrm{id}_C \otimes f \otimes \mathrm{id}_{C \otimes C}) \circ \left(\mathrm{id}_C \otimes \Delta^{(2)} \right)}_{=\mathrm{id}_C \otimes \left((f \otimes \mathrm{id}_{C \otimes C}) \circ \Delta^{(2)} \right)} \circ \Delta \\
&= \left(\mathrm{id}_C \otimes \left((f \otimes \mathrm{id}_{C \otimes C}) \circ \underbrace{\Delta^{(2)}}_{=(\mathrm{id}_C \otimes \Delta) \circ \Delta} \right) \right) \circ \Delta \\
&= \left(\mathrm{id}_C \otimes \left(\underbrace{(f \otimes \mathrm{id}_{C \otimes C}) \circ (\mathrm{id}_C \otimes \Delta)}_{\substack{=f \otimes \Delta \\ =(\mathrm{id}_U \otimes \Delta) \circ (f \otimes \mathrm{id}_C)}} \circ \Delta \right) \right) \circ \Delta \\
&= \underbrace{(\mathrm{id}_C \otimes ((\mathrm{id}_U \otimes \Delta) \circ (f \otimes \mathrm{id}_C) \circ \Delta))}_{=(\mathrm{id}_C \otimes (\mathrm{id}_U \otimes \Delta)) \circ (\mathrm{id}_C \otimes (f \otimes \mathrm{id}_C)) \circ (\mathrm{id}_C \otimes \Delta)} \circ \Delta \\
&= \underbrace{(\mathrm{id}_C \otimes (\mathrm{id}_U \otimes \Delta))}_{=\mathrm{id}_C \otimes \mathrm{id}_U \otimes \Delta} \circ \underbrace{(\mathrm{id}_C \otimes (f \otimes \mathrm{id}_C))}_{=\mathrm{id}_C \otimes f \otimes \mathrm{id}_C} \circ \underbrace{(\mathrm{id}_C \otimes \Delta) \circ \Delta}_{=\Delta^{(2)}} \\
&= (\mathrm{id}_C \otimes \mathrm{id}_U \otimes \Delta) \circ \underbrace{(\mathrm{id}_C \otimes f \otimes \mathrm{id}_C) \circ \Delta^{(2)}}_{=\widetilde{f}} \\
&= (\mathrm{id}_C \otimes \mathrm{id}_U \otimes \Delta) \circ \widetilde{f}.
\end{aligned}$$

Hence,

$$\ker \left(\left(\widetilde{f} \otimes \mathrm{id}_C \right) \circ \Delta \right) = \ker \left((\mathrm{id}_C \otimes \mathrm{id}_U \otimes \Delta) \circ \widetilde{f} \right) \supset \ker \widetilde{f} = K,$$

so that

$$K \subset \ker \left(\left(\widetilde{f} \otimes \mathrm{id}_C \right) \circ \Delta \right) = \Delta^{-1} \left(\ker \left(\widetilde{f} \otimes \mathrm{id}_C \right) \right),$$

so that

$$\begin{aligned}
\Delta(K) \subset \ker \left(\widetilde{f} \otimes \mathrm{id}_C \right) &= \underbrace{\left(\ker \widetilde{f} \right)}_{=K} \otimes C \qquad \text{(since tensoring with } C \text{ is a left-exact functor)} \\
&= K \otimes C.
\end{aligned}$$

Similarly, $\Delta(K) \subset C \otimes K$. But K is a direct summand of C as a $\mathbf{k}$-module (since $\mathbf{k}$ is a field). Hence, Exercise 1.4.32 (applied to $D = K$) yields that there is a canonically defined $\mathbf{k}$-coalgebra structure on K which makes K a subcoalgebra of C. In other words, K is a $\mathbf{k}$-subcoalgebra of C. This solves Exercise 1.4.33(a).

(b) Let E be a $\mathbf{k}$-subcoalgebra of C which is a subset of $\ker f$. We must show that E is a subset of K. We have $E \subset \ker f$, so that $f(E) = 0$. Since E is a $\mathbf{k}$-subcoalgebra of C, we have $\Delta(E) \subset E \otimes E$, and

$$\underbrace{\Delta^{(2)}}_{=(\mathrm{id}_C \otimes \Delta)\circ\Delta}(E) = ((\mathrm{id}_C \otimes \Delta) \circ \Delta)(E) = (\mathrm{id}_C \otimes \Delta)\left(\underbrace{\Delta(E)}_{\subset E\otimes E}\right)$$
$$\subset (\mathrm{id}_C \otimes \Delta)(E \otimes E) = \underbrace{\mathrm{id}_C(E)}_{=E} \otimes \underbrace{\Delta(E)}_{\subset E\otimes E} \subset E \otimes E \otimes E$$

and thus

$$\left((\mathrm{id}_C \otimes f \otimes \mathrm{id}_C) \circ \Delta^{(2)}\right)(E) = (\mathrm{id}_C \otimes f \otimes \mathrm{id}_C)\left(\underbrace{\Delta^{(2)}(E)}_{\subset E\otimes E\otimes E}\right)$$
$$\subset (\mathrm{id}_C \otimes f \otimes \mathrm{id}_C)(E \otimes E \otimes E)$$
$$= \mathrm{id}_C(E) \otimes \underbrace{f(E)}_{=0} \otimes \mathrm{id}_C(E) = 0.$$

Hence, $E \subset \ker\left((\mathrm{id}_C \otimes f \otimes \mathrm{id}_C) \circ \Delta^{(2)}\right) = K$, which means that E is a subset of K. This solves Exercise 1.4.33(b).

Remark: Exercise 1.4.33(a) would not hold if we allowed $\mathbf{k}$ to be an arbitrary commutative ring rather than a field. (For a stupid counterexample, try $\mathbf{k} = \mathbb{Z}$, $C = \mathbf{k}$ and $U = \mathbb{Z}/2$, with f being the canonical projection.) It might be an interesting question to figure out how much freedom we can allow without breaking correctness. Here is one case which definitely works: If $\mathbf{k}$ is a principal ideal domain and C and U are finite free $\mathbf{k}$-modules, then Exercise 1.4.33 is correct. (In fact, our above solution works in this case, after one notices that $(\mathrm{id}_C \otimes f \otimes \mathrm{id}_C) \circ \Delta^{(2)}$ is a homomorphism of finite free $\mathbf{k}$-modules, and the kernel of every homomorphism of finite free $\mathbf{k}$-modules over a principal ideal domain is a direct summand of its domain.)

13.23. **Solution to Exercise 1.4.34.** *Solution to Exercise 1.4.34.*

Lemma 13.23.1. *Let C be a $\mathbf{k}$-coalgebra, and let A be a $\mathbf{k}$-algebra. Let D be a subcoalgebra of C. Let $f : C \to A$ and $g : C \to A$ be two $\mathbf{k}$-linear maps. Then, $(f \star g)\mid_D = (f\mid_D) \star (g\mid_D)$.*

Proof of Lemma 13.23.1. Recall the following classical formula from linear algebra:

- If U, V, W, U', V' and W' are six $\mathbf{k}$-modules and $\alpha : U \to V$, $\beta : V \to W$, $\alpha' : U' \to V'$ and $\beta' : V' \to W'$ are four $\mathbf{k}$-linear maps, then $(\beta \circ \alpha) \otimes (\beta' \circ \alpha') = (\beta \otimes \beta') \circ (\alpha \otimes \alpha')$.

We will refer to this formula as the *composition-tensor relation.*

Let $i : D \to C$ be the inclusion map from D to C. Then, i is a $\mathbf{k}$-coalgebra homomorphism (since D is a subcoalgebra of C). In other words, i is a $\mathbf{k}$-linear map that makes the diagrams

$$\begin{tikzcd} D \arrow[r, "i"] \arrow[d, "\Delta_D"'] & C \arrow[d, "\Delta_C"] \\ D \otimes D \arrow[r, "i\otimes i"] & C \otimes C \end{tikzcd} \qquad \begin{tikzcd} D \arrow[rr, "i"] \arrow[rd, "\epsilon_D"'] & & C \arrow[ld, "\epsilon_C"] \\ & \mathbf{k} & \end{tikzcd}$$

commute. In other words, i is a $\mathbf{k}$-linear map satisfying $(i \otimes i) \circ \Delta_D = \Delta_C \circ i$ and $\epsilon_D = \epsilon_C \circ i$.

Recall that i is the inclusion map from D to C. Hence, if $h : C \to A$ is any map, then

$$h\mid_D = h \circ i. \tag{13.23.1}$$

Now, the definition of convolution yields

$$\begin{aligned}(f\mid_D)\star(g\mid_D) &= m_A\circ\left(\underbrace{(f\mid_D)}_{\substack{=f\circ i\\ \text{(by (13.23.1),}\\ \text{applied to } h=f)}}\otimes\underbrace{(g\mid_D)}_{\substack{=g\circ i\\ \text{(by (13.23.1),}\\ \text{applied to } h=g)}}\right)\circ\Delta_D\\ &= m_A\circ\underbrace{((f\circ i)\otimes(g\circ i))}_{\substack{=(f\otimes g)\circ(i\otimes i)\\ \text{(by the composition-tensor relation)}}}\circ\Delta_D\\ &= m_A\circ(f\otimes g)\circ\underbrace{(i\otimes i)\circ\Delta_D}_{=\Delta_C\circ i} = m_A\circ(f\otimes g)\circ\Delta_C\circ i.\end{aligned}$$

Comparing this with

$$\begin{aligned}(f\star g)\mid_D &= \underbrace{(f\star g)}_{\substack{=m_A\circ(f\otimes g)\circ\Delta_C\\ \text{(by the definition}\\ \text{of convolution)}}}\circ i \qquad \text{(by (13.23.1), applied to } h=f\star g)\\ &= m_A\circ(f\otimes g)\circ\Delta_C\circ i,\end{aligned}$$

we obtain $(f\star g)\mid_D=(f\mid_D)\star(g\mid_D)$. This proves Lemma 13.23.1. □

Our next lemma is a particular case of Exercise 1.4.34(a), which we will use as a stepping stone towards solving the exercise in the general case:

Lemma 13.23.2. *Let $C=\bigoplus_{n\geq 0}C_n$ be a graded $\mathbf{k}$-coalgebra, and A be any $\mathbf{k}$-algebra. Let $h:C\to A$ be a $\mathbf{k}$-linear map such that $h\mid_{C_0}=(u_A\circ\epsilon_C)\mid_{C_0}$. Then, h is a $\star$-invertible map in $\operatorname{Hom}(C,A)$.*

Proof of Lemma 13.23.2. Define a $\mathbf{k}$-linear map $f:C\to A$ as $h-u_A\circ\epsilon_C$. Thus, $h=u_A\circ\epsilon_C+f=u\epsilon+f$. Furthermore, the map f annihilates C_0, because

$$\underbrace{f}_{=h-u_A\circ\epsilon_C}\mid_{C_0}=(h-u_A\circ\epsilon_C)\mid_{C_0}=(h\mid_{C_0})-((u_A\circ\epsilon_C)\mid_{C_0})=0 \qquad \left(\text{since } h\mid_{C_0}=(u_A\circ\epsilon_C)\mid_{C_0}\right).$$

Thus, we can proceed as in the proof of Proposition 1.4.24 to show that $\sum_{k\geq 0}(-1)^k f^{\star k}$ is a well-defined linear map $C\to A$ and a two-sided $\star$-inverse for $u\epsilon+f$. Thus, $u\epsilon+f$ is $\star$-invertible. In other words, h is $\star$-invertible (since $h=u\epsilon+f$). This proves Lemma 13.23.2.

An alternative proof of Lemma 13.23.2 proceeds by mimicking the proof of Proposition 1.4.16. □

We now step to the solution of the exercise.

(a) Here is Takeuchi's argument: We know that the map $h\mid_{C_0}\in\operatorname{Hom}(C_0,A)$ is $\star$-invertible; let $\widetilde{g}$ be its $\star$-inverse. Thus, $\widetilde{g}:C_0\to A$ is a $\mathbf{k}$-linear map satisfying

$$\widetilde{g}\star(h\mid_{C_0})=(h\mid_{C_0})\star\widetilde{g}=u_A\circ\underbrace{\epsilon_{C_0}}_{=\epsilon_C\mid_{C_0}}=u_A\circ(\epsilon_C\mid_{C_0})=(u_A\circ\epsilon_C)\mid_{C_0}.$$

Extend $\widetilde{g}$ to a $\mathbf{k}$-linear map $g:C\to A$ by defining it to be 0 on every C_n for $n>0$. Then, $g\mid_{C_0}=\widetilde{g}$, so that

$$\begin{aligned}(g\star h)\mid_{C_0} &= \underbrace{(g\mid_{C_0})}_{=\widetilde{g}}\star(h\mid_{C_0}) \qquad \text{(by Lemma 13.23.1, applied to } C_0,\ g \text{ and } h \text{ instead of } D,\ f \text{ and } g)\\ &= \widetilde{g}\star(h\mid_{C_0})=(u_A\circ\epsilon_C)\mid_{C_0}\end{aligned}$$

and similarly

$$(h\star g)\mid_{C_0}=(u_A\circ\epsilon_C)\mid_{C_0}.$$

Now, Lemma 13.23.2 (applied to $g\star h$ instead of h) shows that $g\star h$ is a $\star$-invertible map in $\operatorname{Hom}(C,A)$ (since $(g\star h)\mid_{C_0}=(u_A\circ\epsilon_C)\mid_{C_0}$). In other words, there exists a map $p\in\operatorname{Hom}(C,A)$ such that $(g\star h)\star p=p\star(g\star h)=u_A\epsilon_C$. Consider this p. Also, Lemma 13.23.2 (applied to $h\star g$ instead of h) shows that $h\star g$ is a $\star$-invertible map in $\operatorname{Hom}(C,A)$ (since $(h\star g)\mid_{C_0}=(u_A\circ\epsilon_C)\mid_{C_0}$). In other words, there exists a map

$q \in \operatorname{Hom}(C, A)$ such that $(h \star g) \star q = q \star (h \star g) = u_A \epsilon_C$. Consider this q. We conclude that $p \star g$ is a left $\star$-inverse to h (since $(p \star g) \star h = p \star (g \star h) = u_A \epsilon_C$), so that the map h has a left $\star$-inverse. We also conclude that $g \star q$ is a right $\star$-inverse to h (since $h \star (g \star q) = (h \star g) \star q = u_A \epsilon_C$), so that the map h has a right $\star$-inverse. Altogether, we now know that the map h has a left $\star$-inverse and a right $\star$-inverse; therefore, it is $\star$-invertible (because an element of an algebra that has a left inverse and a right inverse must always be invertible). This solves part (a) of the exercise.

(b) Apply part (a) to $C = A$ and the map $\operatorname{id}_A : A \to A$.

(c) Let A be a connected graded bialgebra. We must show that A has a unique antipode S, which is a graded map $A \xrightarrow{S} A$, and that this endows A with a Hopf structure.

The map $u : \mathbf{k} \to A$ is a $\mathbf{k}$-algebra homomorphism and a $\mathbf{k}$-coalgebra homomorphism. Hence, u is a $\mathbf{k}$-bialgebra homomorphism. Hence, $\mathbf{k} \xrightarrow{u} A_0$ is a $\mathbf{k}$-bialgebra homomorphism. Since we know from Exercise 1.3.20(b) that the map u is an isomorphism $\mathbf{k} \xrightarrow{u} A_0$, we thus conclude that $\mathbf{k} \xrightarrow{u} A_0$ is a $\mathbf{k}$-bialgebra isomorphism. Hence, $A_0 \cong \mathbf{k}$ as $\mathbf{k}$-bialgebras. Since $\mathbf{k}$ is a Hopf algebra (with antipode $\operatorname{id}_{\mathbf{k}}$), we thus conclude that A_0 is a Hopf algebra. Hence, part (b) of this exercise shows that A is a Hopf algebra. Therefore, A has an antipode. It is clear that this antipode is unique (since an antipode of A is the same thing as a $\star$-inverse to the map $\operatorname{id} : A \to A$, but it is clear that the $\star$-inverse of any given map is unique). Thus, we know that A has a unique antipode S, endowing it with a Hopf structure. In order to conclude the (new) proof of Proposition 1.4.16, it thus remains to show that this antipode S is graded.

The antipode S is the $\star$-inverse of the map $\operatorname{id} : A \to A$. Hence, the map $\operatorname{id} : A \to A$ is $\star$-invertible. This map id is also graded (obviously). Thus, Exercise 1.4.29(e) (applied to $C = A$ and $r = \operatorname{id}$) shows that the $\star$-inverse $\operatorname{id}^{\star(-1)}$ of id is also graded. In other words, S is graded (since the $\star$-inverse $\operatorname{id}^{\star(-1)}$ of id is S). This concludes the (new) proof of Proposition 1.4.16. Thus, part (c) of the exercise is solved.

13.24. **Solution to Exercise 1.4.35.** *Solution to Exercise 1.4.35.* (a) Let I be a two-sided coideal of A such that $I \cap \mathfrak{p} = 0$ and such that $I = \bigoplus_{n \geq 0} (I \cap A_n)$. Let $I_n = I \cap A_n$ for every $n \in \mathbb{N}$. Then, $I = \bigoplus_{n \geq 0} \underbrace{(I \cap A_n)}_{= I_n} = \bigoplus_{n \geq 0} I_n$.

Since I is a two-sided coideal, we have $\epsilon(I) = 0$.

We now will prove that

$$\text{every } n \in \mathbb{N} \text{ satisfies } I_n = 0. \tag{13.24.1}$$

Proof of (13.24.1)*:* Let us prove (13.24.1) by strong induction over n:

Let $N \in \mathbb{N}$. Assume that (13.24.1) holds for every $n \in \mathbb{N}$ satisfying $n < N$. We must then prove that (13.24.1) holds for $n = N$.

We know that (13.24.1) holds for every $n \in \mathbb{N}$ satisfying $n < N$. In other words,

$$\text{for every } n \in \mathbb{N} \text{ satisfying } n < N, \text{ we have } I_n = 0. \tag{13.24.2}$$

We have $I_N = I \cap A_N \subset I$ and thus $\epsilon \left(\underbrace{I_N}_{\subset I} \right) \subset \epsilon(I) = 0$, hence $\epsilon(I_N) = 0$.

Let $\pi_N : A \otimes A \to (A \otimes A)_N$ be the projection from the graded $\mathbf{k}$-module $A \otimes A$ to its N-th graded component $(A \otimes A)_N$. Then,

$$\pi_N(t) = t \qquad \text{for every } t \in (A \otimes A)_N, \tag{13.24.3}$$

and

$$\pi_N(t) = 0 \qquad \text{for every } \ell \in \mathbb{N} \setminus \{N\} \text{ and every } t \in (A \otimes A)_\ell. \tag{13.24.4}$$

As a consequence,

$$\text{every } (n, m) \in \mathbb{N}^2 \text{ such that } n + m \neq N \text{ satisfy } \pi_N(A_n \otimes A_m) = 0. \tag{13.24.5}$$

Let $i \in I_N$ be arbitrary. Then, $i \in A_N$, so that $\Delta\left(i\right) \in \Delta\left(A_N\right) \subset \left(A \otimes A\right)_N$ (since Δ is a graded map). Thus, $\pi_N\left(\Delta\left(i\right)\right) = \Delta\left(i\right)$ (by (13.24.3), applied to $t = \Delta\left(i\right)$). On the other hand, since $i \in I_N \subset I$, we have

$$\begin{aligned}
\Delta\left(i\right) \in \Delta\left(I\right) &\subset \underbrace{I}_{=\bigoplus_{n\geq 0} I_n} \otimes \underbrace{A}_{=\bigoplus_{m\geq 0} A_m} + \underbrace{A}_{=\bigoplus_{m\geq 0} A_m} \otimes \underbrace{I}_{=\bigoplus_{n\geq 0} I_n} \qquad \text{(since } I \text{ is a two-sided coideal)} \\
&= \underbrace{\left(\bigoplus_{n\geq 0} I_n\right) \otimes \left(\bigoplus_{m\geq 0} A_m\right)}_{=\sum\limits_{(m,n)\in\mathbb{N}^2} I_n \otimes A_m} + \underbrace{\left(\bigoplus_{m\geq 0} A_m\right) \otimes \left(\bigoplus_{n\geq 0} I_n\right)}_{=\sum\limits_{(m,n)\in\mathbb{N}^2} A_m \otimes I_n} = \sum_{(m,n)\in\mathbb{N}^2} I_n \otimes A_m + \sum_{(m,n)\in\mathbb{N}^2} A_m \otimes I_n .
\end{aligned}$$

Thus,

$$
\begin{aligned}
&\pi_N\left(\Delta\left(i\right)\right)\\
&\in \pi_N\left(\sum_{(m,n)\in\mathbb{N}^2} I_n\otimes A_m+\sum_{(m,n)\in\mathbb{N}^2} A_m\otimes I_n\right)\\
&=\underbrace{\sum_{(m,n)\in\mathbb{N}^2}\pi_N\left(I_n\otimes A_m\right)}_{=\sum_{\substack{(m,n)\in\mathbb{N}^2;\\ n+m=N}}\pi_N(I_n\otimes A_m)+\sum_{\substack{(m,n)\in\mathbb{N}^2;\\ n+m\neq N}}\pi_N(I_n\otimes A_m)}+\underbrace{\sum_{(m,n)\in\mathbb{N}^2}\pi_N\left(A_m\otimes I_n\right)}_{=\sum_{\substack{(m,n)\in\mathbb{N}^2;\\ m+n=N}}\pi_N(A_m\otimes I_n)+\sum_{\substack{(m,n)\in\mathbb{N}^2;\\ m+n\neq N}}\pi_N(A_m\otimes I_n)}\\
&=\sum_{\substack{(m,n)\in\mathbb{N}^2;\\ n+m=N}}\pi_N\left(I_n\otimes A_m\right)+\sum_{\substack{(m,n)\in\mathbb{N}^2;\\ n+m\neq N}}\pi_N\left(\underbrace{I_n}_{\subset A_n}\otimes A_m\right)\\
&\qquad+\sum_{\substack{(m,n)\in\mathbb{N}^2;\\ m+n=N}}\pi_N\left(A_m\otimes I_n\right)+\sum_{\substack{(m,n)\in\mathbb{N}^2;\\ m+n\neq N}}\pi_N\left(A_m\otimes\underbrace{I_n}_{\subset A_n}\right)\\
&\subset\sum_{\substack{(m,n)\in\mathbb{N}^2;\\ n+m=N}}\pi_N\left(I_n\otimes A_m\right)+\sum_{\substack{(m,n)\in\mathbb{N}^2;\\ n+m\neq N}}\underbrace{\pi_N\left(A_n\otimes A_m\right)}_{\substack{=0\\ \text{(by (13.24.5))}}}\\
&\qquad+\sum_{\substack{(m,n)\in\mathbb{N}^2;\\ m+n=N}}\pi_N\left(A_m\otimes I_n\right)+\sum_{\substack{(m,n)\in\mathbb{N}^2;\\ m+n\neq N}}\underbrace{\pi_N\left(A_m\otimes A_n\right)}_{\substack{=0\\ \text{(by (13.24.5), applied}\\ \text{to } (m,n) \text{ instead of } (n,m))}}\\
&=\underbrace{\sum_{\substack{(m,n)\in\mathbb{N}^2;\\ n+m=N}}\pi_N\left(I_n\otimes A_m\right)}_{=\sum_{\substack{(m,n)\in\mathbb{N}^2;\\ n+m=N;\\ n<N}}\pi_N(I_n\otimes A_m)+\sum_{\substack{(m,n)\in\mathbb{N}^2;\\ n+m=N;\\ n\geq N}}\pi_N(I_n\otimes A_m)}+\underbrace{\sum_{\substack{(m,n)\in\mathbb{N}^2;\\ m+n=N}}\pi_N\left(A_m\otimes I_n\right)}_{=\sum_{\substack{(m,n)\in\mathbb{N}^2;\\ m+n=N;\\ n<N}}\pi_N(A_m\otimes I_n)+\sum_{\substack{(m,n)\in\mathbb{N}^2;\\ m+n=N;\\ n\geq N}}\pi_N(A_m\otimes I_n)}\\
&=\sum_{\substack{(m,n)\in\mathbb{N}^2;\\ n+m=N;\\ n<N}}\pi_N\left(\underbrace{I_n}_{\substack{=0\\ \text{(by (13.24.2))}}}\otimes A_m\right)+\underbrace{\sum_{\substack{(m,n)\in\mathbb{N}^2;\\ n+m=N;\\ n\geq N}}\pi_N\left(I_n\otimes A_m\right)}_{\substack{=I_N\otimes A_0\\ \text{(since the only pair } (m,n)\in\mathbb{N}^2\\ \text{satisfying } n+m=N \text{ and } n\geq N \text{ is } (0,N))}}\\
&\qquad+\sum_{\substack{(m,n)\in\mathbb{N}^2;\\ m+n=N;\\ n<N}}\pi_N\left(A_m\otimes\underbrace{I_n}_{\substack{=0\\ \text{(by (13.24.2))}}}\right)+\underbrace{\sum_{\substack{(m,n)\in\mathbb{N}^2;\\ m+n=N;\\ n\geq N}}\pi_N\left(A_m\otimes I_n\right)}_{\substack{=A_0\otimes I_N\\ \text{(since the only pair } (m,n)\in\mathbb{N}^2\\ \text{satisfying } m+n=N \text{ and } n\geq N \text{ is } (0,N))}}\\
&=\sum_{\substack{(m,n)\in\mathbb{N}^2;\\ n+m=N;\\ n<N}}\underbrace{\pi_N\left(0\otimes A_m\right)}_{=0}+I_N\otimes\underbrace{A_0}_{\substack{=\mathbf{k}1_A\\ \text{(since } A \text{ is connected)}}}+\sum_{\substack{(m,n)\in\mathbb{N}^2;\\ m+n=N;\\ n<N}}\underbrace{\pi_N\left(A_m\otimes 0\right)}_{=0}+\underbrace{A_0}_{\substack{=\mathbf{k}1_A\\ \text{(since } A \text{ is connected)}}}\otimes I_N\\
&=I_N\otimes\mathbf{k}1_A+\mathbf{k}1_A\otimes I_N.
\end{aligned}
$$

Since $\pi_N(\Delta(i)) = \Delta(i)$, this rewrites as $\Delta(i) \in I_N \otimes \mathbf{k}1_A + \mathbf{k}1_A \otimes I_N$. In other words, there exists some $y \in I_N \otimes \mathbf{k}1_A$ and some $z \in \mathbf{k}1_A \otimes I_N$ such that $\Delta(i) = y + z$. Consider these y and z.

Since $y \in I_N \otimes \mathbf{k}1_A = I_N \otimes 1_A$, there exists a $j \in I_N$ such that $y = j \otimes 1_A$. Consider this j.

Since $z \in \mathbf{k}1_A \otimes I_N = 1_A \otimes I_N$, there exists a $k \in I_N$ such that $z = 1_A \otimes k$. Consider this k.

Now,

$$\Delta(i) = \underbrace{y}_{=j\otimes 1_A} + \underbrace{z}_{=1_A\otimes k} = j \otimes 1_A + 1_A \otimes k.$$

Applying the map $\epsilon \otimes \mathrm{id} : A \otimes A \to A$ to both sides of this equation, we obtain

$$(\epsilon \otimes \mathrm{id})(\Delta(i)) = (\epsilon \otimes \mathrm{id})(j \otimes 1_A + 1_A \otimes k) = \underbrace{\epsilon(j)}_{\substack{=0 \\ \text{(since } j\in I_N\text{, so that} \\ \epsilon(j)\in\epsilon(I_N)=0)}} \mathrm{id}(1_A) + \underbrace{\epsilon(1_A)}_{=1}\underbrace{\mathrm{id}(k)}_{=k} = \underbrace{0\,\mathrm{id}(1_A)}_{=0} + k = k.$$

Since $(\epsilon \otimes \mathrm{id})(\Delta(i)) = i$ (by the axioms of a coalgebra), this rewrites as $i = k$.

Similarly, applying $\mathrm{id} \otimes \epsilon$ to both sides of the equation $\Delta(i) = j \otimes 1_A + 1_A \otimes k$ and simplifying, we obtain $i = j$.

Now,

$$\Delta(i) = \underbrace{j}_{=i} \otimes 1_A + 1_A \otimes \underbrace{k}_{=i} = i \otimes 1_A + 1_A \otimes i.$$

Hence, i is primitive. In other words, $i \in \mathfrak{p}$. Combined with $i \in I_N = I \cap A_N \subset I$, this yields $i \in I \cap \mathfrak{p} = 0$, so that $i = 0$.

Now, forget that we fixed i. We thus have shown that every $i \in I_N$ satisfies $i = 0$. In other words, $I_N = 0$. In other words, (13.24.1) holds for $n = N$. This completes the induction proof of (13.24.1).

Now, $I = \bigoplus_{n\geq 0} \underbrace{I_n}_{\substack{=0 \\ \text{(by (13.24.1))}}} = \bigoplus_{n\geq 0} 0 = 0$. This solves part (a) of the exercise.

(b) By Exercise 1.3.13(a), we know that $\ker f$ is a two-sided coideal of A. It further satisfies $\ker f = \bigoplus_{n\geq 0} ((\ker f) \cap A_n)$ (since f is graded). If $f\mid_{\mathfrak{p}}$ is injective, then $(\ker f) \cap \mathfrak{p} = 0$, and thus part (a) of the current exercise (applied to $I = \ker f$) yields that $\ker f = 0$, so that f is injective. This solves part (b) of the exercise.

(c) The solution of part (c) proceeds precisely as the solution of part (b), except that instead of using Exercise 1.3.13(a) we now must use Exercise 1.3.13(b).

13.25. **Solution to Exercise 1.5.4.** *Solution to Exercise 1.5.4.*

(a) Let A be any associative $\mathbf{k}$-algebra. Define a $\mathbf{k}$-bilinear map $[\cdot,\cdot] : A \times A \to A$ by setting

$$[a, b] = ab - ba \qquad \text{for all } a, b \in A.$$

We must prove that this $\mathbf{k}$-bilinear map $[\cdot,\cdot]$ makes A into a Lie algebra. In order to do so, it is clearly enough to prove that

$$[x, x] = 0 \qquad \text{for all } x \in A, \tag{13.25.1}$$

and that

$$[x, [y, z]] + [z, [x, y]] + [y, [z, x]] = 0 \qquad \text{for all } x, y, z \in A. \tag{13.25.2}$$

Proof of (13.25.1)*:* Every $x \in A$ satisfies

$$\begin{aligned} [x, x] &= xx - xx \qquad \text{(by the definition of } [x, x]\text{)} \\ &= 0. \end{aligned}$$

This proves (13.25.1).

Proof of (13.25.2)*:* Every $x, y, z \in A$ satisfy

$$\begin{aligned}
&\underbrace{[x,[y,z]]}_{\substack{=x[y,z]-[y,z]x\\ \text{(by the definition}\\ \text{of }[x,[y,z]])}} + \underbrace{[z,[x,y]]}_{\substack{=z[x,y]-[x,y]z\\ \text{(by the definition}\\ \text{of }[z,[x,y]])}} + \underbrace{[y,[z,x]]}_{\substack{=y[z,x]-[z,x]y\\ \text{(by the definition}\\ \text{of }[y,[z,x]])}} \\
&= x \underbrace{[y,z]}_{\substack{=yz-zy\\ \text{(by the definition}\\ \text{of }[y,z])}} - \underbrace{[y,z]}_{\substack{=yz-zy\\ \text{(by the definition}\\ \text{of }[y,z])}} x + z \underbrace{[x,y]}_{\substack{=xy-yx\\ \text{(by the definition}\\ \text{of }[x,y])}} - \underbrace{[x,y]}_{\substack{=xy-yx\\ \text{(by the definition}\\ \text{of }[x,y])}} z + y \underbrace{[z,x]}_{\substack{=zx-xz\\ \text{(by the definition}\\ \text{of }[z,x])}} - \underbrace{[z,x]}_{\substack{=zx-xz\\ \text{(by the definition}\\ \text{of }[z,x])}} y \\
&= x\left(yz-zy\right) - \left(yz-zy\right)x + z\left(xy-yx\right) - \left(xy-yx\right)z + y\left(zx-xz\right) - \left(zx-xz\right)y \\
&= xyz - xzy - yzx + zyx + zxy - zyx - xyz + yxz + yzx - yxz - zxy + xzy \\
&= 0.
\end{aligned}$$

This proves (13.25.2).

Now, both (13.25.1) and (13.25.2) are proven. Hence, the $\mathbf{k}$-module A endowed with the $\mathbf{k}$-bilinear map $[\cdot,\cdot]$ satisfies the axioms of a Lie algebra, and therefore is a Lie algebra. This solves part (a) of the exercise.

(b) Let A be a bialgebra. Let $\mathfrak{p}$ be the set of all primitive elements of A. Then, $\mathfrak{p}$ is a $\mathbf{k}$-submodule of A (because it is easy to see that $0 \in \mathfrak{p}$, that $\lambda a \in \mathfrak{p}$ for every $\lambda \in \mathbf{k}$ and $a \in \mathfrak{p}$, and that $a + b \in \mathfrak{p}$ for all $a \in \mathfrak{p}$ and $b \in \mathfrak{p}$). A simple computation shows that every $x \in \mathfrak{p}$ and $y \in \mathfrak{p}$ satisfy $[x,y] \in \mathfrak{p}$ [436]. Hence, $[\mathfrak{p},\mathfrak{p}] \subset \mathfrak{p}$ (since $\mathfrak{p}$ is a $\mathbf{k}$-submodule of A). Therefore, $\mathfrak{p}$ is a Lie subalgebra of A. This solves part (b) of the exercise.

(c) For every subset S of a $\mathbf{k}$-module U, we let $\langle S \rangle$ denote the $\mathbf{k}$-submodule of U spanned by S.

We have defined J as the two-sided ideal of $T(\mathfrak{p})$ generated by all elements $xy - yx - [x,y]$ for x, y in $\mathfrak{p}$. In other words,

$$J = \underbrace{T(\mathfrak{p})}_{\supset \mathbf{k}} \cdot \left\langle xy - yx - [x,y] \ \mid \ x, y \in \mathfrak{p} \right\rangle \cdot \underbrace{T(\mathfrak{p})}_{\supset \mathbf{k}} \tag{13.25.3}$$

$$\begin{aligned}
&\supset \mathbf{k} \cdot \left\langle xy - yx - [x,y] \ \mid \ x, y \in \mathfrak{p} \right\rangle \cdot \mathbf{k} = \left\langle xy - yx - [x,y] \ \mid \ x, y \in \mathfrak{p} \right\rangle \qquad (13.25.4) \\
&\supset \left\{ xy - yx - [x,y] \ \mid \ x, y \in \mathfrak{p} \right\}.
\end{aligned}$$

Thus,

$$xy - yx - [x,y] \in J \qquad \text{for all } x, y \in \mathfrak{p}. \tag{13.25.5}$$

It is also easy to show that

$$xy - yx - [x,y] \text{ is a primitive element of } T(\mathfrak{p}) \text{ for all } x, y \in \mathfrak{p}. \tag{13.25.6}$$

[437]

From this it is easy to obtain

$$\Delta\left(\left\langle xy - yx - [x,y] \ \mid \ x, y \in \mathfrak{p} \right\rangle\right) \subset J \otimes T(\mathfrak{p}) + T(\mathfrak{p}) \otimes J. \tag{13.25.7}$$

[436] *Proof.* Let $x \in \mathfrak{p}$ and $y \in \mathfrak{p}$. Then, $\Delta[x,y] = 1 \otimes [x,y] + [x,y] \otimes 1$ (by (1.3.7)). In other words, the element $[x,y]$ of A is primitive. In other words, $[x,y] \in \mathfrak{p}$ (since $\mathfrak{p}$ is the set of all primitive elements of A), qed.

[437] *Proof.* Let $x, y \in \mathfrak{p}$. Then, $\Delta(xy - yx) = 1 \otimes (xy - yx) + (xy - yx) \otimes 1$ (this can be proven just as in the proof of (1.3.7)). But $[x,y]$ is an element of $\mathfrak{p}$, and thus (by the definition of the comultiplication of $T(\mathfrak{p})$) satisfies $\Delta[x,y] = 1 \otimes [x,y] + [x,y] \otimes 1$ in $T(\mathfrak{p}) \otimes T(\mathfrak{p})$. Now, since Δ is a $\mathbf{k}$-linear map, we have

$$\begin{aligned}
\Delta(xy - yx - [x,y]) &= \underbrace{\Delta(xy - yx)}_{=1\otimes(xy-yx)+(xy-yx)\otimes 1} - \underbrace{\Delta[x,y]}_{=1\otimes[x,y]+[x,y]\otimes 1} \\
&= \left(1 \otimes (xy - yx) + (xy - yx) \otimes 1\right) - \left(1 \otimes [x,y] + [x,y] \otimes 1\right) \\
&= \left(\underbrace{(xy - yx) \otimes 1 - [x,y] \otimes 1}_{=(xy-yx-[x,y])\otimes 1}\right) + \left(\underbrace{1 \otimes (xy - yx) - 1 \otimes [x,y]}_{=1\otimes(xy-yx-[x,y])}\right) \\
&= (xy - yx - [x,y]) \otimes 1 + 1 \otimes (xy - yx - [x,y]) \\
&= 1 \otimes (xy - yx - [x,y]) + (xy - yx - [x,y]) \otimes 1.
\end{aligned}$$

In other words, the element $xy - yx - [x,y]$ of $T(\mathfrak{p})$ is primitive. This proves (13.25.6).

[438] But applying Δ to both sides of the equality (13.25.3), we obtain

$$\begin{align}
\Delta\left(J\right) &= \Delta\left(T\left(\mathfrak{p}\right)\cdot\left\langle xy-yx-\left[x,y\right]\ \mid\ x,y\in\mathfrak{p}\right\rangle\cdot T\left(\mathfrak{p}\right)\right) \nonumber\\
&\subset \underbrace{\Delta\left(T\left(\mathfrak{p}\right)\right)}_{\subset T(\mathfrak{p})\otimes T(\mathfrak{p})}\cdot\underbrace{\Delta\left(\left\langle xy-yx-\left[x,y\right]\ \mid\ x,y\in\mathfrak{p}\right\rangle\right)}_{\substack{\subset J\otimes T(\mathfrak{p})+T(\mathfrak{p})\otimes J\\ \text{(by (13.25.7))}}}\cdot\underbrace{\Delta\left(T\left(\mathfrak{p}\right)\right)}_{\subset T(\mathfrak{p})\otimes T(\mathfrak{p})} \nonumber\\
&\qquad\left(\text{since }\Delta\text{ is a }\mathbf{k}\text{-algebra homomorphism}\right) \nonumber\\
&\subset \left(T\left(\mathfrak{p}\right)\otimes T\left(\mathfrak{p}\right)\right)\cdot\left(J\otimes T\left(\mathfrak{p}\right)+T\left(\mathfrak{p}\right)\otimes J\right)\cdot\left(T\left(\mathfrak{p}\right)\otimes T\left(\mathfrak{p}\right)\right) \nonumber\\
&= \underbrace{\left(T\left(\mathfrak{p}\right)\otimes T\left(\mathfrak{p}\right)\right)\cdot\left(J\otimes T\left(\mathfrak{p}\right)\right)\cdot\left(T\left(\mathfrak{p}\right)\otimes T\left(\mathfrak{p}\right)\right)}_{=(T(\mathfrak{p})\cdot J\cdot T(\mathfrak{p}))\otimes(T(\mathfrak{p})\cdot T(\mathfrak{p})\cdot T(\mathfrak{p}))}+\underbrace{\left(T\left(\mathfrak{p}\right)\otimes T\left(\mathfrak{p}\right)\right)\cdot\left(T\left(\mathfrak{p}\right)\otimes J\right)\cdot\left(T\left(\mathfrak{p}\right)\otimes T\left(\mathfrak{p}\right)\right)}_{=(T(\mathfrak{p})\cdot T(\mathfrak{p})\cdot T(\mathfrak{p}))\otimes(T(\mathfrak{p})\cdot J\cdot T(\mathfrak{p}))} \nonumber\\
&= \underbrace{\left(T\left(\mathfrak{p}\right)\cdot J\cdot T\left(\mathfrak{p}\right)\right)}_{\substack{=J\\ \text{(since J is a two-sided ideal of $T(\mathfrak{p})$)}}}\otimes\underbrace{\left(T\left(\mathfrak{p}\right)\cdot T\left(\mathfrak{p}\right)\cdot T\left(\mathfrak{p}\right)\right)}_{=T(\mathfrak{p})} \nonumber\\
&\qquad+\underbrace{\left(T\left(\mathfrak{p}\right)\cdot T\left(\mathfrak{p}\right)\cdot T\left(\mathfrak{p}\right)\right)}_{=T(\mathfrak{p})}\otimes\underbrace{\left(T\left(\mathfrak{p}\right)\cdot J\cdot T\left(\mathfrak{p}\right)\right)}_{\substack{=J\\ \text{(since J is a two-sided ideal of $T(\mathfrak{p})$)}}} \nonumber\\
&= J\otimes T\left(\mathfrak{p}\right)+T\left(\mathfrak{p}\right)\otimes J. \tag{13.25.8}
\end{align}$$

Also,

$$\epsilon\left(\left\langle xy-yx-\left[x,y\right]\ \mid\ x,y\in\mathfrak{p}\right\rangle\right)=0. \tag{13.25.9}$$

[439] Applying the map ϵ to both sides of the equality (13.25.3), we obtain

$$\begin{aligned}
\epsilon\left(J\right) &= \epsilon\left(T\left(\mathfrak{p}\right)\cdot\left\langle xy-yx-\left[x,y\right]\ \mid\ x,y\in\mathfrak{p}\right\rangle\cdot T\left(\mathfrak{p}\right)\right)\subset\epsilon\left(T\left(\mathfrak{p}\right)\right)\cdot\underbrace{\epsilon\left(\left\langle xy-yx-\left[x,y\right]\ \mid\ x,y\in\mathfrak{p}\right\rangle\right)}_{\substack{=0\\ \text{(by (13.25.9))}}}\cdot\epsilon\left(T\left(\mathfrak{p}\right)\right)\\
&\qquad\left(\text{since }\epsilon\text{ is a }\mathbf{k}\text{-algebra homomorphism}\right)\\
&= 0.
\end{aligned}$$

Thus, $\epsilon\left(J\right)=0$. Combined with (13.25.8), this yields that J is a two-sided coideal of $T\left(\mathfrak{p}\right)$. Thus, the quotient $T\left(\mathfrak{p}\right)/J$ becomes a $\mathbf{k}$-coalgebra. Also, $T\left(\mathfrak{p}\right)/J$ is a $\mathbf{k}$-algebra, since J is a two-sided ideal of $T\left(\mathfrak{p}\right)$.

Now we want to show that $T\left(\mathfrak{p}\right)/J$ has a structure of a cocommutative $\mathbf{k}$-bialgebra inherited from $T\left(\mathfrak{p}\right)$. We already know that $T\left(\mathfrak{p}\right)/J$ is a $\mathbf{k}$-algebra and a $\mathbf{k}$-coalgebra, with both structures being inherited from $T\left(\mathfrak{p}\right)$. The remaining axioms of a cocommutative $\mathbf{k}$-bialgebra that need to be checked for $T\left(\mathfrak{p}\right)/J$ are the commutativity of the diagrams (1.3.4) and the commutativity of the diagram (1.5.2); these axioms are clearly preserved under taking quotients. Hence, $T\left(\mathfrak{p}\right)/J$ is a cocommutative $\mathbf{k}$-bialgebra, with its structure being inherited from $T\left(\mathfrak{p}\right)$. In other words, $\mathcal{U}\left(\mathfrak{p}\right)$ is a cocommutative $\mathbf{k}$-bialgebra, with its structure being inherited from $T\left(\mathfrak{p}\right)$ (since $\mathcal{U}\left(\mathfrak{p}\right)=T\left(\mathfrak{p}\right)/J$). This solves part (c) of the exercise.

(d) Proposition 1.4.10 shows that the antipode S of $T\left(\mathfrak{p}\right)$ is a $\mathbf{k}$-algebra anti-homomorphism from $T\left(\mathfrak{p}\right)$ to $T\left(\mathfrak{p}\right)$.

[438] *Proof.* Let $x,y\in\mathfrak{p}$ be arbitrary. Then, $xy-yx-\left[x,y\right]$ is a primitive element of $T\left(\mathfrak{p}\right)$ (by (13.25.6)). In other words,

$$\begin{aligned}
\Delta\left(xy-yx-\left[x,y\right]\right) &= 1\otimes\left(xy-yx-\left[x,y\right]\right)+\left(xy-yx-\left[x,y\right]\right)\otimes1\\
&= \underbrace{\left(xy-yx-\left[x,y\right]\right)}_{\substack{\in J\\ \text{(by (13.25.5))}}}\otimes\underbrace{1}_{\in T(\mathfrak{p})}+\underbrace{1}_{\in T(\mathfrak{p})}\otimes\underbrace{\left(xy-yx-\left[x,y\right]\right)}_{\substack{\in J\\ \text{(by (13.25.5))}}}\\
&\in J\otimes T\left(\mathfrak{p}\right)+T\left(\mathfrak{p}\right)\otimes J.
\end{aligned}$$

Now forget that we fixed x,y. We thus have proven that $\Delta\left(xy-yx-\left[x,y\right]\right)\in J\otimes T\left(\mathfrak{p}\right)+T\left(\mathfrak{p}\right)\otimes J$ for all $x,y\in\mathfrak{p}$. Since $J\otimes T\left(\mathfrak{p}\right)+T\left(\mathfrak{p}\right)\otimes J$ is a $\mathbf{k}$-submodule of $T\left(\mathfrak{p}\right)\otimes T\left(\mathfrak{p}\right)$, this yields that $\Delta\left(\left\langle xy-yx-\left[x,y\right]\ \mid\ x,y\in\mathfrak{p}\right\rangle\right)\subset J\otimes T\left(\mathfrak{p}\right)+T\left(\mathfrak{p}\right)\otimes J$. Thus, (13.25.7) is proven.

[439] *Proof.* Let $x,y\in\mathfrak{p}$. Then, $xy-yx\in\mathfrak{p}^{\otimes2}$ and thus $\epsilon\left(xy-yx\right)=0$ (by the definition of the comultiplication ϵ on $T\left(\mathfrak{p}\right)$). Also, $\left[x,y\right]\in\mathfrak{p}^{\otimes1}$ and thus $\epsilon\left(\left[x,y\right]\right)=0$ (again by the definition of the comultiplication ϵ on $T\left(\mathfrak{p}\right)$). Since ϵ is $\mathbf{k}$-linear, we have

$$\epsilon\left(xy-yx-\left[x,y\right]\right)=\underbrace{\epsilon\left(xy-yx\right)}_{=0}-\underbrace{\epsilon\left(\left[x,y\right]\right)}_{=0}=0-0=0.$$

Now, forget that we fixed x,y. We thus have shown that $\epsilon\left(xy-yx-\left[x,y\right]\right)=0$ for all $x,y\in\mathfrak{p}$. By linearity, this yields $\epsilon\left(\left\langle xy-yx-\left[x,y\right]\ \mid\ x,y\in\mathfrak{p}\right\rangle\right)=0$, so that (13.25.9) is proven.

It is easy to see that

$$S\left(\langle xy-yx-[x,y]\ \mid\ x,y\in\mathfrak{p}\rangle\right)\subset J. \tag{13.25.10}$$

[440] But recall that if A and B are two $\mathbf{k}$-algebras and $\varphi : A\to B$ is a $\mathbf{k}$-algebra homomorphism, then any $\mathbf{k}$-submodules V_1, V_2, ..., V_n of A satisfy

$$\varphi\left(V_1V_2...V_n\right)=\varphi\left(V_1\right)\cdot\varphi\left(V_2\right)\cdot...\cdot\varphi\left(V_n\right).$$

Similarly, if A and B are two $\mathbf{k}$-algebras and $\varphi : A\to B$ is a $\mathbf{k}$-algebra anti-homomorphism, then any $\mathbf{k}$-submodules V_1, V_2, ..., V_n of A satisfy

$$\varphi\left(V_1V_2...V_n\right)=\varphi\left(V_n\right)\cdot\varphi\left(V_{n-1}\right)\cdot...\cdot\varphi\left(V_1\right).$$

Applying this to $A=T(\mathfrak{p})$, $B=T(\mathfrak{p})$, $\varphi=S$, $n=3$, $V_1=T(\mathfrak{p})$, $V_2=\langle xy-yx-[x,y]\ \mid\ x,y\in\mathfrak{p}\rangle$ and $V_3=T(\mathfrak{p})$, we obtain

$$\begin{aligned}S\left(T(\mathfrak{p})\cdot\langle xy-yx-[x,y]\ \mid\ x,y\in\mathfrak{p}\rangle\cdot T(\mathfrak{p})\right)&=\underbrace{S\left(T(\mathfrak{p})\right)}_{\subset T(\mathfrak{p})}\cdot\underbrace{S\left(\langle xy-yx-[x,y]\ \mid\ x,y\in\mathfrak{p}\rangle\right)}_{\subset J}\cdot\underbrace{S\left(T(\mathfrak{p})\right)}_{\subset T(\mathfrak{p})}\\&\subset T(\mathfrak{p})\cdot J\cdot T(\mathfrak{p})\subset J\end{aligned}$$

(since J is a two-sided ideal). Due to (13.25.3), this rewrites as $S(J)\subset J$. Hence, the $\mathbf{k}$-linear map $S:T(\mathfrak{p})\to T(\mathfrak{p})$ induces a $\mathbf{k}$-linear map $\overline{S}:T(\mathfrak{p})/J\to T(\mathfrak{p})/J$ on the quotient $\mathbf{k}$-modules. This resulting map $\overline{S}:T(\mathfrak{p})/J\to T(\mathfrak{p})/J$ is an antipode for the $\mathbf{k}$-bialgebra $T(\mathfrak{p})/J$ (in fact, the diagram (1.4.3) with A and S replaced by $T(\mathfrak{p})/J$ and $\overline{S}$ commutes, because the diagram (1.4.3) with A replaced by $T(\mathfrak{p})$ commutes). Hence, the $\mathbf{k}$-bialgebra $T(\mathfrak{p})/J$ has an antipode, and thus is a Hopf algebra. In other words, $\mathcal{U}(\mathfrak{p})$ is a Hopf algebra (since $\mathcal{U}(\mathfrak{p})=T(\mathfrak{p})/J$). Since we already know that $\mathcal{U}(\mathfrak{p})$ is cocommutative, this yields that $\mathcal{U}(\mathfrak{p})$ is a cocommutative Hopf algebra. This solves part (d) of the exercise.

13.26. **Solution to Exercise 1.5.5.** *Solution to Exercise 1.5.5.* Let $f,g\in\operatorname{Hom}(C,A)$. We must show that $f\star g=g\star f$.

There are several ways to do this. The slickest one is perhaps the following:

The $\mathbf{k}$-algebra A is commutative. In other words, the diagram (1.5.1) commutes. In other words, $m_A=m_A\circ T_A$, where T_A denotes the twist map $A\otimes A\to A\otimes A,\ b\otimes a\mapsto a\otimes b$.

The $\mathbf{k}$-coalgebra C is cocommutative. In other words, the diagram (1.5.2) commutes. In other words, $\Delta_C=T_C\circ\Delta_C$, where T_C denotes the twist map $C\otimes C\to C\otimes C,\ c\otimes d\mapsto d\otimes c$.

It is straightforward to see that

$$T_A\circ(f\otimes g)=(g\otimes f)\circ T_C. \tag{13.26.1}$$

[441]

[440] *Proof.* Let $x,y\in\mathfrak{p}$. Then, $xy-yx-[x,y]$ is a primitive element of $T(\mathfrak{p})$ (by (13.25.6)). Hence, Proposition 1.4.17 (applied to $xy-yx-[x,y]$ instead of x) yields

$$S\left(xy-yx-[x,y]\right)=-\underbrace{\left(xy-yx-[x,y]\right)}_{\substack{\in J\\ \text{(by (13.25.5))}}}\in -J=J$$

(since J is a two-sided ideal).

Now, forget that we fixed x,y. We thus have shown that $S\left(xy-yx-[x,y]\right)\in J$ for all $x,y\in\mathfrak{p}$. Since J is a $\mathbf{k}$-module, this yields that $S\left(\langle xy-yx-[x,y]\ \mid\ x,y\in\mathfrak{p}\rangle\right)\subset J$. Thus, (13.25.10) is proven.

[441] *Proof of (13.26.1):* Let $z\in C\otimes C$. We shall prove the equality $\left(T_A\circ(f\otimes g)\right)(z)=\left((g\otimes f)\circ T_C\right)(z)$.

Since this equality is $\mathbf{k}$-linear in z, we can WLOG assume that z is a pure tensor (since each tensor $C\otimes C$ is a $\mathbf{k}$-linear combination of pure tensors). Assume this. Hence, $z=c\otimes d$ for some $c\in C$ and $d\in C$. Consider these c and d.

Now,

$$\begin{aligned}\left(T_A\circ(f\otimes g)\right)\left(\underbrace{z}_{=c\otimes d}\right)&=\left(T_A\circ(f\otimes g)\right)(c\otimes d)=T_A\left(\underbrace{(f\otimes g)(c\otimes d)}_{=f(c)\otimes g(d)}\right)=T_A\left(f(c)\otimes g(d)\right)\\&=g(d)\otimes f(c)\qquad\text{(by the definition of the map }T_A).\end{aligned}$$

Now, the definition of convolution yields $f \star g = m_A \circ (f \otimes g) \circ \Delta_C$ and $g \star f = m_A \circ (g \otimes f) \circ \Delta_C$. Hence,

$$\begin{aligned} f \star g &= \underbrace{m_A}_{=m_A \circ T_A} \circ (f \otimes g) \circ \Delta_C = m_A \circ \underbrace{T_A \circ (f \otimes g)}_{=(g \otimes f) \circ T_C} \circ \Delta_C \\ &= m_A \circ (g \otimes f) \circ \underbrace{T_C \circ \Delta_C}_{=\Delta_C} = m_A \circ (g \otimes f) \circ \Delta_C = g \star f. \end{aligned}$$

This solves Exercise 1.5.5.

13.27. **Solution to Exercise 1.5.6.** *Solution to Exercise 1.5.6.* Recall that our abstract definition of a **k**-algebra (Definition 1.1.1) and our definition of a **k**-coalgebra (Definition 1.2.1) differ from each other only in the directions of the arrows. More precisely, reversing all arrows in the former definition yields the latter definition. Similarly, our definition of a cocommutative **k**-coalgebra is obtained by reversing all arrows in our abstract definition of a commutative **k**-algebra, and our definition of a **k**-coalgebra homomorphism is obtained by reversing all arrows in our abstract definition of a **k**-algebra homomorphism. Hence, the statements of parts (a) and (b) of this exercise can be obtained from each other by reversing all arrows. Therefore, if we can solve part (b) of this exercise in an element-free way[442], then we can clearly apply the same argument "with all arrows reversed" (and, of course, with A, m_A and u_A replaced by C, Δ_C and ϵ_C) to solve part (a). Hence, in order to solve this exercise, it is enough to find an element-free solution to its part (b). Let us do this now.

Let us first show how to solve part (b) using computations with elements (i.e., not in an element-free way). This is very easy. We need to prove the following statements:

> *Statement 1:* If a **k**-algebra A is commutative, then its multiplication $m_A : A \otimes A \to A$ is a **k**-algebra homomorphism.
>
> *Statement 2:* If a **k**-algebra A has the property that its multiplication $m_A : A \otimes A \to A$ is a **k**-algebra homomorphism, then A is commutative.

Proof of Statement 1 using computations with elements: Assume that a **k**-algebra A is commutative. We need to prove that $m_A : A \otimes A \to A$ is a **k**-algebra homomorphism. To do so, it is enough to show that m_A preserves products and that m_A maps the unity of $A \otimes A$ to the unity of A.

Let us show that m_A preserves products. This means proving that $m_A(uv) = m_A(u) \cdot m_A(v)$ for all $u \in A \otimes A$ and $v \in A \otimes A$. So let $u \in A \otimes A$ and $v \in A \otimes A$. Since the equality in question ($m_A(uv) = m_A(u) \cdot m_A(v)$) is linear in each of u and v, we can WLOG assume that u and v are pure tensors. Having made this assumption, we can write $u = a \otimes b$ for some $a \in A$ and $b \in A$, and we can write $v = c \otimes d$ for some $c \in A$ and $d \in A$. Now,

$$m_A\left(\underbrace{u}_{=a\otimes b}\underbrace{v}_{=c\otimes d}\right) = m_A\left(\underbrace{(a \otimes b)(c \otimes d)}_{=ac\otimes bd}\right) = m_A(ac \otimes bd) = (ac)(bd)$$

Comparing this with

$$((g \otimes f) \circ T_C)\left(\underbrace{z}_{=c\otimes d}\right) = ((g \otimes f) \circ T_C)(c \otimes d) = (g \otimes f)\left(\underbrace{T_C(c \otimes d)}_{\substack{=d\otimes c \\ \text{(by the definition of the map } T_C)}}\right) = (g \otimes f)(d \otimes c) = g(d) \otimes f(c),$$

we obtain $(T_A \circ (f \otimes g))(z) = ((g \otimes f) \circ T_C)(z)$.

Now, forget that we fixed z. We thus have shown that $(T_A \circ (f \otimes g))(z) = ((g \otimes f) \circ T_C)(z)$ for each $z \in C \otimes C$. In other words, $T_A \circ (f \otimes g) = (g \otimes f) \circ T_C$. This proves (13.26.1).

[442] By an "element-free" argument, we mean an argument which only talks about linear maps, but never talks about elements of modules such as A and $A \otimes A$. For instance, the Second solution of Exercise 1.4.4(a) that we gave above was element-free, whereas the First solution of Exercise 1.4.4(a) (which we also showed above) was not element-free (since it involved elements c and d of C and D). A synonym for "element-free argument" is "argument by pure diagram chasing", although it is not required that one actually draws any diagrams in the argument (commutative diagrams are just shortcuts for identities between maps).

and

$$m_A\left(\underbrace{u}_{=a\otimes b}\right)\cdot m_A\left(\underbrace{v}_{=c\otimes d}\right)=\underbrace{m_A\left(a\otimes b\right)}_{=ab}\cdot\underbrace{m_A\left(c\otimes d\right)}_{=cd}=\left(ab\right)\left(cd\right).$$

But[443]

$$(ac)\,(bd)=((ac)\,b)\,d=\left(a\underbrace{(cb)}_{\substack{=bc\\ \text{(since } A \text{ is}\\ \text{commutative)}}}\right)d=(a\,(bc))\,d=((ab)\,c)\,d=(ab)\,(cd)\,. \qquad (13.27.1)$$

Hence, altogether,

$$m_A\left(uv\right)=\left(ac\right)\left(bd\right)=\left(ab\right)\left(cd\right)=m_A\left(u\right)\cdot m_A\left(v\right).$$

Thus, we have proven that m_A preserves products. In order to prove that m_A maps the unity of $A\otimes A$ to the unity of A, we recall that the former unity is $1_A\otimes 1_A$ and the latter unity is 1_A, which satisfy $m_A\left(1_A\otimes 1_A\right)=1_A\cdot 1_A=1_A$. We have thus shown that m_A preserves products and that m_A maps the unity of $A\otimes A$ to the unity of A. In other words, m_A is a $\mathbf{k}$-algebra homomorphism, and Statement 1 is proven.

Proof of Statement 2 using computations with elements: Assume that a $\mathbf{k}$-algebra A has the property that its multiplication $m_A : A\otimes A\to A$ is a $\mathbf{k}$-algebra homomorphism. Let a and b be elements of A. Then,

$$m_A\left(\underbrace{\left(1_A\otimes b\right)\left(a\otimes 1_A\right)}_{=(1_Aa)\otimes(b1_A)=a\otimes b}\right)=m_A\left(a\otimes b\right)=ab.$$

Comparing this with

$$\begin{aligned} m_A\left(\left(1_A\otimes b\right)\left(a\otimes 1_A\right)\right)&=\underbrace{m_A\left(1_A\otimes b\right)}_{=1_Ab=b}\underbrace{m_A\left(a\otimes 1_A\right)}_{=a1_A=a} \qquad \text{(since } m_A \text{ is a } \mathbf{k}\text{-algebra homomorphism)}\\ &=ba,\end{aligned}$$

this becomes $ab=ba$. We thus have shown that $ab=ba$ for all $a\in A$ and $b\in A$. In other words, A is commutative, and thus Statement 2 is proven.

We thus have solved part (b) of the exercise using computations with elements. But we want an element-free solution of part (b). It turns out that we can obtain such a solution from our above solution by a more or less straightforward rewriting procedure. Let us show how this works.

Again, we need to prove Statements 1 and 2 made above.

Element-free proof of Statement 1: Assume that a $\mathbf{k}$-algebra A is commutative. We need to prove that $m_A : A\otimes A\to A$ is a $\mathbf{k}$-algebra homomorphism. To do so, it is enough to show that the diagrams

$$\begin{tikzcd} A\otimes A \arrow[r, "m_A"] & A \\ A\otimes A\otimes A\otimes A \arrow[u, "m_{A\otimes A}"] \arrow[r, "m_A\otimes m_A"] & A\otimes A \arrow[u, "m_A"'] \end{tikzcd}$$

and

$$\begin{tikzcd} A\otimes A \arrow[rr, "m_A"] & & A \\ & \mathbf{k} \arrow[ul, "u_{A\otimes A}"] \arrow[ur, "u_A"'] & \end{tikzcd}$$

are commutative. In other words, it is enough to show that

$$m_A\circ m_{A\otimes A}=m_A\circ\left(m_A\otimes m_A\right) \qquad (13.27.2)$$

and

$$u_A=m_A\circ u_{A\otimes A}. \qquad (13.27.3)$$

Let us prove (13.27.2) first. If we were allowed to compute with elements, then we could prove (13.27.2) by evaluating both sides of (13.27.2) at a pure tensor $a\otimes b\otimes c\otimes d\in A\otimes A\otimes A\otimes A$; this would leave us with the task of showing that $(ac)\,(bd)=(ab)\,(cd)$, which we already have done in the computation which proved

[443] In the following computation, we are deliberately being painstakingly slow and writing down every single step, including every application of associativity. This is to simplify our job later on (when we will translate this computation to an element-free argument).

(13.27.1). However, we are not allowed to do this, because we want this proof to be element-free. But what we *can* do is computing with maps instead of elements. We just need to replace the computation which proved (13.27.1) by a computation which uses maps instead of elements. If we replace every expression in (13.27.1) by the **k**-linear map which sends every $a \otimes b \otimes c \otimes d \in A \otimes A \otimes A \otimes A$ to said expression, then we obtain:

$$\begin{aligned}
&(\text{the } \mathbf{k}\text{-linear map sending every } a \otimes b \otimes c \otimes d \in A \otimes A \otimes A \otimes A \text{ to } (ac)(bd)) \\
&= (\text{the } \mathbf{k}\text{-linear map sending every } a \otimes b \otimes c \otimes d \in A \otimes A \otimes A \otimes A \text{ to } ((ac)b)d) \\
&= (\text{the } \mathbf{k}\text{-linear map sending every } a \otimes b \otimes c \otimes d \in A \otimes A \otimes A \otimes A \text{ to } (a(cb))d) \\
&= (\text{the } \mathbf{k}\text{-linear map sending every } a \otimes b \otimes c \otimes d \in A \otimes A \otimes A \otimes A \text{ to } (a(bc))d) \\
&= (\text{the } \mathbf{k}\text{-linear map sending every } a \otimes b \otimes c \otimes d \in A \otimes A \otimes A \otimes A \text{ to } ((ab)c)d) \\
&= (\text{the } \mathbf{k}\text{-linear map sending every } a \otimes b \otimes c \otimes d \in A \otimes A \otimes A \otimes A \text{ to } (ab)(cd)).
\end{aligned} \tag{13.27.4}$$

We are not yet done, because we still are using elements (in describing the maps). So we should rewrite the maps appearing in the computation (13.27.4) in such a way that no elements occur in them anymore. Denoting by T the twist map $A \otimes A \to A \otimes A$ (sending every $a \otimes b$ to $b \otimes a$) [444], we have

$$\begin{aligned}
&(\text{the } \mathbf{k}\text{-linear map sending every } a \otimes b \otimes c \otimes d \in A \otimes A \otimes A \otimes A \text{ to } (ac)(bd)) \\
&= m_A \circ (m_A \otimes m_A) \circ (\mathrm{id}_A \otimes T \otimes \mathrm{id}_A);
\end{aligned}$$

$$\begin{aligned}
&(\text{the } \mathbf{k}\text{-linear map sending every } a \otimes b \otimes c \otimes d \in A \otimes A \otimes A \otimes A \text{ to } ((ac)b)d) \\
&= m_A \circ (m_A \otimes \mathrm{id}_A) \circ (m_A \otimes \mathrm{id}_A \otimes \mathrm{id}_A) \circ (\mathrm{id}_A \otimes T \otimes \mathrm{id}_A);
\end{aligned}$$

$$\begin{aligned}
&(\text{the } \mathbf{k}\text{-linear map sending every } a \otimes b \otimes c \otimes d \in A \otimes A \otimes A \otimes A \text{ to } (a(cb))d) \\
&= m_A \circ (m_A \otimes \mathrm{id}_A) \circ (\mathrm{id}_A \otimes m_A \otimes \mathrm{id}_A) \circ (\mathrm{id}_A \otimes T \otimes \mathrm{id}_A);
\end{aligned}$$

$$\begin{aligned}
&(\text{the } \mathbf{k}\text{-linear map sending every } a \otimes b \otimes c \otimes d \in A \otimes A \otimes A \otimes A \text{ to } (a(bc))d) \\
&= m_A \circ (m_A \otimes \mathrm{id}_A) \circ (\mathrm{id}_A \otimes m_A \otimes \mathrm{id}_A);
\end{aligned}$$

$$\begin{aligned}
&(\text{the } \mathbf{k}\text{-linear map sending every } a \otimes b \otimes c \otimes d \in A \otimes A \otimes A \otimes A \text{ to } ((ab)c)d) \\
&= m_A \circ (m_A \otimes \mathrm{id}_A) \circ (m_A \otimes \mathrm{id}_A \otimes \mathrm{id}_A);
\end{aligned}$$

$$\begin{aligned}
&(\text{the } \mathbf{k}\text{-linear map sending every } a \otimes b \otimes c \otimes d \in A \otimes A \otimes A \otimes A \text{ to } (ab)(cd)) \\
&= m_A \circ (m_A \otimes m_A).
\end{aligned}$$

Hence, the computation (13.27.4) rewrites as

$$\begin{aligned}
&m_A \circ (m_A \otimes m_A) \circ (\mathrm{id}_A \otimes T \otimes \mathrm{id}_A) \\
&= m_A \circ (m_A \otimes \mathrm{id}_A) \circ (m_A \otimes \mathrm{id}_A \otimes \mathrm{id}_A) \circ (\mathrm{id}_A \otimes T \otimes \mathrm{id}_A) \\
&= m_A \circ (m_A \otimes \mathrm{id}_A) \circ (\mathrm{id}_A \otimes m_A \otimes \mathrm{id}_A) \circ (\mathrm{id}_A \otimes T \otimes \mathrm{id}_A) \\
&= m_A \circ (m_A \otimes \mathrm{id}_A) \circ (\mathrm{id}_A \otimes m_A \otimes \mathrm{id}_A) \\
&= m_A \circ (m_A \otimes \mathrm{id}_A) \circ (m_A \otimes \mathrm{id}_A \otimes \mathrm{id}_A) \\
&= m_A \circ (m_A \otimes m_A).
\end{aligned} \tag{13.27.5}$$

However, now that we are no longer talking about elements, our computation has become significantly harder to follow. For example, the first equality in the computation (13.27.5) is far less obvious than the corresponding equality $(ac)(bd) = ((ac)b)d$ in the computation (13.27.1). So, in order to justify the computation (13.27.5), we need to recall where exactly we used associativity or commutativity in (13.27.4), and translate these uses into element-free language. Let us do this step by step:

[444] We do not count the use of this map T as a use of elements (even though we just defined it using elements). Twist maps like T are one of the basic features of tensor products (along with associativity isomorphisms $(U \otimes V) \otimes W \to U \otimes (V \otimes W)$, which we are suppressing, and with trivial isomorphisms of the form $\mathbf{k} \otimes U \to U$), and their use is allowed in element-free arguments. They don't interfere with "reversing the arrows" because arrows like T are very easy to reverse.

The first equality in (13.27.5)*:* The first equality in (13.27.5) is simply an element-free way to state $(ac)(bd) = ((ac)b)d$ for all $a, b, c, d \in A$. On the level of elements, this follows from applying associativity to the elements ac, b and d of A. In other words, this follows from applying the associativity law $m_A \circ (\operatorname{id}_A \otimes m_A) = m_A \circ (m_A \otimes \operatorname{id}_A)$ to the tensor $ac \otimes b \otimes d = ((m_A \otimes \operatorname{id}_A \otimes \operatorname{id}_A) \circ (\operatorname{id}_A \otimes T \otimes \operatorname{id}_A))(a \otimes b \otimes c \otimes d)$. Therefore, the first equality in (13.27.5) should follow from $m_A \circ (\operatorname{id}_A \otimes m_A) = m_A \circ (m_A \otimes \operatorname{id}_A)$ by composition with the map $(m_A \otimes \operatorname{id}_A \otimes \operatorname{id}_A) \circ (\operatorname{id}_A \otimes T \otimes \operatorname{id}_A)$ on the right. And indeed, this is how it is proven:

$$\begin{aligned} & m_A \circ \underbrace{(m_A \otimes m_A)}_{\substack{=(\operatorname{id}_A \otimes m_A)\circ(m_A \otimes \operatorname{id}_{A\otimes A}) \\ =(\operatorname{id}_A \otimes m_A)\circ(m_A \otimes \operatorname{id}_A \otimes \operatorname{id}_A)}} \circ (\operatorname{id}_A \otimes T \otimes \operatorname{id}_A) \\ &= \underbrace{m_A \circ (\operatorname{id}_A \otimes m_A)}_{=m_A \circ (m_A \otimes \operatorname{id}_A)} \circ (m_A \otimes \operatorname{id}_A \otimes \operatorname{id}_A) \circ (\operatorname{id}_A \otimes T \otimes \operatorname{id}_A) \\ &= m_A \circ (m_A \otimes \operatorname{id}_A) \circ (m_A \otimes \operatorname{id}_A \otimes \operatorname{id}_A) \circ (\operatorname{id}_A \otimes T \otimes \operatorname{id}_A). \end{aligned}$$

Thus, we have found an element-free proof of the first equality in (13.27.5).

The second equality in (13.27.5)*:* The second equality in (13.27.5) is simply an element-free way to state $((ac)b)d = (a(cb))d$ for all $a, b, c, d \in A$. On the level of elements, this follows from applying associativity to the elements a, c and b of A, and then multiplying with d on the right. In other words, this follows from applying the associativity law $m_A \circ (m_A \otimes \operatorname{id}_A) = m_A \circ (\operatorname{id}_A \otimes m_A)$ to the first three tensorands of the tensor $a \otimes c \otimes b \otimes d = (\operatorname{id}_A \otimes T \otimes \operatorname{id}_A)(a \otimes b \otimes c \otimes d)$, and then applying m_A. In other words, this follows from applying the equality $(m_A \circ (m_A \otimes \operatorname{id}_A)) \otimes \operatorname{id}_A = (m_A \circ (\operatorname{id}_A \otimes m_A)) \otimes \operatorname{id}_A$ to the tensor $(\operatorname{id}_A \otimes T \otimes \operatorname{id}_A)(a \otimes b \otimes c \otimes d)$, and then applying m_A. Therefore, the second equality in (13.27.5) should follow from $m_A \circ (m_A \otimes \operatorname{id}_A) = m_A \circ (\operatorname{id}_A \otimes m_A)$ by tensoring both sides with id_A on the right and then composing them with the map $\operatorname{id}_A \otimes T \otimes \operatorname{id}_A$ on the right and with m_A on the left. And indeed, this is how it is proven:

$$\begin{aligned} & m_A \circ \underbrace{(m_A \otimes \operatorname{id}_A) \circ (m_A \otimes \operatorname{id}_A \otimes \operatorname{id}_A)}_{=(m_A \circ (m_A \otimes \operatorname{id}_A)) \otimes \operatorname{id}_A} \circ (\operatorname{id}_A \otimes T \otimes \operatorname{id}_A) \\ &= m_A \circ \left(\underbrace{(m_A \circ (m_A \otimes \operatorname{id}_A))}_{=m_A \circ (\operatorname{id}_A \otimes m_A)} \otimes \operatorname{id}_A \right) \circ (\operatorname{id}_A \otimes T \otimes \operatorname{id}_A) \\ &= m_A \circ \underbrace{((m_A \circ (\operatorname{id}_A \otimes m_A)) \otimes \operatorname{id}_A)}_{=(m_A \otimes \operatorname{id}_A) \circ (\operatorname{id}_A \otimes m_A \otimes \operatorname{id}_A)} \circ (\operatorname{id}_A \otimes T \otimes \operatorname{id}_A) \\ &= m_A \circ (m_A \otimes \operatorname{id}_A) \circ (\operatorname{id}_A \otimes m_A \otimes \operatorname{id}_A) \circ (\operatorname{id}_A \otimes T \otimes \operatorname{id}_A). \end{aligned}$$

Thus, we have obtained an element-free proof of the second equality in (13.27.5).

The third equality in (13.27.5)*:* The third equality in (13.27.5) is simply an element-free way to state $(a(cb))d = (a(bc))d$ for all $a, b, c, d \in A$. On the level of elements, this follows from applying commutativity to the elements c and b of A, then multiplying with a on the left, and then multiplying with d on the right. In other words, this follows from applying the commutativity law $m_A = m_A \circ T$ to the second and third tensorands of the tensor $a \otimes c \otimes b \otimes d = (\operatorname{id}_A \otimes T \otimes \operatorname{id}_A)(a \otimes b \otimes c \otimes d)$, and then applying $m_A \circ (m_A \otimes \operatorname{id}_A)$. In other words, this follows from applying the equality $\operatorname{id}_A \otimes m_A \otimes \operatorname{id}_A = \operatorname{id}_A \otimes (m_A \circ T) \otimes \operatorname{id}_A$ to the tensor $(\operatorname{id}_A \otimes T \otimes \operatorname{id}_A)(a \otimes b \otimes c \otimes d)$, and then applying $m_A \circ (m_A \otimes \operatorname{id}_A)$. Therefore, the third equality in (13.27.5) should follow from $m_A = m_A \circ T$ by tensoring both sides with id_A on the left and on the right and then composing them with the map $\operatorname{id}_A \otimes T \otimes \operatorname{id}_A$ on the right and with the map $m_A \circ (m_A \otimes \operatorname{id}_A)$ on the left.

And indeed, this is how it is proven:

$$\begin{aligned}
&m_A \circ (m_A \otimes \mathrm{id}_A) \circ \left(\mathrm{id}_A \otimes \underbrace{m_A}_{= m_A \circ T} \otimes \mathrm{id}_A \right) \circ (\mathrm{id}_A \otimes T \otimes \mathrm{id}_A) \\
&= m_A \circ (m_A \otimes \mathrm{id}_A) \circ \underbrace{(\mathrm{id}_A \otimes (m_A \circ T) \otimes \mathrm{id}_A)}_{= (\mathrm{id}_A \otimes m_A \otimes \mathrm{id}_A) \circ (\mathrm{id}_A \otimes T \otimes \mathrm{id}_A)} \circ (\mathrm{id}_A \otimes T \otimes \mathrm{id}_A) \\
&= m_A \circ (m_A \otimes \mathrm{id}_A) \circ (\mathrm{id}_A \otimes m_A \otimes \mathrm{id}_A) \circ \underbrace{(\mathrm{id}_A \otimes T \otimes \mathrm{id}_A) \circ (\mathrm{id}_A \otimes T \otimes \mathrm{id}_A)}_{\substack{= \mathrm{id}_{A \otimes A \otimes A \otimes A} \\ (\text{since } T \circ T = \mathrm{id}_{A \otimes A})}} \\
&= m_A \circ (m_A \otimes \mathrm{id}_A) \circ (\mathrm{id}_A \otimes m_A \otimes \mathrm{id}_A) .
\end{aligned}$$

Thus, we have obtained an element-free proof of the third equality in (13.27.5).

The fourth equality in (13.27.5)*:* The fourth equality in (13.27.5) is simply an element-free way to state $(a(bc))d = ((ab)c)d$ for all $a, b, c, d \in A$. On the level of elements, this follows from applying associativity to the elements a, b and c of A, and then multiplying with d on the right. In other words, this follows from applying the associativity law $m_A \circ (\mathrm{id}_A \otimes m_A) = m_A \circ (m_A \otimes \mathrm{id}_A)$ to the first three tensorands of the tensor $a \otimes b \otimes c \otimes d$, and then applying the map m_A. In other words, this follows from applying the equality $(m_A \circ (\mathrm{id}_A \otimes m_A)) \otimes \mathrm{id}_A = (m_A \circ (m_A \otimes \mathrm{id}_A)) \otimes \mathrm{id}_A$ to the tensor $a \otimes b \otimes c \otimes d$, and then applying the map m_A. Therefore, the fourth equality in (13.27.5) should follow from $m_A \circ (\mathrm{id}_A \otimes m_A) = m_A \circ (m_A \otimes \mathrm{id}_A)$ by tensoring both sides with id_A on the right and then composing them with the map m_A on the left. And indeed, this is how it is proven:

$$\begin{aligned}
&m_A \circ \underbrace{(m_A \otimes \mathrm{id}_A) \circ (\mathrm{id}_A \otimes m_A \otimes \mathrm{id}_A)}_{= (m_A \circ (\mathrm{id}_A \otimes m_A)) \otimes \mathrm{id}_A} \\
&= m_A \circ \left(\underbrace{(m_A \circ (\mathrm{id}_A \otimes m_A))}_{= m_A \circ (m_A \otimes \mathrm{id}_A)} \otimes \mathrm{id}_A \right) \\
&= m_A \circ \underbrace{((m_A \circ (m_A \otimes \mathrm{id}_A)) \otimes \mathrm{id}_A)}_{= (m_A \otimes \mathrm{id}_A) \circ (m_A \otimes \mathrm{id}_A \otimes \mathrm{id}_A)} \\
&= m_A \circ (m_A \otimes \mathrm{id}_A) \circ (m_A \otimes \mathrm{id}_A \otimes \mathrm{id}_A) .
\end{aligned}$$

Thus, we have obtained an element-free proof of the fourth equality in (13.27.5).

The fifth equality in (13.27.5)*:* The fifth equality in (13.27.5) is simply an element-free way to state $((ab)c)d = (ab)(cd)$ for all $a, b, c, d \in A$. On the level of elements, this follows from applying associativity to the elements ab, c and d of A. In other words, this follows from applying the associativity law $m_A \circ (m_A \otimes \mathrm{id}_A) = m_A \circ (\mathrm{id}_A \otimes m_A)$ to the tensor $ab \otimes c \otimes d = (m_A \otimes \mathrm{id}_A \otimes \mathrm{id}_A)(a \otimes b \otimes c \otimes d)$. Therefore, the fifth equality in (13.27.5) should follow from $m_A \circ (m_A \otimes \mathrm{id}_A) = m_A \circ (\mathrm{id}_A \otimes m_A)$ by composing both sides of this equality with $m_A \otimes \mathrm{id}_A \otimes \mathrm{id}_A$ on the right. And indeed, this is how it is proven:

$$\begin{aligned}
&\underbrace{m_A \circ (m_A \otimes \mathrm{id}_A)}_{= m_A \circ (\mathrm{id}_A \otimes m_A)} \circ (m_A \otimes \mathrm{id}_A \otimes \mathrm{id}_A) \\
&= m_A \circ \underbrace{(\mathrm{id}_A \otimes m_A) \circ (m_A \otimes \mathrm{id}_A \otimes \mathrm{id}_A)}_{= (\mathrm{id}_A \otimes m_A) \circ (m_A \otimes \mathrm{id}_{A \otimes A}) = m_A \otimes m_A} \\
&= m_A \circ (m_A \otimes m_A) .
\end{aligned}$$

Thus, we have obtained an element-free proof of the fifth equality in (13.27.5).

Now, all five equalities in the computation (13.27.5) are proven without reference to elements. Hence, we have shown $m_A \circ (m_A \otimes m_A) \circ (\mathrm{id}_A \otimes T \otimes \mathrm{id}_A) = m_A \circ (m_A \otimes m_A)$ in an element-free way. In other words, $m_A \circ m_{A \otimes A} = m_A \circ (m_A \otimes m_A)$ is proven in an element-free way (because the definition of $m_{A \otimes A}$ yields $m_{A \otimes A} = (m_A \otimes m_A) \circ (\mathrm{id}_A \otimes T \otimes \mathrm{id}_A)$). In other words, (13.27.2) is proven.

It remains to prove (13.27.3). We leave this very simple proof to the reader (noticing that it does not require the commutativity of A).

Now we know that both (13.27.2) and (13.27.3) hold. Statement 1 is thus proven in an element-free way.

We leave it to the reader to prove Statement 2 in an element-free way. Altogether, Statements 1 and 2 have now been proven. Therefore, part (b) of the exercise has been solved in an element-free way. Therefore, a solution of part (a) can be obtained mechanically by "reversing all arrows".

Of course, part (a) can alternatively be solved using the Sweedler notation.

13.28. **Solution to Exercise 1.5.8.** *Solution to Exercise 1.5.8.*

We shall use the notations introduced in Definition 1.4.8. Let us first state two simple facts from linear algebra:

Proposition 13.28.1. *Let U, V, U' and V' be four* **k***-modules. Let $x : U \to U'$ and $y : V \to V'$ be two* **k***-linear maps. Then,*

$$(y \otimes x) \circ T_{U,V} = T_{U',V'} \circ (x \otimes y) . \tag{13.28.1}$$

Proposition 13.28.2. *Let U and V be two* **k***-modules. Then,*

$$T_{V,U} \circ T_{U,V} = \operatorname{id}_{U \otimes V} . \tag{13.28.2}$$

We shall further need the following lemma, which follows easily from the definition of commutativity:

Lemma 13.28.3. *Let A be a commutative* **k***-algebra. Then, $m_A = m_A \circ T_{A,A}$.*

Proof of Lemma 13.28.3. Let T denote the twist map $T_{A,A}$. The **k**-algebra A is commutative if and only if the diagram (1.5.1) commutes (by Definition 1.5.1). Hence, the diagram (1.5.1) commutes (since A is commutative). In other words, we have $m = m \circ T$. In other words, we have $m_A = m_A \circ T_{A,A}$ (since $m = m_A$ and $T = T_{A,A}$). This proves Lemma 13.28.3. □

We now come to the solution of the exercise.

(a) This is easy to prove by computing with elements, but let us give an "element-free" proof.

Let $f : A \to B$ be a **k**-linear map. We shall prove the logical equivalence

$$(f \circ m_A = m_B \circ (f \otimes f) \circ T_{A,A}) \iff (f \circ m_A = m_B \circ (f \otimes f)) . \tag{13.28.3}$$

[*Proof of* (13.28.3)*:* We have assumed that at least one of the **k**-algebras A and B is commutative. In other words, the **k**-algebra A is commutative or the **k**-algebra B is commutative. Thus, we are in one of the following two cases:

Case 1: The **k**-algebra A is commutative.

Case 2: The **k**-algebra B is commutative.

Let us first consider Case 1. In this case, the **k**-algebra A is commutative. Thus, Lemma 13.28.3 yields $m_A = m_A \circ T_{A,A}$.

But Proposition 13.28.2 (applied to $U = A$ and $V = A$) yields $T_{A,A} \circ T_{A,A} = \operatorname{id}_{A \otimes A}$. Thus, if we have $f \circ m_A = m_B \circ (f \otimes f) \circ T_{A,A}$, then we also have

$$f \circ \underbrace{m_A}_{= m_A \circ T_{A,A}} = \underbrace{f \circ m_A}_{= m_B \circ (f \otimes f) \circ T_{A,A}} \circ T_{A,A} = m_B \circ (f \otimes f) \circ \underbrace{T_{A,A} \circ T_{A,A}}_{= \operatorname{id}_{A \otimes A}} = m_B \circ (f \otimes f) .$$

Thus, we have proved the implication

$$(f \circ m_A = m_B \circ (f \otimes f) \circ T_{A,A}) \implies (f \circ m_A = m_B \circ (f \otimes f)) . \tag{13.28.4}$$

On the other hand, if we have $f \circ m_A = m_B \circ (f \otimes f)$, then we also have

$$f \circ \underbrace{m_A}_{= m_A \circ T_{A,A}} = \underbrace{f \circ m_A}_{= m_B \circ (f \otimes f)} \circ T_{A,A} = m_B \circ (f \otimes f) \circ T_{A,A} .$$

Hence, we have proved the implication

$$(f \circ m_A = m_B \circ (f \otimes f)) \implies (f \circ m_A = m_B \circ (f \otimes f) \circ T_{A,A}) .$$

Combining this implication with (13.28.4), we obtain the logical equivalence

$$(f \circ m_A = m_B \circ (f \otimes f) \circ T_{A,A}) \iff (f \circ m_A = m_B \circ (f \otimes f)) .$$

Thus, (13.28.3) is proved in Case 1.

Let us next consider Case 2. In this case, the **k**-algebra B is commutative. Thus, Lemma 13.28.3 (applied to B instead of A) yields $m_B = m_B \circ T_{B,B}$, so that $m_B \circ T_{B,B} = m_B$. On the other hand, Proposition 13.28.1 (applied to $U = A$, $V = A$, $U' = B$, $V' = B$, $x = f$ and $y = f$) yields

$$(f \otimes f) \circ T_{A,A} = T_{B,B} \circ (f \otimes f) .$$

Now, we have the following chain of logical equivalences:

$$\left(f \circ m_A = m_B \circ \underbrace{(f \otimes f) \circ T_{A,A}}_{=T_{B,B} \circ (f \otimes f)} \right) \iff \left(f \circ m_A = \underbrace{m_B \circ T_{B,B}}_{=m_B} \circ (f \otimes f) \right) \iff (f \circ m_A = m_B \circ (f \otimes f)) .$$

Hence, the equivalence (13.28.3) holds. Thus, (13.28.3) is proved in Case 2.

We have now proved (13.28.3) in both Cases 1 and 2. Hence, the proof of (13.28.3) is complete (since Cases 1 and 2 cover all possibilities).]

Now, forget that we fixed f. We thus have proved the equivalence (13.28.3) for any **k**-linear map $f : A \to B$.

Now, the **k**-algebra homomorphisms from A to B are precisely the **k**-linear maps $f : A \to B$ that make the two diagrams

$$\begin{tikzcd} A \arrow[r, "f"] & B \\ A \otimes A \arrow[u, "m_A"] \arrow[r, "f \otimes f"] & B \otimes B \arrow[u, "m_B"'] \end{tikzcd} \qquad \text{and} \qquad \begin{tikzcd} A \arrow[rr, "f"] & & B \\ & \mathbf{k} \arrow[ul, "u_A"] \arrow[ur, "u_B"'] & \end{tikzcd}$$

commute (by Definition 1.3.1). In other words, the **k**-algebra homomorphisms from A to B are the **k**-linear maps $f : A \to B$ that satisfy $f \circ m_A = m_B \circ (f \otimes f)$ and $f \circ u_A = u_B$. Thus,

$$\begin{aligned} &\{\text{the } \mathbf{k}\text{-algebra homomorphisms from } A \text{ to } B\} \\ &= \{\text{the } \mathbf{k}\text{-linear maps } f : A \to B \text{ that satisfy } f \circ m_A = m_B \circ (f \otimes f) \text{ and } f \circ u_A = u_B\} . \end{aligned} \tag{13.28.5}$$

On the other hand, the **k**-algebra anti-homomorphisms from A to B are the **k**-linear maps $f : A \to B$ that satisfy $f \circ m_A = m_B \circ (f \otimes f) \circ T_{A,A}$ and $f \circ u_A = u_B$ (by Definition 1.4.8(b)). Hence,

$$\begin{aligned} &\{\text{the } \mathbf{k}\text{-algebra anti-homomorphisms from } A \text{ to } B\} \\ &= \left\{ \text{the } \mathbf{k}\text{-linear maps } f : A \to B \text{ that satisfy } \underbrace{f \circ m_A = m_B \circ (f \otimes f) \circ T_{A,A}}_{\substack{\iff (f \circ m_A = m_B \circ (f \otimes f)) \\ \text{(by (13.28.3))}}} \text{ and } f \circ u_A = u_B \right\} \\ &= \{\text{the } \mathbf{k}\text{-linear maps } f : A \to B \text{ that satisfy } f \circ m_A = m_B \circ (f \otimes f) \text{ and } f \circ u_A = u_B\} \\ &= \{\text{the } \mathbf{k}\text{-algebra homomorphisms from } A \text{ to } B\} \qquad \text{(by (13.28.5))} . \end{aligned}$$

In other words, the **k**-algebra anti-homomorphisms from A to B are the same as the **k**-algebra homomorphisms from A to B. This solves Exercise 1.5.8(a).

(b) We have solved Exercise 1.5.8(a) in an element-free fashion. Thus, by "reversing all arrows" in this solution of Exercise 1.5.8(a) (and replacing "algebra" and "commutative" by "coalgebra" and "cocommutative", respectively), we can obtain a solution to the dual of Exercise 1.5.8(a). Consequently, the dual of Exercise 1.5.8(a) holds.

The notion of a cocommutative **k**-coalgebra is dual to the notion of a commutative **k**-algebra (i.e., is obtained from the latter notion by "reversing all arrows"). Furthermore, it is easy to see that the notion of a **k**-coalgebra anti-homomorphism is dual to the notion of a **k**-algebra anti-homomorphism, whereas the notion of a **k**-coalgebra homomorphism is dual to the notion of a **k**-algebra homomorphism. Thus, the dual of Exercise 1.5.8(a) is the following exercise:

> *Exercise A:* Let A and B be two **k**-coalgebras, at least one of which is cocommutative. Prove that the **k**-coalgebra anti-homomorphisms from A to B are the same as the **k**-coalgebra homomorphisms from A to B.

As we have seen above, this dual holds. Thus, Exercise 1.5.8(b) is solved.

13.29. **Solution to Exercise 1.5.9.** *Solution to Exercise 1.5.9.* For every $1 \leq i < j \leq k$, let $t_{i,j}$ be the transposition in $\mathfrak{S}_k$ which transposes i with j. It is well-known that the symmetric group $\mathfrak{S}_k$ is generated by the transpositions $t_{i,i+1}$ with i ranging over $\{1, 2, \ldots, k-1\}$.

But let us notice that $(\rho(\pi)) \circ (\rho(\psi)) = \rho(\pi\psi)$ for any two elements π and ψ of $\mathfrak{S}_k$. Hence, the set of all $\pi \in \mathfrak{S}_k$ satisfying $m^{(k-1)} \circ (\rho(\pi)) = m^{(k-1)}$ is closed under multiplication. Since this set also contains the trivial permutation $\mathrm{id} = 1_{\mathfrak{S}_k} \in \mathfrak{S}_k$ and is closed under taking inverses (this is easy to check), this shows that this set is a subgroup of $\mathfrak{S}_k$. Therefore, if this set contains a set of generators of $\mathfrak{S}_k$, then this set must be the whole $\mathfrak{S}_k$. Hence, if we can show that this set contains the transposition $t_{i,i+1}$ for every $i \in \{1, 2, \ldots, k-1\}$, then it will follow that this set must be the whole $\mathfrak{S}_k$ (because the group $\mathfrak{S}_k$ is generated by the transpositions $t_{i,i+1}$ with i ranging over $\{1, 2, \ldots, k-1\}$), and this will entail that every $\pi \in \mathfrak{S}_k$ satisfies $m^{(k-1)} \circ (\rho(\pi)) = m^{(k-1)}$, which will solve the problem. Hence, in order to complete this solution, it is enough to check that the set of all $\pi \in \mathfrak{S}_k$ satisfying $m^{(k-1)} \circ (\rho(\pi)) = m^{(k-1)}$ contains the transposition $t_{i,i+1}$ for every $i \in \{1, 2, \ldots, k-1\}$. In other words, it is enough to check that

$$m^{(k-1)} \circ (\rho(t_{i,i+1})) = m^{(k-1)} \qquad \text{for all } i \in \{1, 2, \ldots, k-1\}. \tag{13.29.1}$$

Proof of (13.29.1)*:* Let $i \in \{1, 2, \ldots, k-1\}$. Let T denote the twist map $A \otimes A \to A \otimes A$ sending every pure tensor $a \otimes b$ to $b \otimes a$. Any k vectors $v_1, v_2, \ldots, v_k$ in A satisfy

$$\begin{aligned}
&(\rho(t_{i,i+1}))(v_1 \otimes v_2 \otimes \cdots \otimes v_k) \\
&= t_{i,i+1}(v_1 \otimes v_2 \otimes \cdots \otimes v_k) = \underbrace{v_1 \otimes v_2 \otimes \cdots \otimes v_{i-1}}_{=\mathrm{id}_{A^{\otimes(i-1)}}(v_1 \otimes v_2 \otimes \cdots \otimes v_{i-1})} \otimes \underbrace{v_{i+1} \otimes v_i}_{=T(v_i \otimes v_{i+1})} \otimes \underbrace{v_{i+2} \otimes v_{i+3} \otimes \cdots \otimes v_k}_{=\mathrm{id}_{A^{\otimes(k-1-i)}}(v_{i+2} \otimes v_{i+3} \otimes \cdots \otimes v_k)} \\
&= \mathrm{id}_{A^{\otimes(i-1)}}(v_1 \otimes v_2 \otimes \cdots \otimes v_{i-1}) \otimes T(v_i \otimes v_{i+1}) \otimes \mathrm{id}_{A^{\otimes(k-1-i)}}(v_{i+2} \otimes v_{i+3} \otimes \cdots \otimes v_k) \\
&= \left(\mathrm{id}_{A^{\otimes(i-1)}} \otimes T \otimes \mathrm{id}_{A^{\otimes(k-1-i)}}\right) \left(\underbrace{v_1 \otimes v_2 \otimes \cdots \otimes v_{i-1} \otimes v_i \otimes v_{i+1} \otimes v_{i+2} \otimes v_{i+3} \otimes \cdots \otimes v_k}_{=v_1 \otimes v_2 \otimes \cdots \otimes v_k} \right) \\
&= \left(\mathrm{id}_{A^{\otimes(i-1)}} \otimes T \otimes \mathrm{id}_{A^{\otimes(k-1-i)}}\right)(v_1 \otimes v_2 \otimes \cdots \otimes v_k).
\end{aligned}$$

In other words, the two maps $\rho(t_{i,i+1})$ and $\mathrm{id}_{A^{\otimes(i-1)}} \otimes T \otimes \mathrm{id}_{A^{\otimes(k-1-i)}}$ are equal to each other on every pure tensor. Being **k**-linear maps, these two maps must therefore be identical, i.e., we have $\rho(t_{i,i+1}) = \mathrm{id}_{A^{\otimes(i-1)}} \otimes T \otimes \mathrm{id}_{A^{\otimes(k-1-i)}}$.

But recall that A is commutative, whence the diagram (1.5.1) commutes. Thus, $m \circ T = m$.

We have $i \in \{1, 2, \ldots, k-1\}$, so that $1 \leq i \leq k-1$ and therefore $0 \leq i-1 \leq (k-1)-1$. Hence, in particular, $k-1 \geq 1 \geq 0$. We can thus apply Exercise 1.4.19(c) to $k-1$ and $i-1$ instead of k and i. As a result, we obtain

$$m^{(k-1)} = m^{((k-1)-1)} \circ \left(\mathrm{id}_{A^{\otimes(i-1)}} \otimes m \otimes \mathrm{id}_{A^{\otimes((k-1)-1-(i-1))}}\right) = m^{((k-1)-1)} \circ \left(\mathrm{id}_{A^{\otimes(i-1)}} \otimes m \otimes \mathrm{id}_{A^{\otimes(k-1-i)}}\right).$$

Hence,

$$\begin{aligned}
&\underbrace{m^{(k-1)}}_{=m^{((k-1)-1)} \circ \left(\mathrm{id}_{A^{\otimes(i-1)}} \otimes m \otimes \mathrm{id}_{A^{\otimes(k-1-i)}}\right)} \circ \underbrace{(\rho(t_{i,i+1}))}_{=\mathrm{id}_{A^{\otimes(i-1)}} \otimes T \otimes \mathrm{id}_{A^{\otimes(k-1-i)}}} \\
&= m^{((k-1)-1)} \circ \underbrace{\left(\mathrm{id}_{A^{\otimes(i-1)}} \otimes m \otimes \mathrm{id}_{A^{\otimes(k-1-i)}}\right) \circ \left(\mathrm{id}_{A^{\otimes(i-1)}} \otimes T \otimes \mathrm{id}_{A^{\otimes(k-1-i)}}\right)}_{=\left(\mathrm{id}_{A^{\otimes(i-1)}} \otimes (m \circ T) \otimes \mathrm{id}_{A^{\otimes(k-1-i)}}\right)} \\
&= m^{((k-1)-1)} \circ \left(\mathrm{id}_{A^{\otimes(i-1)}} \otimes \underbrace{(m \circ T)}_{=m} \otimes \mathrm{id}_{A^{\otimes(k-1-i)}} \right) \\
&= m^{((k-1)-1)} \circ \left(\mathrm{id}_{A^{\otimes(i-1)}} \otimes m \otimes \mathrm{id}_{A^{\otimes(k-1-i)}}\right) \\
&= m^{(k-1)}.
\end{aligned}$$

Thus, (13.29.1) is proven. This completes the solution of Exercise 1.5.9.

13.30. Solution to Exercise 1.5.10.

Solution to Exercise 1.5.10. Here is the statement:

> **Exercise.** Let C be a cocommutative $\mathbf{k}$-coalgebra, and let $k \in \mathbb{N}$. The symmetric group $\mathfrak{S}_k$ acts on the k-fold tensor power $C^{\otimes k}$ by permuting the tensor factors: $\sigma(v_1 \otimes v_2 \otimes \cdots \otimes v_k) = v_{\sigma^{-1}(1)} \otimes v_{\sigma^{-1}(2)} \otimes \cdots \otimes v_{\sigma^{-1}(k)}$ for all $v_1, v_2, \ldots, v_k \in C$ and $\sigma \in \mathfrak{S}_k$. For every $\pi \in \mathfrak{S}_k$, denote by $\rho(\pi)$ the action of π on $C^{\otimes k}$ (this is an endomorphism of $C^{\otimes k}$). Show that every $\pi \in \mathfrak{S}_k$ satisfies $(\rho(\pi)) \circ \Delta^{(k-1)} = \Delta^{(k-1)}$. (Recall that $\Delta^{(k-1)} : C \to C^{\otimes k}$ is defined as in Exercise 1.4.20 for $k \geq 1$, and by $\Delta^{(-1)} = \epsilon : C \to \mathbf{k}$ for $k = 0$.)

The solution of this exercise can be obtained from the above solution of Exercise 1.5.9 by reversing all arrows (and replacing A, m and $m^{(i)}$ by C, Δ and $\Delta^{(i)}$).

13.31. Solution to Exercise 1.5.11.

Solution to Exercise 1.5.11. (a) Let H be a $\mathbf{k}$-bialgebra and A be a commutative $\mathbf{k}$-algebra. Let f and g be two $\mathbf{k}$-algebra homomorphisms $H \to A$.

Since A is commutative, the map $m_A : A \otimes A \to A$ is a $\mathbf{k}$-algebra homomorphism (according to Exercise 1.5.6(b)). Also, $f \otimes g$ is a $\mathbf{k}$-algebra homomorphism (by Exercise 1.3.6(a), since f and g are $\mathbf{k}$-algebra homomorphisms). Finally, the axioms of a $\mathbf{k}$-bialgebra show that $\Delta_H : H \to H \otimes H$ is a $\mathbf{k}$-algebra homomorphism (since H is a $\mathbf{k}$-bialgebra). Now, the definition of convolution yields that $f \star g = m_A \circ (f \otimes g) \circ \Delta_H$. Thus, $f \star g$ is a composition of three $\mathbf{k}$-algebra homomorphisms (namely, of m_A, $f \otimes g$ and Δ_H), and hence a $\mathbf{k}$-algebra homomorphism itself. This solves Exercise 1.5.11(a).

(b) Exercise 1.5.11(b) can be solved by straightforward induction over k using (in the induction step) the result of Exercise 1.5.11(a) and (in the induction base) the fact that $u_A \circ \epsilon_H$ is a $\mathbf{k}$-algebra homomorphism (since u_A and ϵ_H are $\mathbf{k}$-algebra homomorphisms). The details are left to the reader.

(c) Let H be a Hopf algebra, and let A be a commutative $\mathbf{k}$-algebra. Let $f : H \to A$ be a $\mathbf{k}$-algebra homomorphism.

Proposition 1.4.10 shows that the antipode S of H is an algebra anti-endomorphism. Combined with the fact that f is a $\mathbf{k}$-algebra homomorphism, this yields that the composition $f \circ S$ is a $\mathbf{k}$-algebra anti-homomorphism[445]. But since algebra anti-homomorphisms $H \to A$ are the same as algebra homomorphisms $H \to A$ (since A is commutative), this yields that $f \circ S$ is a $\mathbf{k}$-algebra homomorphism. It remains to prove that $f \circ S$ is $\star$-inverse to f. But this follows from Proposition 1.4.26(a) (applied to $\alpha = f$). [446] This solves Exercise 1.5.11(c).

(d) Let A be a commutative $\mathbf{k}$-algebra. Then, the map $m_A : A \otimes A \to A$ is a $\mathbf{k}$-algebra homomorphism (according to Exercise 1.5.6(b)).

We need to prove that $m_A^{(k)}$ is a $\mathbf{k}$-algebra homomorphism (since $m^{(k)} = m_A^{(k)}$). To do so, it suffices to adapt the solution of Exercise 1.4.22(a) with only very minor changes (mainly we have to change H into A, replace "coalgebra" by "algebra" and use Exercise 1.3.6(a) instead of using Exercise 1.3.6(b)). The details of this adaptation are left to the reader.

[445] Here, we have used the (easily proved) fact that a composition of a $\mathbf{k}$-algebra homomorphism with a $\mathbf{k}$-algebra anti-homomorphism (in either order) always is a $\mathbf{k}$-algebra anti-homomorphism.

[446] Just for the sake of it, here is a standalone proof of this claim: The definition of convolution yields

$$\begin{aligned}
(f \circ S) \star f &= m_A \circ \left((f \circ S) \otimes \underbrace{f}_{= f \circ \mathrm{id}} \right) \circ \Delta_H = m_A \circ \underbrace{((f \circ S) \otimes (f \circ \mathrm{id}))}_{= (f \otimes f) \circ (S \otimes \mathrm{id})} \circ \Delta_H \\
&= \underbrace{m_A \circ (f \otimes f)}_{\substack{= f \circ m_H \\ \text{(since } f \text{ is a } \mathbf{k}\text{-algebra homomorphism)}}} \circ (S \otimes \mathrm{id}) \circ \Delta_H = f \circ \underbrace{m_H \circ (S \otimes \mathrm{id}) \circ \Delta_H}_{\substack{= u_H \circ \epsilon_H \\ \text{(by the commutativity of (1.4.3))}}} \\
&= \underbrace{f \circ u_H}_{\substack{= u_A \\ \text{(since } f \text{ is a } \mathbf{k}\text{-algebra homomorphism)}}} \circ \epsilon_H = u_A \circ \epsilon_H ,
\end{aligned}$$

and similarly $f \star (f \circ S) = u_A \circ \epsilon_H$, so that $f \circ S$ is $\star$-inverse to f.

(e) Let C' and C be two $\mathbf{k}$-coalgebras. Let $\gamma : C \to C'$ be a $\mathbf{k}$-coalgebra homomorphism. Let A and A' be two $\mathbf{k}$-algebras. Let $\alpha : A \to A'$ be a $\mathbf{k}$-algebra homomorphism. Let $f_1, f_2, \ldots, f_k$ be several $\mathbf{k}$-linear maps $C' \to A$.

Proposition 1.4.3 shows that the map

$$\operatorname{Hom}(C', A) \to \operatorname{Hom}(C, A'), \qquad f \mapsto \alpha \circ f \circ \gamma$$

is a $\mathbf{k}$-algebra homomorphism $(\operatorname{Hom}(C', A), \star) \to (\operatorname{Hom}(C, A'), \star)$. Denote this $\mathbf{k}$-algebra homomorphism by φ.

Since φ is a $\mathbf{k}$-algebra homomorphism, we have

$$\begin{aligned}\varphi(f_1 \star f_2 \star \cdots \star f_k) &= \underbrace{\varphi(f_1)}_{\substack{=\alpha\circ f_1\circ\gamma \\ \text{(by the definition of } \varphi)}} \star \underbrace{\varphi(f_2)}_{\substack{=\alpha\circ f_2\circ\gamma \\ \text{(by the definition of } \varphi)}} \star \cdots \star \underbrace{\varphi(f_k)}_{\substack{=\alpha\circ f_k\circ\gamma \\ \text{(by the definition of } \varphi)}} \\ &= (\alpha \circ f_1 \circ \gamma) \star (\alpha \circ f_2 \circ \gamma) \star \cdots \star (\alpha \circ f_k \circ \gamma).\end{aligned}$$

Hence,

$$(\alpha \circ f_1 \circ \gamma) \star (\alpha \circ f_2 \circ \gamma) \star \cdots \star (\alpha \circ f_k \circ \gamma) = \varphi(f_1 \star f_2 \star \cdots \star f_k) = \alpha \circ (f_1 \star f_2 \star \cdots \star f_k) \circ \gamma$$

(by the definition of φ). This solves Exercise 1.5.11(e).

(f) Let H be a commutative $\mathbf{k}$-bialgebra. Let k and ℓ be two nonnegative integers. Then, Exercise 1.5.11(b) (applied to $A = H$ and $f_i = \operatorname{id}_H$) yields that $\underbrace{\operatorname{id}_H \star \operatorname{id}_H \star \cdots \star \operatorname{id}_H}_{k \text{ times}}$ is a $\mathbf{k}$-algebra homomorphism $H \to H$. Since $\underbrace{\operatorname{id}_H \star \operatorname{id}_H \star \cdots \star \operatorname{id}_H}_{k \text{ times}} = \operatorname{id}_H^{\star k}$, this shows that $\operatorname{id}_H^{\star k}$ is a $\mathbf{k}$-algebra homomorphism $H \to H$. We can thus apply Exercise 1.5.11(e) to H, H, H, H, ℓ, id_H, $\operatorname{id}_H^{\star k}$ and id_H instead of C, C', A, A', k, f_i, α and γ. As a result, we obtain

$$\operatorname{id}_H^{\star k} \circ \left(\underbrace{\operatorname{id}_H \star \operatorname{id}_H \star \cdots \star \operatorname{id}_H}_{\ell \text{ times}}\right) \circ \operatorname{id}_H = \underbrace{\left(\operatorname{id}_H^{\star k} \circ \operatorname{id}_H \circ \operatorname{id}_H\right) \star \left(\operatorname{id}_H^{\star k} \circ \operatorname{id}_H \circ \operatorname{id}_H\right) \star \cdots \star \left(\operatorname{id}_H^{\star k} \circ \operatorname{id}_H \circ \operatorname{id}_H\right)}_{\ell \text{ times}}.$$

Since $\operatorname{id}_H^{\star k} \circ \operatorname{id}_H \circ \operatorname{id}_H = \operatorname{id}_H^{\star k}$ and $\underbrace{\operatorname{id}_H \star \operatorname{id}_H \star \cdots \star \operatorname{id}_H}_{\ell \text{ times}} = \operatorname{id}_H^{\star \ell}$, this rewrites as

$$\operatorname{id}_H^{\star k} \circ \operatorname{id}_H^{\star \ell} = \underbrace{\operatorname{id}_H^{\star k} \star \operatorname{id}_H^{\star k} \star \cdots \star \operatorname{id}_H^{\star k}}_{\ell \text{ times}} = \left(\operatorname{id}_H^{\star k}\right)^{\star \ell} = \operatorname{id}_H^{\star (k\ell)}.$$

This solves Exercise 1.5.11(f).

(g) Let H be a commutative $\mathbf{k}$-Hopf algebra.

First, it is easy to see that

$$\text{(13.31.1)} \qquad \operatorname{id}_H^{\star k} \text{ is a } \mathbf{k}\text{-algebra homomorphism } H \to H \qquad \text{for every } k \in \mathbb{Z}.$$

[447]

Furthermore, we have

$$\text{(13.31.2)} \qquad \operatorname{id}_H^{\star(-k)} = \operatorname{id}_H^{\star k} \circ S \qquad \text{for every } k \in \mathbb{Z}.$$

[447] *Proof of* (13.31.1)*:* Let $k \in \mathbb{Z}$. If k is nonnegative, then (13.31.1) can be proven just as in the solution to Exercise 1.5.11(f). Hence, for the rest of this proof of (13.31.1), we assume WLOG that k is not nonnegative. Thus, $k < 0$, so that $-k$ is nonnegative. Hence, Exercise 1.5.11(b) (applied to H, $-k$ and id_H instead of A, k and f_i) yields that $\underbrace{\operatorname{id}_H \star \operatorname{id}_H \star \cdots \star \operatorname{id}_H}_{-k \text{ times}}$ is a $\mathbf{k}$-algebra homomorphism $H \to H$. Since $\underbrace{\operatorname{id}_H \star \operatorname{id}_H \star \cdots \star \operatorname{id}_H}_{-k \text{ times}} = \operatorname{id}_H^{\star(-k)}$, this shows that $\operatorname{id}_H^{\star(-k)}$ is a $\mathbf{k}$-algebra homomorphism $H \to H$. Thus, Exercise 1.5.11(c) (applied to $A = H$ and $f = \operatorname{id}_H^{\star(-k)}$) yields that $\operatorname{id}_H^{\star(-k)} \circ S : H \to H$ is again a $\mathbf{k}$-algebra homomorphism, and is a $\star$-inverse to $\operatorname{id}_H^{\star(-k)}$.

Since $\operatorname{id}_H^{\star(-k)} \circ S$ is a $\star$-inverse to $\operatorname{id}_H^{\star(-k)}$, we have $\operatorname{id}_H^{\star(-k)} \circ S = \left(\operatorname{id}_H^{\star(-k)}\right)^{\star(-1)} = \operatorname{id}_H^{\star((-k)(-1))} = \operatorname{id}_H^{\star k}$. Hence, $\operatorname{id}_H^{\star k}$ is a $\mathbf{k}$-algebra homomorphism (since $\operatorname{id}_H^{\star(-k)} \circ S$ is a $\mathbf{k}$-algebra homomorphism), and thus (13.31.1) is proven.

[448]

Now, fix two integers k and ℓ. Due to (13.31.1), we know that $\operatorname{id}_H^{\star k}$ is a $\mathbf{k}$-algebra homomorphism $H \to H$. Hence, if ℓ is nonnegative, then we can prove $\operatorname{id}_H^{\star k} \circ \operatorname{id}_H^{\star \ell} = \operatorname{id}_H^{\star (k\ell)}$ just as we did in the solution to Exercise 1.5.11(f) (and thus finish the solution to Exercise 1.5.11(g)). Hence, for the rest of this solution, we WLOG assume that ℓ is not nonnegative. Thus, $\ell < 0$, so that $-\ell$ is nonnegative. We can thus apply Exercise 1.5.11(e) to H, H, H, H, $-\ell$, id_H, $\operatorname{id}_H^{\star k}$ and id_H instead of C, C', A, A', k, f_i, α and γ. As a result, we obtain

$$\operatorname{id}_H^{\star k} \circ \left(\underbrace{\operatorname{id}_H \star \operatorname{id}_H \star \cdots \star \operatorname{id}_H}_{-\ell \text{ times}} \right) \circ \operatorname{id}_H = \underbrace{\left(\operatorname{id}_H^{\star k} \circ \operatorname{id}_H \circ \operatorname{id}_H \right) \star \left(\operatorname{id}_H^{\star k} \circ \operatorname{id}_H \circ \operatorname{id}_H \right) \star \cdots \star \left(\operatorname{id}_H^{\star k} \circ \operatorname{id}_H \circ \operatorname{id}_H \right)}_{-\ell \text{ times}} .$$

Since $\operatorname{id}_H^{\star k} \circ \operatorname{id}_H \circ \operatorname{id}_H = \operatorname{id}_H^{\star k}$ and $\underbrace{\operatorname{id}_H \star \operatorname{id}_H \star \cdots \star \operatorname{id}_H}_{-\ell \text{ times}} = \operatorname{id}_H^{\star (-\ell)}$, this rewrites as

$$\operatorname{id}_H^{\star k} \circ \operatorname{id}_H^{\star (-\ell)} \circ \operatorname{id}_H = \underbrace{\operatorname{id}_H^{\star k} \star \operatorname{id}_H^{\star k} \star \cdots \star \operatorname{id}_H^{\star k}}_{-\ell \text{ times}} = \left(\operatorname{id}_H^{\star k} \right)^{\star (-\ell)} = \operatorname{id}_H^{\star (k(-\ell))} = \operatorname{id}_H^{\star (-k\ell)} .$$

In view of $\operatorname{id}_H^{\star (-\ell)} \circ \operatorname{id}_H = \operatorname{id}_H^{\star (-\ell)}$, this rewrites as

$$\operatorname{id}_H^{\star k} \circ \operatorname{id}_H^{\star (-\ell)} = \operatorname{id}_H^{\star (-k\ell)} .$$

Now, $\operatorname{id}_H^{\star \ell} = \operatorname{id}_H^{\star (-(-\ell))} = \operatorname{id}_H^{\star (-\ell)} \circ S$ (by (13.31.2), applied to $-\ell$ instead of k), and thus

$$\operatorname{id}_H^{\star k} \circ \underbrace{\operatorname{id}_H^{\star \ell}}_{= \operatorname{id}_H^{\star (-\ell)} \circ S} = \underbrace{\operatorname{id}_H^{\star k} \circ \operatorname{id}_H^{\star (-\ell)}}_{= \operatorname{id}_H^{\star (-k\ell)}} \circ S = \operatorname{id}_H^{\star (-k\ell)} \circ S.$$

Compared with $\operatorname{id}_H^{\star (k\ell)} = \operatorname{id}_H^{\star (-(-k\ell))} = \operatorname{id}_H^{\star (-k\ell)} \circ S$ (by (13.31.2), applied to $-k\ell$ instead of k), this yields $\operatorname{id}_H^{\star k} \circ \operatorname{id}_H^{\star \ell} = \operatorname{id}_H^{\star (k\ell)}$. This solves Exercise 1.5.11(g).

(h) The dual of Exercise 1.5.11(a) is the following exercise:

> If H is a $\mathbf{k}$-bialgebra and C is a cocommutative $\mathbf{k}$-coalgebra, and if f and g are two $\mathbf{k}$-coalgebra homomorphisms $C \to H$, then prove that $f \star g$ also is a $\mathbf{k}$-coalgebra homomorphism $C \to H$.

The solution of this exercise is obtained from our above solution of Exercise 1.5.11(a) by "reversing arrows" (and replacing "algebra" by "coalgebra", and applying Exercise 1.5.6(a) instead of Exercise 1.5.6(b), and using Exercise 1.3.6(b) instead of Exercise 1.3.6(a)).

The dual of Exercise 1.5.11(b) is the following exercise:

> If H is a $\mathbf{k}$-bialgebra and C is a cocommutative $\mathbf{k}$-coalgebra, and if $f_1, f_2, \ldots, f_k$ are several $\mathbf{k}$-coalgebra homomorphisms $C \to H$, then prove that $f_1 \star f_2 \star \cdots \star f_k$ also is a $\mathbf{k}$-coalgebra homomorphism $C \to H$.

This can be solved by induction over k in the same way as Exercise 1.5.11(b) (but now using the dual of Exercise 1.5.11(a) instead of Exercise 1.5.11(a) itself).

The dual of Exercise 1.5.11(c) is the following exercise:

> If H is a Hopf algebra and C is a cocommutative $\mathbf{k}$-coalgebra, and if $f : C \to H$ is a $\mathbf{k}$-coalgebra homomorphism, then prove that $S \circ f : C \to H$ (where S is the antipode of H) is again a $\mathbf{k}$-coalgebra homomorphism, and is a $\star$-inverse to f.

A solution of this can be obtained by reversing all arrows in the above solution of Exercise 1.5.11(c) (and using Exercise 1.4.28 in lieu of Proposition 1.4.10, and using Proposition 1.4.26(b) in lieu of Proposition 1.4.26(a)).

The dual of Exercise 1.5.11(d) is the following exercise:

[448] *Proof of* (13.31.2)*:* Let $k \in \mathbb{Z}$. We know that $\operatorname{id}_H^{\star k}$ is a $\mathbf{k}$-algebra homomorphism $H \to H$ (according to (13.31.1)). Thus, Exercise 1.5.11(c) (applied to $A = H$ and $f = \operatorname{id}_H^{\star k}$) yields that $\operatorname{id}_H^{\star k} \circ S : H \to H$ is again a $\mathbf{k}$-algebra homomorphism, and is a $\star$-inverse to $\operatorname{id}_H^{\star k}$.

Since $\operatorname{id}_H^{\star k} \circ S$ is a $\star$-inverse to $\operatorname{id}_H^{\star k}$, we have $\operatorname{id}_H^{\star k} \circ S = \left(\operatorname{id}_H^{\star k} \right)^{\star (-1)} = \operatorname{id}_H^{\star (k(-1))} = \operatorname{id}_H^{\star (-k)}$. This proves (13.31.2).

> If C is a cocommutative $\mathbf{k}$-coalgebra, then show that $\Delta^{(k)}$ is a $\mathbf{k}$-coalgebra homomorphism for every $k \in \mathbb{N}$. (The map $\Delta^{(k)} : C \to C^{\otimes(k+1)}$ is defined as in Exercise 1.4.20.)

This can be solved just as we solved Exercise 1.5.11(d), but again with all arrows reversed (and referring to Exercise 1.5.6(a) and Exercise 1.4.22(b) instead of Exercise 1.5.6(b) and Exercise 1.4.22(a), respectively).

The dual of Exercise 1.5.11(e) is Exercise 1.5.11(e) itself (up to renaming objects and maps).

The dual of Exercise 1.5.11(f) is the following exercise:

> If H is a cocommutative $\mathbf{k}$-bialgebra, and k and ℓ are two nonnegative integers, then prove that $\operatorname{id}_H^{\star \ell} \circ \operatorname{id}_H^{\star k} = \operatorname{id}_H^{\star(\ell k)}$.

This can be solved just as we solved Exercise 1.5.11(f), but again with all arrows reversed.

The dual of Exercise 1.5.11(g) is the following exercise:

> If H is a cocommutative $\mathbf{k}$-Hopf algebra, and k and ℓ are two integers, then prove that $\operatorname{id}_H^{\star \ell} \circ \operatorname{id}_H^{\star k} = \operatorname{id}_H^{\star(\ell k)}$.

This can be solved just as we solved Exercise 1.5.11(g), but again with all arrows reversed.

13.32. **Solution to Exercise 1.5.13.** *Solution to Exercise 1.5.13.* We will use the concepts of "$\star$-invertible" maps and their "$\star$-inverses" as defined in Exercise 1.4.29. We will also use the notations introduced in Definition 1.4.8.

Let A be a cocommutative Hopf algebra. Then, the $\mathbf{k}$-linear map $\operatorname{id}_A : A \to A$ is $\star$-invertible (since A is a Hopf algebra), and its $\star$-inverse $\operatorname{id}_A^{\star(-1)}$ is the antipode S of A. That is, $\operatorname{id}_A^{\star(-1)} = S$. Applying Exercise 1.4.29(b) to $C = A$ and $r = \operatorname{id}_A$, we now conclude that $\operatorname{id}_A^{\star(-1)}$ is a $\mathbf{k}$-coalgebra anti-homomorphism $A \to A$ (since id_A is a $\mathbf{k}$-coalgebra homomorphism $A \to A$). Since a $\mathbf{k}$-coalgebra anti-homomorphism $A \to A$ is the same thing as a $\mathbf{k}$-coalgebra homomorphism $A \to A$ (by Exercise 1.5.8(b), because A is cocommutative), this yields that $\operatorname{id}_A^{\star(-1)}$ is a $\mathbf{k}$-coalgebra homomorphism $A \to A$. In other words, S is a $\mathbf{k}$-coalgebra homomorphism $A \to A$ (since $\operatorname{id}_A^{\star(-1)} = S$). Now, Proposition 1.4.26(b) (applied to $H = A$, $C = A$ and $\gamma = S$) yields $S \circ S = S^{\star(-1)}$. But $S^{\star(-1)} = \operatorname{id}_A$ (since $S = \operatorname{id}_A^{\star(-1)}$). Hence, $S^2 = S \circ S = S^{\star(-1)} = \operatorname{id}_A$. This solves Exercise 1.5.13.

13.33. **Solution to Exercise 1.5.14.** *Solution to Exercise 1.5.14.* (a) We shall solve Exercise 1.5.14(a) in two ways.

First, here is a messy computational solution:

Fix $a \in A$, and assume WLOG that a is homogeneous of degree $n \in \mathbb{N}$. Let us prove that $(S \star E)(a)$ is primitive.

In fact, using the Sweedler notation, we can write $\Delta(a) = \sum_{(a)} a_1 \otimes a_2$ with all a_1 and a_2 homogeneous. Thus,

$$(S \star E)(a) = \sum_{(a)} S(a_1) \cdot \underbrace{E(a_2)}_{=(\deg a_2) \cdot a_2} = \sum_{(a)} (\deg a_2) S(a_1) \cdot a_2.$$

Hence,

$$\begin{aligned}
\Delta\left(\left(S\star E\right)\left(a\right)\right) &= \sum_{(a)}\left(\deg a_2\right)\Delta\left(S\left(a_1\right)\cdot a_2\right) = \sum_{(a)}\left(\deg a_2\right)\underbrace{\sum_{(a_1)}\sum_{(a_2)}\left(S\left(a_1\right)\right)_1\cdot\left(a_2\right)_1\otimes\left(S\left(a_1\right)\right)_2\cdot\left(a_2\right)_2}_{\substack{=\sum_{(a_1)}\sum_{(a_2)}\left(S\left(\left(a_1\right)_2\right)\right)\cdot\left(a_2\right)_1\otimes\left(S\left(\left(a_1\right)_1\right)\right)\cdot\left(a_2\right)_2 \\ \text{(here we used } \left(S(a_1)\right)_1\otimes\left(S(a_1)\right)_2=S\left(\left(a_1\right)_2\right)\otimes S\left(\left(a_1\right)_1\right)\text{,} \\ \text{which is a consequence of Exercise 1.4.28)}}} \\
&= \sum_{(a)}\left(\deg a_2\right)\sum_{(a_1)}\sum_{(a_2)}\left(S\left(\left(a_1\right)_2\right)\right)\cdot\left(a_2\right)_1\otimes\left(S\left(\left(a_1\right)_1\right)\right)\cdot\left(a_2\right)_2 \\
&= \sum_{(a)}\left(\deg a_3+\deg a_4\right)\left(S\left(a_2\right)\right)\cdot a_3\otimes\left(S\left(a_1\right)\right)\cdot a_4 \\
&= \sum_{(a)}\left(\deg a_3\right)\left(S\left(a_2\right)\right)\cdot a_3\otimes\left(S\left(a_1\right)\right)\cdot a_4+\sum_{(a)}\left(\deg a_4\right)\left(S\left(a_2\right)\right)\cdot a_3\otimes\left(S\left(a_1\right)\right)\cdot a_4 \\
&= \sum_{(a)}\left(\deg a_2\right)\left(S\left(a_1\right)\right)\cdot a_2\otimes\underbrace{\left(S\left(a_3\right)\right)\cdot a_4}_{=\epsilon(a_3)}+\sum_{(a)}\left(\deg a_2\right)\underbrace{\left(S\left(a_3\right)\right)\cdot a_4}_{=\epsilon(a_3)}\otimes\left(S\left(a_1\right)\right)\cdot a_2 \\
&\qquad \text{(by the cocommutativity of } A\text{)} \\
&= \underbrace{\sum_{(a)}\left(\deg a_2\right)\left(S\left(a_1\right)\right)\cdot a_2}_{=(S\star E)(a)}\otimes 1+1\otimes\underbrace{\sum_{(a)}\left(\deg a_2\right)\left(S\left(a_1\right)\right)\cdot a_2}_{=(S\star E)(a)} \\
&= \left(S\star E\right)\left(a\right)\otimes 1+1\otimes\left(S\star E\right)\left(a\right).
\end{aligned}$$

This (slightly unclean but easily formalizable) computation shows that $\left(S\star E\right)\left(a\right)$ is primitive, and similarly the same can be shown for $\left(E\star S\right)\left(a\right)$. This proves (a).

There is, however, a nicer proof: A *coderivation* of a $\mathbf{k}$-coalgebra (C,Δ,ϵ) is defined as a $\mathbf{k}$-linear map $F : C \to C$ such that $\Delta\circ F = (F\otimes\operatorname{id}+\operatorname{id}\otimes F)\circ\Delta$. (The reader can check that this axiom is the result of writing the axiom for a derivation in element-free terms and reversing all arrows. Nothing less should be expected.) It is easy to see (by checking on each homogeneous component) that E is a coderivation. Hence, it will be enough to check that $\left(S\star f\right)\left(a\right)$ and $\left(f\star S\right)\left(a\right)$ are primitive whenever $f : A\to A$ is a coderivation and $a\in A$. So fix a coderivation $f : A\to A$. Notice that the antipode S of A is a coalgebra anti-endomorphism (by Exercise 1.4.28), thus a coalgebra endomorphism (because coalgebra anti-endomorphisms of a cocommutative coalgebra are precisely the same as its coalgebra endomorphisms[449]). Thus, $\Delta\circ S = (S\otimes S)\circ\Delta$. Moreover, $\Delta : A\to A\otimes A$ is a coalgebra homomorphism[450] and an algebra homomorphism (since A is a bialgebra). Applying (1.4.2) to $A\otimes A$, A, A, Δ, id_A, S and f instead of A', C, C', α, γ, f and g, we obtain

$$\begin{aligned}
\Delta\circ\left(S\star f\right) &= \underbrace{\left(\Delta\circ S\right)}_{=(S\otimes S)\circ\Delta}\star\underbrace{\left(\Delta\circ f\right)}_{\substack{=(f\otimes\operatorname{id}+\operatorname{id}\otimes f)\circ\Delta \\ \text{(since } f \text{ is a coderivation)}}} \\
&= \left(\left(S\otimes S\right)\circ\Delta\right)\star\left(\left(f\otimes\operatorname{id}+\operatorname{id}\otimes f\right)\circ\Delta\right) = \left(\left(S\otimes S\right)\star\left(f\otimes\operatorname{id}+\operatorname{id}\otimes f\right)\right)\circ\Delta \\
&\qquad \left(\begin{array}{c}\text{by (1.4.2), applied to } A\otimes A,\ A\otimes A,\ A\otimes A\otimes A\otimes A,\ A\otimes A, \\ \operatorname{id},\ \Delta,\ S\otimes S \text{ and } f\otimes\operatorname{id}+\operatorname{id}\otimes f \text{ instead of } A,\ A',\ C,\ C',\ \alpha,\ \gamma,\ f \text{ and } g\end{array}\right) \\
&= \left(\left(S\otimes S\right)\star\left(f\otimes\operatorname{id}\right)\right)\circ\Delta+\left(\left(S\otimes S\right)\star\left(\operatorname{id}\otimes f\right)\right)\circ\Delta. \qquad (13.33.1)
\end{aligned}$$

[449] This is the result of Exercise 1.5.8(b).

[450] In fact, Exercise 1.5.6(a) (applied to $C = A$) shows that A is cocommutative if and only if $\Delta : A\to A\otimes A$ is a coalgebra homomorphism. But we know that A is cocommutative, so that $\Delta : A\to A\otimes A$ is a coalgebra homomorphism.

But Exercise 1.4.4(a) yields $(S \otimes S) \star (f \otimes \operatorname{id}) = (S \star f) \otimes \underbrace{(S \star \operatorname{id})}_{= u\epsilon} = (S \star f) \otimes u\epsilon$ and similarly $(S \otimes S) \star (\operatorname{id} \otimes f) = u\epsilon \otimes (S \star f)$. Now, (13.33.1) becomes

$$\begin{aligned} \Delta \circ (S \star f) &= \underbrace{((S \otimes S) \star (f \otimes \operatorname{id}))}_{=(S\star f)\otimes u\epsilon} \circ \Delta + \underbrace{((S \otimes S) \star (\operatorname{id} \otimes f))}_{=u\epsilon \otimes (S \star f)} \circ \Delta \\ &= ((S \star f) \otimes u\epsilon) \circ \Delta + (u\epsilon \otimes (S \star f)) \circ \Delta . \end{aligned}$$

Hence, every $a \in A$ satisfies (using the Sweedler notation)

$$\begin{aligned} (\Delta \circ (S \star f))(a) &= (((S \star f) \otimes u\epsilon) \circ \Delta + (u\epsilon \otimes (S \star f)) \circ \Delta)(a) \\ &= ((S \star f) \otimes u\epsilon)(\Delta(a)) + (u\epsilon \otimes (S \star f))(\Delta(a)) \\ &= \sum_{(a)} (S \star f)(a_1) \otimes (u\epsilon)(a_2) + \sum_{(a)} (u\epsilon)(a_1) \otimes (S \star f)(a_2) \\ &= \underbrace{\sum_{(a)} (S \star f)(a_1)\, \epsilon(a_2)}_{=(S\star f)(a)} \otimes 1 + 1 \otimes \underbrace{\sum_{(a)} \epsilon(a_1)\,(S \star f)(a_2)}_{=(S\star f)(a)} \\ &= (S \star f)(a) \otimes 1 + 1 \otimes (S \star f)(a) . \end{aligned}$$

In other words, for every $a \in A$, the element $(S \star f)(a)$ is primitive. Similarly the same can be shown for $(f \star S)(a)$, and so we are done.

(b) is a very simple computation. (Alternatively, the $(S \star E)(p) = E(p)$ part follows from applying part (c) to $a = 1$, and similarly one can show $(E \star S)(p) = E(p)$.)

(c) This is computational again: It is straightforward to check that E is a derivation of the algebra A. Now,

$$\begin{aligned} \Delta(ap) &= \underbrace{\Delta(a)}_{=\sum_{(a)} a_1 \otimes a_2} \underbrace{\Delta(p)}_{=p\otimes 1 + 1 \otimes p} = \left(\sum_{(a)} a_1 \otimes a_2 \right) (p \otimes 1 + 1 \otimes p) \\ &= \sum_{(a)} a_1 p \otimes a_2 + \sum_{(a)} a_1 \otimes a_2 p, \end{aligned}$$

so that

$$\begin{aligned} (S \star E)(ap) &= \sum_{(a)} \underbrace{S(a_1 p)}_{\substack{=S(p)S(a_1) \\ \text{(since } S \text{ is an algebra} \\ \text{anti-endomorphism)}}} E(a_2) + \sum_{(a)} S(a_1) \underbrace{E(a_2 p)}_{\substack{=E(a_2)p + a_2 E(p) \\ \text{(since } E \text{ is a derivation)}}} \\ &= \sum_{(a)} \underbrace{S(p)}_{\substack{=-p \\ \text{(by Proposition 1.4.17)}}} S(a_1)\, E(a_2) + \sum_{(a)} S(a_1)\,(E(a_2)\, p + a_2 E(p)) \\ &= -p \underbrace{\sum_{(a)} S(a_1)\, E(a_2)}_{=(S\star E)(a)} + \underbrace{\sum_{(a)} S(a_1)\, E(a_2)}_{=(S\star E)(a)} p + \underbrace{\sum_{(a)} S(a_1)\, a_2}_{=u(\epsilon(a))} E(p) \\ &= \underbrace{-p\,(S \star E)(a) + (S \star E)(a)\, p}_{=[(S\star E)(a),p]} + \underbrace{u(\epsilon(a))\, E(p)}_{=\epsilon(a)E(p)} \\ &= [(S \star E)(a), p] + \epsilon(a)\, E(p) , \end{aligned}$$

thus proving part (c).

(d) Assume that A is connected and that $\mathbb{Q}$ is a subring of $\mathbf{k}$. Let B be the $\mathbf{k}$-subalgebra of A generated by $\mathfrak{p}$. In order to prove part (d), we need to show that $A \subset B$. Clearly, $\mathfrak{p} \subset B$.

Consider the grading $A = \bigoplus_{n \geq 0} A_n$ of A. Now, we need to prove that $A \subset B$. In order to prove this, it is clearly enough to show that $A_n \subset B$ for every $n \in \mathbb{N}$. We will prove this by strong induction over n.

Let $n \in \mathbb{N}$, and assume that we have shown that $A_m \subset B$ for every nonnegative integer $m < n$. We need to prove that $A_n \subset B$.

If $n = 0$, then this is obvious (since $A_0 = \mathbf{k} \cdot 1_A \subset B$). Thus, for the rest of this proof, we WLOG assume that $n \neq 0$. Hence, n is a positive integer.

Let $a \in A_n$. We know that $(S \star E)(a)$ is primitive (by part (a)), so that $(S \star E)(a) \in \mathfrak{p} \subset B$. On the other hand, $a \in A_n$ shows that

$$\begin{aligned}\Delta(a) &\in \Delta(A_n) \subset (A \otimes A)_n \qquad \text{(since } \Delta \text{ is a graded map)} \\ &= \sum_{i=0}^{n} A_i \otimes A_{n-i} = A_0 \otimes A_n + \sum_{i=1}^{n} A_i \otimes A_{n-i}.\end{aligned}$$

We can thus write $\Delta(a)$ in the form $\Delta(a) = u + v$ for some $u \in A_0 \otimes A_n$ and $v \in \sum_{i=1}^{n} A_i \otimes A_{n-i}$. Consider these u and v.

We are first going to prove that $u = 1 \otimes a$. Indeed, applying $\epsilon \otimes \mathrm{id}$ to the equality $\Delta(a) = u + v$, we obtain $(\epsilon \otimes \mathrm{id})(\Delta(a)) = (\epsilon \otimes \mathrm{id})(u + v) = (\epsilon \otimes \mathrm{id})(u) + (\epsilon \otimes \mathrm{id})(v)$. But since $(\epsilon \otimes \mathrm{id})(\Delta(a)) = a$ (by the commutativity of the diagram (1.2.2)) and $(\epsilon \otimes \mathrm{id})(v) = 0$ (because

$$\begin{aligned}(\epsilon \otimes \mathrm{id})(v) &\in (\epsilon \otimes \mathrm{id})\left(\sum_{i=1}^{n} A_i \otimes A_{n-i}\right) \qquad \left(\text{since } v \in \sum_{i=1}^{n} A_i \otimes A_{n-i}\right) \\ &= \sum_{i=1}^{n} \underbrace{\epsilon(A_i)}_{\substack{=0 \\ (\text{since } i>0)}} \otimes A_{n-i} = \sum_{i=1}^{n} 0 \otimes A_{n-i} = 0\end{aligned}$$

), this rewrites as $a = (\epsilon \otimes \mathrm{id})(u) + 0$. In other words, $a = (\epsilon \otimes \mathrm{id})(u)$. But the element u has the form $u = 1_A \otimes u'$ for some $u' \in A_n$ (because $u \in \underbrace{A_0}_{=\mathbf{k} \cdot 1_A} \otimes A_n = \mathbf{k} \cdot 1_A \otimes A_n = 1_A \otimes A_n$). This $u' \in A_n$ can be recovered by $u' = (\epsilon \otimes \mathrm{id})(u)$ (since it satisfies $(\epsilon \otimes \mathrm{id})\left(\underbrace{u}_{=1_A \otimes u'}\right) = (\epsilon \otimes \mathrm{id})(1_A \otimes u') = \underbrace{\epsilon(1_A)}_{=1} u' = u'$), and thus simply equals a (since $a = (\epsilon \otimes \mathrm{id})(u)$). Thus, $u = 1_A \otimes a$.

Now,

$$\begin{aligned}\underbrace{(S \star E)}_{=m \circ (S \otimes E) \circ \Delta}(a) &= (m \circ (S \otimes E) \circ \Delta)(a) = (m \circ (S \otimes E))\left(\underbrace{\Delta(a)}_{=u+v}\right) \\ &= (m \circ (S \otimes E))(u + v) = (m \circ (S \otimes E))(u) + (m \circ (S \otimes E))(v).\end{aligned}$$

Since

$$\begin{aligned}(m \circ (S \otimes E))\left(\underbrace{u}_{=1_A \otimes a}\right) &= (m \circ (S \otimes E))(1_A \otimes a) = m\left(\underbrace{(S \otimes E)(1_A \otimes a)}_{=S(1_A) \otimes E(a)}\right) \\ &= m(S(1_A) \otimes E(a)) = \underbrace{S(1_A)}_{=1_A} \underbrace{E(a)}_{\substack{=na \\ (\text{since } a \in A_n)}} = na\end{aligned}$$

and

$$\begin{aligned}
&(m\circ(S\otimes E))\left(\underbrace{v}_{\in\sum_{i=1}^{n} A_i\otimes A_{n-i}}\right)\\
&\in (m\circ(S\otimes E))\left(\sum_{i=1}^{n} A_i\otimes A_{n-i}\right)=\sum_{i=1}^{n}(m\circ(S\otimes E))\left(A_i\otimes A_{n-i}\right)\\
&=\sum_{i=1}^{n} m\underbrace{\left(\left(S\otimes E\right)\left(A_i\otimes A_{n-i}\right)\right)}_{=S(A_i)\otimes E(A_{n-i})}=\sum_{i=1}^{n} m\left(S\left(A_i\right)\otimes E\left(A_{n-i}\right)\right)\\
&=\sum_{i=1}^{n} S\left(A_i\right)\cdot E\left(A_{n-i}\right)=\sum_{i=1}^{n-1} S\left(A_i\right)\cdot E\left(A_{n-i}\right)+S\left(A_n\right)\cdot\underbrace{E\left(A_0\right)}_{\substack{=0\\ \text{(by the definition}\\ \text{of } E)}}\\
&=\sum_{i=1}^{n-1}\underbrace{S\left(A_i\right)}_{\substack{\subset A_i\\ \text{(since the map } S \text{ is graded)}}}\cdot\underbrace{E\left(A_{n-i}\right)}_{\substack{\subset A_{n-i}\\ \text{(since the map } E \text{ is graded)}}}\\
&\subset\sum_{i=1}^{n-1}\underbrace{A_i}_{\substack{\subset B\\ \text{(since } A_m\subset B \text{ for every}\\ \text{nonnegative integer } m<n)}}\cdot\underbrace{A_{n-i}}_{\substack{\subset B\\ \text{(since } A_m\subset B \text{ for every}\\ \text{nonnegative integer } m<n)}}\subset\sum_{i=1}^{n-1} B\cdot B\subset B
\end{aligned}$$

(since B is an algebra), this becomes

$$(S\star E)(a)=\underbrace{(m\circ(S\otimes E))(u)}_{=na}+\underbrace{(m\circ(S\otimes E))(v)}_{\in B}\in na+B,$$

so that

$$na\in\underbrace{(S\star E)(a)}_{\in B}+B\subset B+B\subset B.$$

Since n is positive and $\mathbb{Q}$ is a subring of $\mathbf{k}$, we can divide this by n and obtain $a\in B$.

We have thus shown that $a\in B$ for every $a\in A_n$. Hence, $A_n\subset B$. This completes the induction step. Thus, we know that $A_n\subset B$ for every $n\in\mathbb{N}$. Consequently, $A\subset B$, and the proof of (d) is complete.

This solution of part (d) is not the most generalizable one – for instance, (d) also holds if A is connected filtered instead of connected graded, and then a different argument is necessary. This is a part of the Cartier-Milnor-Moore theorem, and appears e.g. in [60, §3.2].

(e) If $a\in T(V)$ is homogeneous of positive degree and $p\in V$, then part (c) yields

$$\begin{aligned}
(S\star E)(ap)&=\left[(S\star E)(a),p\right]+\underbrace{\epsilon(a)}_{=0}E(p)\qquad\text{(since } p \text{ is primitive in } T(V))\\
&=\left[(S\star E)(a),p\right].
\end{aligned}$$

This allows proving (e) by induction over n, with the induction base $n=1$ being a consequence of part (b).

13.34. **Solution to Exercise 1.6.1.** *Solution to Exercise 1.6.1.* (a) Let $\mathfrak{u}$ be the map $\epsilon_C^*\circ s:\mathbf{k}\to C^*$. Let $\mathfrak{m}$ be the map $\Delta_C^*\circ\rho_{C,C}:C^*\otimes C^*\to C^*$. Our goal is to prove that C^*, endowed with $\mathfrak{m}$ as the associative operation and $\mathfrak{u}$ as the unity map, is a $\mathbf{k}$-algebra. In order to achieve this, we need to prove that the diagrams (1.2.1) and (1.2.2) with A, m and u replaced by C^*, $\mathfrak{m}$ and $\mathfrak{u}$ are commutative.

Let us first prove that the diagram (1.2.1) with A, m and u replaced by C^*, $\mathfrak{m}$ and $\mathfrak{u}$ is commutative. In other words, let us prove that the diagram

(13.34.1)
$$\begin{tikzcd}
& C^* \otimes C^* \otimes C^* \arrow[ld, "\mathfrak{m}\otimes \mathrm{id}"'] \arrow[rd, "\mathrm{id} \otimes \mathfrak{m}"] & \\
C^* \otimes C^* \arrow[rd, "\mathfrak{m}"] & & C^* \otimes C^* \arrow[ld, "\mathfrak{m}"'] \\
& C^* &
\end{tikzcd}$$

is commutative.

Indeed, consider the diagram

(13.34.2)
$$\begin{tikzcd}
& & & C^* \otimes C^* \otimes C^* \arrow[ld, "\rho_{C,C} \otimes \mathrm{id}"'] \arrow[rd, "\mathrm{id} \otimes \rho_{C,C}"] \arrow[lldd, bend right, "\mathfrak{m}\otimes \mathrm{id}"'] \arrow[rrdd, bend left, "\mathrm{id} \otimes \mathfrak{m}"] & & & \\
& & \left(C \otimes C\right)^* \otimes C^* \arrow[ld, "\Delta_C^* \otimes \mathrm{id}"'] \arrow[rd, "\rho_{C \otimes C, C}"] & & C^* \otimes \left(C \otimes C\right)^* \arrow[ld, "\rho_{C, C \otimes C}"'] \arrow[rd, "\mathrm{id} \otimes \Delta_C^*"] & & \\
& C^* \otimes C^* \arrow[rd, "\rho_{C,C}"] \arrow[rrdd, bend right, "\mathfrak{m}"'] & & \left(C \otimes C \otimes C\right)^* \arrow[ld, "\left(\Delta_C \otimes \mathrm{id}\right)^*"'] \arrow[rd, "\left(\mathrm{id} \otimes \Delta_C\right)^*"] & & C^* \otimes C^* \arrow[ld, "\rho_{C,C}"'] \arrow[lldd, bend left, "\mathfrak{m}"] & \\
& & \left(C \otimes C\right)^* \arrow[rd, "\Delta_C^*"] & & \left(C \otimes C\right)^* \arrow[ld, "\Delta_C^*"'] & & \\
& & & C^* & & &
\end{tikzcd}$$
.

We are going to show that this diagram is commutative. In order to do so, we will show that its little squares and triangles are commutative.

The triangle

$$\begin{tikzcd}
& & C^* \otimes C^* \otimes C^* \arrow[ld, "\rho_{C,C} \otimes \mathrm{id}"'] \arrow[lldd, bend right, "\mathfrak{m}\otimes \mathrm{id}"'] \\
& \left(C \otimes C\right)^* \otimes C^* \arrow[ld, "\Delta_C^* \otimes \mathrm{id}"'] & \\
C^* \otimes C^* & &
\end{tikzcd}$$

is commutative, since $\mathfrak{m} = \Delta_C^* \circ \rho_{C,C}$. For the same reason, the triangle

$$\begin{tikzcd}
C^* \otimes C^* \otimes C^* \arrow[rd, "\mathrm{id} \otimes \rho_{C,C}"] \arrow[rrdd, bend left, "\mathrm{id} \otimes \mathfrak{m}"] & & \\
& C^* \otimes \left(C \otimes C\right)^* \arrow[rd, "\mathrm{id} \otimes \Delta_C^*"'] & \\
& & C^* \otimes C^*
\end{tikzcd}$$

is commutative. The commutativity of the triangles

$$\begin{tikzcd}
C^* \otimes C^* \arrow[dr, "\rho_{C,C}"] \arrow[ddrr, bend right=60, "\mathfrak{m}"'] & & \\
& (C \otimes C)^* \arrow[dr, "\Delta_C^*"] & \\
& & C^*
\end{tikzcd}
\qquad \text{and} \qquad
\begin{tikzcd}
& & C^* \otimes C^* \arrow[dl, "\rho_{C,C}"'] \arrow[ddll, bend left=60, "\mathfrak{m}"] \\
& (C \otimes C)^* \arrow[dl, "\Delta_C^*"'] & \\
C^* & &
\end{tikzcd}$$

also clearly follows from $\mathfrak{m} = \Delta_C^* \circ \rho_{C,C}$. The commutativity of the square

$$\begin{tikzcd}
& C^* \otimes C^* \otimes C^* \arrow[dl, "\rho_{C,C} \otimes \mathrm{id}"] \arrow[dr, "\mathrm{id} \otimes \rho_{C,C}"] & \\
(C \otimes C)^* \otimes C^* \arrow[dr, "\rho_{C \otimes C, C}"] & & C^* \otimes (C \otimes C)^* \arrow[dl, "\rho_{C, C \otimes C}"] \\
& (C \otimes C \otimes C)^* &
\end{tikzcd}$$

is a basic linear-algebraic fact. The commutativity of the square

$$\begin{tikzcd}
& (C \otimes C)^* \otimes C^* \arrow[dl, "\Delta_C^* \otimes \mathrm{id}"] \arrow[dr, "\rho_{C \otimes C, C}"] & \\
C^* \otimes C^* \arrow[dr, "\rho_{C,C}"] & & (C \otimes C \otimes C)^* \arrow[dl, "(\Delta_C \otimes \mathrm{id})^*"] \\
& (C \otimes C)^* &
\end{tikzcd}$$

is a particular case of the following linear-algebraic fact: If X, Y and B are three $\mathbf{k}$-modules and $f : Y \to X$ is a $\mathbf{k}$-linear map, then the diagram

$$\begin{tikzcd}
& X^* \otimes B^* \arrow[dl, "f^* \otimes \mathrm{id}"] \arrow[dr, "\rho_{X,B}"] & \\
Y^* \otimes B^* \arrow[dr, "\rho_{Y,B}"] & & (X \otimes B)^* \arrow[dl, "(f \otimes \mathrm{id})^*"] \\
& (Y \otimes B)^* &
\end{tikzcd}$$

is commutative.[451] Similarly, the commutativity of the square

$$\begin{tikzcd}
& C^* \otimes (C \otimes C)^* \arrow[dl, "\rho_{C, C \otimes C}"] \arrow[dr, "\mathrm{id} \otimes \Delta_C^*"'] & \\
(C \otimes C \otimes C)^* \arrow[dr, "(\mathrm{id} \otimes \Delta_C)^*"] & & C^* \otimes C^* \arrow[dl, "\rho_{C,C}"'] \\
& (C \otimes C)^* &
\end{tikzcd}$$

[451] This is part of the reason why $\rho_{U,V}$ is functorial in U.

can be shown. Finally, the square

$$\begin{tikzcd}
& (C\otimes C\otimes C)^* \arrow[dl, "(\Delta_C\otimes \mathrm{id})^*"'] \arrow[dr, "(\mathrm{id}\otimes \Delta_C)^*"] & \\
(C\otimes C)^* \arrow[dr, "\Delta_C^*"] & & (C\otimes C)^* \arrow[dl, "\Delta_C^*"] \\
& C^* &
\end{tikzcd}$$

is commutative, because it is obtained by dualizing the commutative diagram (1.2.1).

We thus have shown that every little square and every little triangle of the diagram (13.34.2) is commutative. Hence, the whole diagram (13.34.2) is commutative. In particular, the square formed by the four long curved arrows in (13.34.2) is commutative. But this square is precisely the diagram (13.34.1). Hence, we have shown that the diagram (13.34.1) is commutative. In other words, the diagram (1.2.1) with A, m and u replaced by C^*, $\mathfrak{m}$ and $\mathfrak{u}$ is commutative.

It now remains to prove that the diagram (1.2.2) with A, m and u replaced by C^*, $\mathfrak{m}$ and $\mathfrak{u}$ is commutative. In other words, it remains to prove that the diagram

$$\begin{tikzcd}
C^*\otimes \mathbf{k} \arrow[d, "\mathrm{id}\otimes \mathfrak{u}"'] & C^* \arrow[l] \arrow[r] \arrow[d, "\mathrm{id}"'] & \mathbf{k}\otimes C^* \arrow[d, "\mathfrak{u}\otimes \mathrm{id}"] \\
C^*\otimes C^* \arrow[r, "\mathfrak{m}"] & C^* & C^*\otimes C^* \arrow[l, "\mathfrak{m}"']
\end{tikzcd}$$

is commutative. We will only prove the commutativity of the left square of this diagram (since the right square is analogous). That is, we will only prove the commutativity of the square

$$\begin{tikzcd}
C^*\otimes \mathbf{k} \arrow[d, "\mathrm{id}\otimes \mathfrak{u}"'] & C^* \arrow[l] \arrow[d, "\mathrm{id}"] \\
C^*\otimes C^* \arrow[r, "\mathfrak{m}"] & C^*
\end{tikzcd} . \tag{13.34.3}$$

Indeed, the proof is similar to our proof of (13.34.1), but instead of the big diagram (13.34.2) we now have the diagram

$$\begin{tikzcd}
C^*\otimes \mathbf{k} \arrow[d, "\mathrm{id}\otimes s"] \arrow[dd, bend right=60, "\mathrm{id}\otimes \mathfrak{u}"'] & & C^* \arrow[ll, "\cong"'] \arrow[dl] \arrow[dd, "\mathrm{id}^*"] \arrow[dd, bend left=60, "\mathrm{id}"] \\
C^*\otimes \mathbf{k}^* \arrow[r, "\rho_{C,\mathbf{k}}"] \arrow[d, "\mathrm{id}\otimes \epsilon_C^*"] & (C\otimes \mathbf{k})^* \arrow[d, "(\mathrm{id}\otimes \epsilon_C)^*"] & \\
C^*\otimes C^* \arrow[r, "\rho_{C,C}"] \arrow[rr, bend right=40, "\mathfrak{m}"'] & (C\otimes C)^* \arrow[r, "\Delta_C^*"] & C^*
\end{tikzcd} \tag{13.34.4}$$

in which the arrow $C^* \to (C\otimes \mathbf{k})^*$ is the adjoint map of the canonical isomorphism $C\otimes \mathbf{k} \to C$ (and in which we identify $\mathbf{k}^*$ with $\mathbf{k}$). The commutativity of the little triangles and squares is again easily proven (the square

$$\begin{tikzcd}
& C^* \arrow[dl] \arrow[dd, "\mathrm{id}^*"] \\
(C\otimes \mathbf{k})^* \arrow[d, "(\mathrm{id}\otimes \epsilon_C)^*"'] & \\
(C\otimes C)^* \arrow[r, "\Delta_C^*"] & C^*
\end{tikzcd}$$

is commutative by virtue of being the dual of the left square in (1.2.2)). The commutativity of the "2-gon"

$$\begin{tikzcd} C^* \arrow[d, "\mathrm{id}^*"] \arrow[d, bend left=80, "\mathrm{id}"] \\ C^* \end{tikzcd}$$

simply says that $\mathrm{id}^* = \mathrm{id}$, which is obvious. Thus, everything in (13.34.4) commutes. By following the "outer quadrilateral" of (13.34.4), we obtain precisely the commutativity of (13.34.3). This completes our solution of Exercise 1.6.1 (a).

Remark: Our solution was not the simplest one (by far). We could have saved much work by doing certain abuses of notation (such as identifying s with the identity map). We could also solve part (a) very easily if we had solved part (b) first. In the above solution, we have avoided all such shortcuts.

(b) Let C be a $\mathbf{k}$-coalgebra. Let us notice that

$$\rho_{C,C}(f \otimes g) = m_{\mathbf{k}} \circ (f \otimes g) \qquad \text{for all } f \in C^* \text{ and } g \in C^*. \tag{13.34.5}$$

[452]

Now, we want to prove that the $\mathbf{k}$-algebra structure defined on C^* in part (a) is precisely the one defined on $\operatorname{Hom}(C, \mathbf{k}) = C^*$ in Definition 1.4.1 applied to $A = \mathbf{k}$. In order to do this, it is clearly enough to show that the product of any two elements of C^* with respect to the former $\mathbf{k}$-algebra structure equals their product with respect to the latter $\mathbf{k}$-algebra structure. In other words, it is enough to prove that for any $f \in C^*$ and $g \in C^*$, we have

$$\begin{aligned} &(\text{the product } fg \text{ with respect to the } \mathbf{k}\text{-algebra } C^* \text{ defined in part (a)}) \\ &= (\text{the product } fg \text{ with respect to the } \mathbf{k}\text{-algebra } \operatorname{Hom}(C,k) \text{ defined in Definition 1.4.1 applied to } A = \mathbf{k}). \end{aligned}$$

But this follows from the following computation:

$$\begin{aligned} &(\text{the product } fg \text{ with respect to the } \mathbf{k}\text{-algebra } C^* \text{ defined in part (a)}) \\ &= \underbrace{m_{C^*}}_{=\Delta_C^* \circ \rho_{C,C}} (f \otimes g) = (\Delta_C^* \circ \rho_{C,C})(f \otimes g) = \Delta_C^* \left(\underbrace{\rho_{C,C}(f \otimes g)}_{\substack{= m_{\mathbf{k}} \circ (f \otimes g) \\ (\text{by (13.34.5)})}} \right) \\ &= \Delta_C^* (m_{\mathbf{k}} \circ (f \otimes g)) = m_{\mathbf{k}} \circ (f \otimes g) \circ \Delta_C \qquad (\text{by the definition of } \Delta_C^*) \\ &= f \star g \qquad (\text{since } f \star g = m_{\mathbf{k}} \circ (f \otimes g) \circ \Delta_C \text{ (by the definition of } f \star g)) \\ &= (\text{the product } fg \text{ with respect to the } \mathbf{k}\text{-algebra } \operatorname{Hom}(C,k) \text{ defined in Definition 1.4.1 applied to } A = \mathbf{k}). \end{aligned}$$

Thus, part (b) of the exercise is solved.

(c) Let C be a graded $\mathbf{k}$-coalgebra. Let $C = \bigoplus_{n \geq 0} C_n$ be its decomposition into homogeneous components. Then, for every $n \in \mathbb{N}$, we identify $(C_n)^*$ with a $\mathbf{k}$-submodule of C^*, namely with the $\mathbf{k}$-submodule $\{f \in C^* \mid f(C_p) = 0 \text{ for all } p \in \mathbb{N} \text{ satisfying } p \neq n\}$. Thus,

$$(C_n)^* = \{f \in C^* \mid f(C_p) = 0 \text{ for all } p \in \mathbb{N} \text{ satisfying } p \neq n\} \tag{13.34.6}$$

for every $n \in \mathbb{N}$.

By the definition of C^o, we have $C^o = \bigoplus_{n \geq 0} (C_n)^* = \sum_{n \geq 0} (C_n)^*$.

Now let a and b be two nonnegative integers. We shall show that $(C_a)^* (C_b)^* \subset (C_{a+b})^*$.

Indeed, let $x \in (C_a)^*$ and $y \in (C_b)^*$ be arbitrary. We have

$$x \in (C_a)^* = \{f \in C^* \mid f(C_p) = 0 \text{ for all } p \in \mathbb{N} \text{ satisfying } p \neq a\}$$

(by (13.34.6), applied to $n = a$). In other words, x is an element of C^* such that

$$x(C_p) = 0 \text{ for all } p \in \mathbb{N} \text{ satisfying } p \neq a. \tag{13.34.7}$$

[452] *Proof of* (13.34.5)*:* Let $f \in C^*$ and $g \in C^*$. Then, the definition of $\rho_{C,C}$ shows that $\rho_{C,C}(f \otimes g)$ is the composition $C \otimes C \xrightarrow{f \otimes g} \mathbf{k} \otimes \mathbf{k} \xrightarrow{m_{\mathbf{k}}} \mathbf{k}$ of the map $f \otimes g$ with the canonical isomorphism $\mathbf{k} \otimes \mathbf{k} \xrightarrow{m_{\mathbf{k}}} \mathbf{k}$. In other words, $\rho_{C,C}(f \otimes g) = m_{\mathbf{k}} \circ (f \otimes g)$. This proves (13.34.5).

Similarly, y is an element of C^* such that

$$y\left(C_p\right) = 0 \text{ for all } p \in \mathbb{N} \text{ satisfying } p \neq b. \tag{13.34.8}$$

Also, (13.34.6) (applied to $n = a + b$) yields

$$\left(C_{a+b}\right)^* = \left\{f \in C^* \ \mid \ f\left(C_p\right) = 0 \text{ for all } p \in \mathbb{N} \text{ satisfying } p \neq a + b\right\}. \tag{13.34.9}$$

Now, part (b) of this exercise yields that the $\mathbf{k}$-algebra structure defined on C^* in part (a) is precisely the one defined on $\operatorname{Hom}\left(C, \mathbf{k}\right) = C^*$ in Definition 1.4.1 applied to $A = \mathbf{k}$. Thus, the product of the $\mathbf{k}$-algebra C^* is precisely the convolution product $\star$ on $\operatorname{Hom}\left(C, \mathbf{k}\right) = C^*$. Hence,

$$xy = x \star y = m_{\mathbf{k}} \circ \left(x \otimes y\right) \circ \Delta_C.$$

Now, fix $p \in \mathbb{N}$ such that $p \neq a + b$. Then, $\Delta_C\left(C_p\right) \subset \sum_{q=0}^{p} C_q \otimes C_{p-q}$ (since the coalgebra C is graded). Now, every $q \in \{0, 1, ..., p\}$ satisfies

$$\left(x \otimes y\right)\left(C_q \otimes C_{p-q}\right) = 0. \tag{13.34.10}$$

(*Proof of* (13.34.10)*:* Let $q \in \{0, 1, ..., p\}$. Then we must have either $q \neq a$ or $p - q \neq b$ (or both), because otherwise we would have $p = \underbrace{q}_{=a} + \underbrace{\left(p - q\right)}_{=b} = a + b$ which would contradict $p \neq a + b$. If $q \neq a$, then $x\left(C_q\right) = 0$ (by (13.34.7)) and therefore $\left(x \otimes y\right)\left(C_q \otimes C_{p-q}\right) = \underbrace{x\left(C_q\right)}_{=0} \otimes y\left(C_{p-q}\right) = 0$. If $p - q \neq b$, then (13.34.8) yields $y\left(C_{p-q}\right) = 0$, and thus $\left(x \otimes y\right)\left(C_q \otimes C_{p-q}\right) = x\left(C_q\right) \otimes \underbrace{y\left(C_{p-q}\right)}_{=0} = 0$. Hence, in either case, we have $\left(x \otimes y\right)\left(C_q \otimes C_{p-q}\right) = 0$, and so (13.34.10) is proven.)

Now,

$$\begin{aligned}
\left(\underbrace{xy}_{=m_{\mathbf{k}} \circ (x \otimes y) \circ \Delta_C}\right)\left(C_p\right) &= \left(m_{\mathbf{k}} \circ \left(x \otimes y\right) \circ \Delta_C\right)\left(C_p\right) = m_{\mathbf{k}}\left(\left(x \otimes y\right)\left(\underbrace{\Delta_C\left(C_p\right)}_{\subset \sum_{q=0}^{p} C_q \otimes C_{p-q}}\right)\right) \\
&\subset m_{\mathbf{k}}\left(\left(x \otimes y\right)\left(\sum_{q=0}^{p} C_q \otimes C_{p-q}\right)\right) = \sum_{q=0}^{p} m_{\mathbf{k}}\left(\underbrace{\left(x \otimes y\right)\left(C_q \otimes C_{p-q}\right)}_{\substack{=0 \\ \text{(by (13.34.10))}}}\right) \\
&= \sum_{q=0}^{p} m_{\mathbf{k}} 0 = 0.
\end{aligned}$$

In other words, $\left(xy\right)\left(C_p\right) = 0$.

Now, forget that we fixed p. We thus have shown that $\left(xy\right)\left(C_p\right) = 0$ for all $p \in \mathbb{N}$ satisfying $p \neq a + b$. In other words,

$$xy \in \left\{f \in C^* \ \mid \ f\left(C_p\right) = 0 \text{ for all } p \in \mathbb{N} \text{ satisfying } p \neq a + b\right\} = \left(C_{a+b}\right)^*$$

(by (13.34.9)).

Now, forget that we fixed x and y. We thus have shown that $xy \in \left(C_{a+b}\right)^*$ for all $x \in \left(C_a\right)^*$ and $y \in \left(C_b\right)^*$. This yields $\left(C_a\right)^* \left(C_b\right)^* \subset \left(C_{a+b}\right)^*$ (since $\left(C_{a+b}\right)^*$ is a $\mathbf{k}$-module).

But this has been proven for all $a \in \mathbb{N}$ and $b \in \mathbb{N}$. From this, it is easy to conclude that $C^o C^o \subset C^o$ (since $C^o = \sum_{n \geq 0} \left(C_n\right)^*$). Hence, the $\mathbf{k}$-submodule C^o of C^* is closed under multiplication. Since we also have $1_{C^*} \in C^o$ (in fact, it is very easy to see that $1_{C^*} = \epsilon_C \in \left(C_0\right)^* \subset \sum_{n \geq 0} \left(C_n\right)^* = C^o$), this shows that C^o is a $\mathbf{k}$-subalgebra of C^*. This solves part (c) of the exercise.

(d) Let C and D be two $\mathbf{k}$-coalgebras. Let $f : C \to D$ be a homomorphism of $\mathbf{k}$-coalgebras. Then, the two diagrams

(13.34.11)

$$\begin{tikzcd}
C \arrow[r, "f"] \arrow[d, "\Delta_C"'] & D \arrow[d, "\Delta_D"] \\
C \otimes C \arrow[r, "f \otimes f"] & D \otimes D
\end{tikzcd}
\qquad \text{and} \qquad
\begin{tikzcd}
C \arrow[rr, "f"] \arrow[dr, "\epsilon_C"'] & & D \arrow[dl, "\epsilon_D"] \\
& \mathbf{k} &
\end{tikzcd}$$

are commutative (since f is a homomorphism of $\mathbf{k}$-coalgebras).

We need to prove that $f^* : D^* \to C^*$ is a homomorphism of $\mathbf{k}$-algebras. In other words, we need to prove that the two diagrams

(13.34.12)
$$\begin{tikzcd} D^* \arrow[r, "f^*"] & C^* \\ D^* \otimes D^* \arrow[u, "m_{D^*}"] \arrow[r, "f^* \otimes f^*"] & C^* \otimes C^* \arrow[u, "m_{C^*}"'] \end{tikzcd} \qquad \text{and} \qquad \begin{tikzcd} D^* \arrow[rr, "f^*"] & & C^* \\ & \mathbf{k} \arrow[ul, "u_{D^*}"] \arrow[ur, "u_{C^*}"'] & \end{tikzcd}$$

are commutative.

By the definition of the $\mathbf{k}$-algebra D^*, we have $m_{D^*} = \Delta_D^* \circ \rho_{D,D}$. Thus, the left triangle (with two vertical edges and one curved edge) in the diagram

(13.34.13)
$$\begin{tikzcd} D^* \arrow[r, "f^*"] & C^* \\ \left(D \otimes D\right)^* \arrow[u, "\Delta_D^*"] \arrow[r, "\left(f \otimes f\right)^*"] & \left(C \otimes C\right)^* \arrow[u, "\Delta_C^*"'] \\ D^* \otimes D^* \arrow[u, "\rho_{D,D}"] \arrow[r, "f^* \otimes f^*"] \arrow[uu, bend left=80, "m_{D^*}"] & C^* \otimes C^* \arrow[u, "\rho_{C,C}"'] \arrow[uu, bend right=80, "m_{C^*}"'] \end{tikzcd}$$

commutes. Similarly, the right triangle in (13.34.13) also commutes. The upper rectangle commutes because it is obtained from the first diagram in (13.34.11) by dualizing. The lower triangle in (13.34.13) commutes by basic linear algebra. Hence, the whole diagram (13.34.13) commutes. But the outer rim of the diagram (13.34.13) is exactly the first diagram in (13.34.12). Thus, the first diagram in (13.34.12) commutes. The (even simpler) task of proving the commutativity of the second diagram in (13.34.12) is left to the reader. Thus, f^* is a homomorphism of $\mathbf{k}$-algebras. Part (d) of the exercise is solved.

(e) This is an exercise in linear algebra (it has nothing to do with Hopf algebras). Here is a rough sketch of how it is solved: Let $U = \bigoplus_{n \geq 0} U_n$ be the decomposition of U into its homogeneous components, and let $V = \bigoplus_{n \geq 0} V_n$ be the decomposition of V into its homogeneous components. Notice that every V_n is finite free (since V is of finite type). Graded $\mathbf{k}$-linear maps $U \to V$ can be regarded as elements of $\prod_{n \geq 0} \operatorname{Hom}\left(U_n, V_n\right)$ (because a graded $\mathbf{k}$-linear map $U \to V$ restricts to a $\mathbf{k}$-linear map $U_n \to V_n$ for every $n \in \mathbb{N}$, and is uniquely determined by the totality of these restrictions). Similarly, graded $\mathbf{k}$-linear maps $V^o \to U^o$ can be regarded as elements of $\prod_{n \geq 0} \operatorname{Hom}\left(\left(V_n\right)^*, \left(U_n\right)^*\right)$. For every $n \in \mathbb{N}$, there is a canonical $\mathbf{k}$-module isomorphism $\operatorname{Hom}\left(U_n, V_n\right) \to \operatorname{Hom}\left(\left(V_n\right)^*, \left(U_n\right)^*\right)$ which sends every $\varphi \in \operatorname{Hom}\left(U_n, V_n\right)$ to $\varphi^* \in \operatorname{Hom}\left(\left(V_n\right)^*, \left(U_n\right)^*\right)$ (since V_n is finite free). Taking the product of these isomorphisms, we obtain a $\mathbf{k}$-module isomorphism from $\prod_{n \geq 0} \operatorname{Hom}\left(U_n, V_n\right)$ to $\prod_{n \geq 0} \operatorname{Hom}\left(\left(V_n\right)^*, \left(U_n\right)^*\right)$, that is, a $\mathbf{k}$-module isomorphism from the $\mathbf{k}$-module of all graded $\mathbf{k}$-linear maps $U \to V$ to the $\mathbf{k}$-module of all graded $\mathbf{k}$-linear maps $V^o \to U^o$. It is easy to see that this isomorphism sends every $f : U \to V$ to $f^* : V^o \to U^o$. Hence, there is a 1-to-1 correspondence between graded $\mathbf{k}$-linear maps $U \to V$ and graded $\mathbf{k}$-linear maps $V^o \to U^o$ given by $f \mapsto f^*$ (namely, this isomorphism). This solves Exercise 1.6.1 (e).

(f) Let $f : C \to D$ be a graded $\mathbf{k}$-linear map. We need to prove that $f : C \to D$ is a $\mathbf{k}$-coalgebra morphism if and only if $f^* : D^o \to C^o$ is a $\mathbf{k}$-algebra morphism. In other words, we need to prove the following two assertions:

> *Assertion 1:* If $f : C \to D$ is a $\mathbf{k}$-coalgebra morphism, then $f^* : D^o \to C^o$ is a $\mathbf{k}$-algebra morphism.
>
> *Assertion 2:* If $f^* : D^o \to C^o$ is a $\mathbf{k}$-algebra morphism, then $f : C \to D$ is a $\mathbf{k}$-coalgebra morphism.

We start by proving Assertion 1. One way to prove it proceeds by repeating the solution of Exercise 1.6.1 (d), except that D^*, $\left(D \otimes D\right)^*$, C^* and $\left(C \otimes C\right)^*$ are replaced by D^o, $\left(D \otimes D\right)^o$, C^o and $\left(C \otimes C\right)^o$ (where, of course, the map $\rho_{D,D} : D^o \to \left(D \otimes D\right)^o$ has to be interpreted as the restriction of the map $\rho_{D,D} : D^* \to \left(D \otimes D\right)^*$ to the $\mathbf{k}$-submodule D^o of D^*, and similarly for the other maps). This replacement can be done completely robotically, and thus is left to the reader. Another way to prove Assertion 1 is by realizing that it follows immediately from Exercise 1.6.1 (d) (since $f^* : D^o \to C^o$ is a restriction of the map $f^* : D^* \to C^*$).

Either way, we do not end up using the assumption that D is of finite type. However, we will need this assumption in our proof of Assertion 2.

Now, let us prove Assertion 2. Assume that $f^* : D^o \to C^o$ is a **k**-algebra morphism. We want to show that $f : C \to D$ is a **k**-coalgebra morphism. In other words, we want to show that the two diagrams (13.34.11) commute. Let us start with the left one of these diagrams.

The graded **k**-module D is of finite type, and therefore the map $\rho_{D,D} : D^o \otimes D^o \to (D \otimes D)^o$ (a restriction of the map $\rho_{D,D} : D^* \otimes D^* \to (D \otimes D)^*$) is an isomorphism. Its inverse $\rho_{D,D}^{-1} : (D \otimes D)^o \to D^o \otimes D^o$ is therefore well-defined[453]. We can thus form the (asymmetric!) diagram

(13.34.14)
$$\begin{tikzcd}
D^o \arrow[rrr, "f^*"] & & & C^o \\
& D^o \otimes D^o \arrow[ul, "m_{D^*}"'] \arrow[r, "f^* \otimes f^*"] & C^o \otimes C^o \arrow[ur, "m_{C^*}"] \arrow[dr, "\rho_{C,C}"] & \\
(D \otimes D)^o \arrow[uu, "\Delta_D^*"] \arrow[ur, "\rho_{D,D}^{-1}"] \arrow[rrr, "(f \otimes f)^*"'] & & & (C \otimes C)^o \arrow[uu, "\Delta_C^*"']
\end{tikzcd} .$$

(The arrows labelled m_{C^*} and m_{D^*} could just as well have been labelled m_{C^o} and m_{D^o}, since the multiplication maps m_{C^o} and m_{D^o} are restrictions of m_{C^*} and m_{D^*}.) The two triangles in (13.34.14) commute due to $m_{D^*} = \Delta_D^* \circ \rho_{D,D}$ and $m_{C^*} = \Delta_C^* \circ \rho_{C,C}$. The upper quadrilateral in (13.34.14) commutes because f^* is a **k**-algebra homomorphism, and the lower quadrilateral in (13.34.14) commutes because of the commutativity of the diagram

$$\begin{tikzcd}
D^o \otimes D^o \arrow[r, "f^* \otimes f^*"] \arrow[d, "\rho_{D,D}"] & C^o \otimes C^o \arrow[d, "\rho_{C,C}"] \\
(D \otimes D)^o \arrow[r, "(f \otimes f)^*"'] & (C \otimes C)^o
\end{tikzcd}$$

(which follows from standard linear algebra). Hence, the whole diagram (13.34.14) commutes. Removing the two interior nodes of this diagram, we obtain the commutative diagram

(13.34.15)
$$\begin{tikzcd}
D^o \arrow[r, "f^*"] & C^o \\
(D \otimes D)^o \arrow[u, "\Delta_D^*"] \arrow[r, "(f \otimes f)^*"'] & (C \otimes C)^o \arrow[u, "\Delta_C^*"']
\end{tikzcd} .$$

This does not immediately yield the commutativity of the first diagram in (13.34.11) (because we cannot revert taking dual **k**-modules), so we are not yet done. But we are close.

We have

$$\begin{aligned}
(\Delta_D \circ f)^* = f^* \circ \Delta_D^* = \Delta_C^* \circ (f \otimes f)^* \qquad & \text{(since the diagram (13.34.15) commutes)} \\
= ((f \otimes f) \circ \Delta_C)^* &
\end{aligned}$$

as maps from $(D \otimes D)^o$ to C^o. But a general linear-algebraic fact states that if U and V are two graded **k**-modules such that V is of finite type, and if α and β are two graded **k**-linear maps $U \to V$ such that $\alpha^* = \beta^*$ as maps from V^o to U^o, then $\alpha = \beta$ [454]. Applying this to $U = C$, $V = D \otimes D$, $\alpha = \Delta_D \circ f$ and $\beta = (f \otimes f) \circ \Delta_C$, we obtain $\Delta_D \circ f = (f \otimes f) \circ \Delta_C$ (since $D \otimes D$ is of finite type[455] and since $(\Delta_D \circ f)^* = ((f \otimes f) \circ \Delta_C)^*$). In other words, the first diagram in (13.34.11) is commutative. The reader can verify that so is the second diagram in (13.34.11) (once again, this is the easier part). So we have shown that both diagrams in (13.34.11) are commutative, and thus f is a **k**-coalgebra morphism. This proves Assertion 2.

Now that Assertions 1 and 2 are both proven, part (f) of the exercise is solved.

[453] Beware: we don't have an inverse of the non-restricted map $\rho_{D,D} : D^* \otimes D^* \to (D \otimes D)^*$.

[454] This follows immediately from Exercise 1.6.1 (e).

[455] This is because D is of finite type.

13.35. **Solution to Exercise 1.6.4.** *Solution to Exercise 1.6.4.* We have $\operatorname{Sym}(V) \cong \mathbf{k}[x]$. Thus, $\left(x^k\right)_{k\in\mathbb{N}}$ is a basis of the $\mathbf{k}$-module $\operatorname{Sym}(V)$. Therefore, $\left(x^k \otimes x^\ell\right)_{(k,\ell)\in\mathbb{N}^2}$ is a basis of the $\mathbf{k}$-module $\operatorname{Sym}(V) \otimes \operatorname{Sym}(V)$.

On the other hand, $\left(f^{(k)}\right)_{k\in\mathbb{N}}$ is a basis of the $\mathbf{k}$-module $(\operatorname{Sym}(V))^o$ (namely, the dual basis to the graded basis $\left(x^k\right)_{k\in\mathbb{N}}$ of $\operatorname{Sym}(V)$). Hence, $\left(f^{(k)} \otimes f^{(\ell)}\right)_{(k,\ell)\in\mathbb{N}^2}$ is a basis of the $\mathbf{k}$-module $(\operatorname{Sym}(V))^o \otimes (\operatorname{Sym}(V))^o$.

We denote by m the multiplication map of $\operatorname{Sym}(V)$. We denote by u the unity map of $\operatorname{Sym}(V)$ (that is, the map $\mathbf{k} \to \operatorname{Sym}(V)$ which sends $1_{\mathbf{k}}$ to $1_{\operatorname{Sym}(V)}$). We denote by Δ the comultiplication of $\operatorname{Sym}(V)$. We denote by ϵ the counit of $\operatorname{Sym}(V)$. We denote by S the antipode of $\operatorname{Sym}(V)$.

We denote by $m_{(\operatorname{Sym}(V))^o}$, $u_{(\operatorname{Sym}(V))^o}$, $\Delta_{(\operatorname{Sym}(V))^o}$, $\epsilon_{(\operatorname{Sym}(V))^o}$ and $S_{(\operatorname{Sym}(V))^o}$ the maps analogous to m, u, Δ, ϵ and S but defined for the Hopf algebra $(\operatorname{Sym}(V))^o$ instead of $\operatorname{Sym}(V)$.

(a) Clearly, $x^i \cdot x^j = x^{i+j}$ for all $i \in \mathbb{N}$ and $j \in \mathbb{N}$.

We have $\Delta_{T(V)}(x) = 1 \otimes x + x \otimes 1$ in $T(V) \otimes T(V)$ (by the definition of $\Delta_{T(V)}$). Projecting this equality down onto $\operatorname{Sym}(V) \otimes \operatorname{Sym}(V)$, we obtain $\Delta_{\operatorname{Sym}(V)}(x) = x \otimes 1 + 1 \otimes x$ (since $\operatorname{Sym}(V)$ is a quotient Hopf algebra of $T(V)$). But we defined Δ to mean the comultiplication of $\operatorname{Sym}(V)$. We thus have $\Delta = \Delta_{\operatorname{Sym}(V)}$, so that $\Delta(x) = \Delta_{\operatorname{Sym}(V)}(x) = x \otimes 1 + 1 \otimes x$. Thus, x is a primitive element of $\operatorname{Sym}(V)$. Proposition 1.4.17 thus yields $S(x) = -x$.

Proposition 1.4.10 yields that the antipode S of $\operatorname{Sym}(V)$ is an algebra anti-endomorphism of $\operatorname{Sym}(V)$. Since an algebra anti-endomorphism of $\operatorname{Sym}(V)$ means the same as an algebra endomorphism of $\operatorname{Sym}(V)$ (by Exercise 1.5.8(a), because $\operatorname{Sym}(V)$ is commutative), this yields that S is an algebra endomorphism of $\operatorname{Sym}(V)$. Consequently, for any $n \in \mathbb{N}$, we have

$$S(x^n) = \Bigg(\underbrace{S(x)}_{=-x}\Bigg)^n = (-x)^n = (-1)^n x^n. \tag{13.35.1}$$

It remains to prove that $\Delta(x^n) = \sum_{i+j=n} \dbinom{n}{i} x^i \otimes x^j$ for every $n \in \mathbb{N}$ (where the summation sign "$\sum_{i+j=n}$" is shorthand for "$\sum\limits_{\substack{(i,j)\in\mathbb{N}^2;\\ i+j=n}}$"). Let $n \in \mathbb{N}$. The $\mathbf{k}$-algebra $\operatorname{Sym}(V) \otimes \operatorname{Sym}(V)$ is commutative (since $\operatorname{Sym}(V)$ is commutative), and thus the binomial formula can be applied in it. We know that $\operatorname{Sym}(V)$ is a bialgebra, and thus Δ is a $\mathbf{k}$-algebra homomorphism (by the axioms of a bialgebra). Hence,

$$\begin{aligned}
\Delta(x^n) &= \Bigg(\underbrace{\Delta(x)}_{=x\otimes 1+1\otimes x}\Bigg)^n = (x \otimes 1 + 1 \otimes x)^n \\
&= \sum_{\ell=0}^{n} \binom{n}{\ell} \underbrace{(x \otimes 1)^\ell}_{=x^\ell \otimes 1} \underbrace{(1 \otimes x)^{n-\ell}}_{=1\otimes x^{n-\ell}} \qquad \text{(by the binomial formula)} \\
&= \sum_{\ell=0}^{n} \binom{n}{\ell} \underbrace{\left(x^\ell \otimes 1\right)\left(1 \otimes x^{n-\ell}\right)}_{=(x^\ell 1)\otimes(1x^{n-\ell})=x^\ell\otimes x^{n-\ell}} = \sum_{\ell=0}^{n} \binom{n}{\ell} x^\ell \otimes x^{n-\ell} = \sum_{i+j=n} \binom{n}{i} x^i \otimes x^j
\end{aligned} \tag{13.35.2}$$

(here, we have substituted (i,j) for $(\ell, n-\ell)$ in the sum). This completes the solution of Exercise 1.6.4(a).

(b) The definition of $f^{(i)}$ shows that

$$f^{(i)}\left(x^j\right) = \delta_{i,j} \qquad \text{for all } i \in \mathbb{N} \text{ and } j \in \mathbb{N}. \tag{13.35.3}$$

But the definition of the $\mathbf{k}$-algebra $(\operatorname{Sym}(V))^o$ shows that the multiplication map $m_{(\operatorname{Sym}(V))^o}$ of the $\mathbf{k}$-algebra $(\operatorname{Sym}(V))^o$ is adjoint to the comultiplication map Δ of the $\mathbf{k}$-coalgebra $\operatorname{Sym}(V)$. In other words, we have

$$\left(m_{(\operatorname{Sym}(V))^o}(a), b\right)_{\operatorname{Sym}(V)} = (a, \Delta(b))_{\operatorname{Sym}(V)\otimes\operatorname{Sym}(V)} \tag{13.35.4}$$

for any $a \in (\operatorname{Sym}(V))^o \otimes (\operatorname{Sym}(V))^o$ and any $b \in \operatorname{Sym}(V)$. Thus, any $p \in (\operatorname{Sym}(V))^o$, $q \in (\operatorname{Sym}(V))^o$ and $b \in \operatorname{Sym}(V)$ satisfy

$$(pq, b)_{\operatorname{Sym}(V)} = (p \otimes q, \Delta(b))_{\operatorname{Sym}(V) \otimes \operatorname{Sym}(V)} . \tag{13.35.5}$$

[456]

On the other hand, the definition of the **k**-coalgebra $(\operatorname{Sym}(V))^o$ shows that the comultiplication map $\Delta_{(\operatorname{Sym}(V))^o}$ of the **k**-coalgebra $(\operatorname{Sym}(V))^o$ is adjoint to the multiplication map m of the **k**-algebra $\operatorname{Sym}(V)$. In other words, we have

$$\left(\Delta_{(\operatorname{Sym}(V))^o}(a), b\right)_{\operatorname{Sym}(V) \otimes \operatorname{Sym}(V)} = (a, m(b))_{\operatorname{Sym}(V)} \tag{13.35.6}$$

for any $a \in (\operatorname{Sym}(V))^o$ and any $b \in \operatorname{Sym}(V) \otimes \operatorname{Sym}(V)$. Thus, any $a \in (\operatorname{Sym}(V))^o$, $p \in \operatorname{Sym}(V)$ and $q \in \operatorname{Sym}(V)$ satisfy

$$\left(\Delta_{(\operatorname{Sym}(V))^o}(a), p \otimes q\right)_{\operatorname{Sym}(V) \otimes \operatorname{Sym}(V)} = (a, pq)_{\operatorname{Sym}(V)} . \tag{13.35.7}$$

[457]

The definition of the **k**-Hopf algebra $(\operatorname{Sym}(V))^o$ shows that the antipode $S_{(\operatorname{Sym}(V))^o}$ of the **k**-Hopf algebra $(\operatorname{Sym}(V))^o$ is adjoint to the antipode S of the **k**-Hopf algebra $\operatorname{Sym}(V)$. In other words, we have

$$\left(S_{(\operatorname{Sym}(V))^o}(a), b\right)_{\operatorname{Sym}(V)} = (a, S(b))_{\operatorname{Sym}(V)} \tag{13.35.8}$$

for any $a \in (\operatorname{Sym}(V))^o$ and $b \in \operatorname{Sym}(V)$.

Let us also notice a simple fact about sums: If u, v and n are three nonnegative integers, if $\mathfrak{A}$ is an additive abelian group, and if $(t_{i,j})_{(i,j) \in \mathbb{N}^2}$ is a family of elements of $\mathfrak{A}$, then

$$\sum_{i+j=n} t_{i,j} \delta_{u,i} \delta_{v,j} = t_{u,v} \delta_{u+v,n} \tag{13.35.9}$$

(where, again, "$\sum_{i+j=n}$" is shorthand for "$\sum\limits_{\substack{(i,j) \in \mathbb{N}^2; \\ i+j=n}}$"). [458]

[456] *Proof of (13.35.5):* Let $p \in (\operatorname{Sym}(V))^o$, $q \in (\operatorname{Sym}(V))^o$ and $b \in \operatorname{Sym}(V)$. Then, $m_{(\operatorname{Sym}(V))^o}(p \otimes q) = pq$ (since $m_{(\operatorname{Sym}(V))^o}$ is the multiplication map of the algebra $(\operatorname{Sym}(V))^o$). Thus,

$$\left(\underbrace{pq}_{= m_{(\operatorname{Sym}(V))^o}(p \otimes q)}, b\right)_{\operatorname{Sym}(V)} = \left(m_{(\operatorname{Sym}(V))^o}(p \otimes q), b\right)_{\operatorname{Sym}(V)} = (p \otimes q, \Delta(b))_{\operatorname{Sym}(V) \otimes \operatorname{Sym}(V)}$$

(by (13.35.4), applied to $a = p \otimes q$). This proves (13.35.5).

[457] *Proof of (13.35.7):* Let $a \in (\operatorname{Sym}(V))^o$, $p \in \operatorname{Sym}(V)$ and $q \in \operatorname{Sym}(V)$. Then, $m(p \otimes q) = pq$ (since m is the multiplication map of the algebra $\operatorname{Sym}(V)$). But (13.35.6) (applied to $b = p \otimes q$) yields

$$\left(\Delta_{(\operatorname{Sym}(V))^o}(a), p \otimes q\right)_{\operatorname{Sym}(V) \otimes \operatorname{Sym}(V)} = \left(a, \underbrace{m(p \otimes q)}_{= pq}\right)_{\operatorname{Sym}(V)} = (a, pq)_{\operatorname{Sym}(V)} .$$

This proves (13.35.7).

[458] *Proof of (13.35.9):* Let u, v and n be three nonnegative integers. Let $\mathfrak{A}$ be an additive abelian group. Let $(t_{i,j})_{(i,j) \in \mathbb{N}^2}$ be a family of elements of $\mathfrak{A}$.

Let us first assume that $u + v \neq n$. Each addend of the sum $\sum_{i+j=n} t_{i,j} \delta_{u,i} \delta_{v,j}$ contains the factor $\delta_{u,i} \delta_{v,j}$, which is 0 unless we have both $u = i$ and $v = j$. Thus, each addend of this sum vanishes unless it satisfies both $u = i$ and $v = j$ at the same time. Since no addend of this sum can satisfy both $u = i$ and $v = j$ at the same time (because such an addend would then also satisfy $\underbrace{u}_{=i} + \underbrace{v}_{=j} = i + j = n$, which would contradict $u + v \neq n$), this shows that each addend of this sum vanishes. Consequently, the sum itself must vanish. That is, we have $\sum_{i+j=n} t_{i,j} \delta_{u,i} \delta_{v,j} = 0$. Compared with $t_{u,v} \underbrace{\delta_{u+v,n}}_{\substack{=0 \\ (\text{since } u+v \neq n)}} = 0$, this yields $\sum_{i+j=n} t_{i,j} \delta_{u,i} \delta_{v,j} = t_{u,v} \delta_{u+v,n}$. Hence, (13.35.9) is proven under the assumption that $u + v \neq n$.

Therefore, for the rest of our proof of (13.35.9), we can WLOG assume that we **don't** have $u + v \neq n$. Thus, $u + v = n$. Again, each addend of the sum $\sum_{i+j=n} t_{i,j} \delta_{u,i} \delta_{v,j}$ contains the factor $\delta_{u,i} \delta_{v,j}$, which is 0 unless we have both $u = i$ and $v = j$. Thus, each addend of this sum vanishes unless it satisfies both $u = i$ and $v = j$ at the same time. But there exists exactly one addend of the sum $\sum_{i+j=n} t_{i,j} \delta_{u,i} \delta_{v,j}$ which satisfies both $u = i$ and $v = j$: namely, the addend for $(i,j) = (u,v)$ (and this is

Now, we have

$$f^{(u)} f^{(v)} = \binom{u+v}{u} f^{(u+v)} \qquad \text{for every } u \in \mathbb{N} \text{ and } v \in \mathbb{N}. \tag{13.35.10}$$

Proof of (13.35.10): Let $u \in \mathbb{N}$ and $v \in \mathbb{N}$.

Let $k \in \mathbb{N}$. Then,

$$\begin{aligned}
\left(f^{(u)} f^{(v)}\right)\left(x^k\right) &= \left(f^{(u)} f^{(v)}, x^k\right)_{\operatorname{Sym}(V)} = \left(f^{(u)} \otimes f^{(v)}, \underbrace{\Delta\left(x^k\right)}_{\substack{=\sum_{i+j=k}\binom{k}{i} x^i \otimes x^j \\ \text{(by (13.35.2), applied to } n=k)}}\right)_{\operatorname{Sym}(V)\otimes\operatorname{Sym}(V)} \\
&\qquad \left(\text{by (13.35.5), applied to } p = f^{(u)},\ q = f^{(v)} \text{ and } b = x^k\right) \\
&= \left(f^{(u)} \otimes f^{(v)}, \sum_{i+j=k} \binom{k}{i} x^i \otimes x^j\right)_{\operatorname{Sym}(V)\otimes\operatorname{Sym}(V)} \\
&= \sum_{i+j=k} \binom{k}{i} \underbrace{\left(f^{(u)} \otimes f^{(v)}, x^i \otimes x^j\right)_{\operatorname{Sym}(V)\otimes\operatorname{Sym}(V)}}_{\substack{=\left(f^{(u)}, x^i\right)_{\operatorname{Sym}(V)}\left(f^{(v)}, x^j\right)_{\operatorname{Sym}(V)} \\ \text{(by the definition of the} \\ \text{bilinear form } (\cdot,\cdot)_{\operatorname{Sym}(V)\otimes\operatorname{Sym}(V)})}} \\
&= \sum_{i+j=k} \binom{k}{i} \underbrace{\left(f^{(u)}, x^i\right)_{\operatorname{Sym}(V)}}_{\substack{=f^{(u)}\left(x^i\right)=\delta_{u,i} \\ \text{(by (13.35.3), applied} \\ \text{to } u \text{ and } i \text{ instead of } i \text{ and } j)}} \underbrace{\left(f^{(v)}, x^j\right)_{\operatorname{Sym}(V)}}_{\substack{=f^{(v)}\left(x^j\right)=\delta_{v,j} \\ \text{(by (13.35.3), applied} \\ \text{to } v \text{ instead of } i)}} \\
&= \sum_{i+j=k} \binom{k}{i} \delta_{u,i} \delta_{v,j} \\
&= \binom{k}{u} \delta_{u+v,k} \qquad \left(\text{by (13.35.9), applied to } n = k,\ \mathfrak{A} = \mathbf{k} \text{ and } t_{i,j} = \binom{k}{i}\right) \\
&= \binom{u+v}{u} \delta_{u+v,k} \qquad \left(\begin{array}{c} \text{because the equality } \binom{k}{u}\delta_{u+v,k} = \binom{u+v}{u}\delta_{u+v,k} \text{ holds} \\ \text{in the case when } u+v=k \text{ (obviously) and in the case} \\ \text{when } u+v \neq k \text{ (since both sides of this equality are 0 in} \\ \text{this case (due to } \delta_{u+v,k} = 0)) \end{array}\right).
\end{aligned}$$

Comparing this with

$$\left(\binom{u+v}{u} f^{(u+v)}\right)\left(x^k\right) = \binom{u+v}{u} \underbrace{f^{(u+v)}\left(x^k\right)}_{\substack{=\delta_{u+v,k} \\ \text{(by (13.35.3), applied} \\ \text{to } u+v \text{ and } k \text{ instead of } i \text{ and } j)}} = \binom{u+v}{u} \delta_{u+v,k},$$

indeed an addend because we have $u+v=n$). This addend equals $t_{u,v} \underbrace{\delta_{u,u}}_{=1} \underbrace{\delta_{v,v}}_{=1} = t_{u,v}$. Because of this, and because all other addends of our sum vanish, we thus conclude that the whole sum must equal $t_{u,v}$. That is, we have $\sum_{i+j=n} t_{i,j} \delta_{u,i} \delta_{v,j} = t_{u,v}$. Compared with $t_{u,v} \underbrace{\delta_{u+v,n}}_{\substack{=1 \\ (\text{since } u+v=n)}} = t_{u,v}$, this yields $\sum_{i+j=n} t_{i,j} \delta_{u,i} \delta_{v,j} = t_{u,v} \delta_{u+v,n}$. Hence, (13.35.9) is proven.

we obtain $\left(f^{(u)}f^{(v)}\right)\left(x^k\right) = \left(\dbinom{u+v}{u} f^{(u+v)}\right)\left(x^k\right)$.

Let us now forget that we fixed k. We thus have shown that $\left(f^{(u)}f^{(v)}\right)\left(x^k\right) = \left(\dbinom{u+v}{u} f^{(u+v)}\right)\left(x^k\right)$ for every $k \in \mathbb{N}$. In other words, the two $\mathbf{k}$-linear maps $f^{(u)}f^{(v)}$ and $\dbinom{u+v}{u} f^{(u+v)}$ are equal to each other on the basis $\left(x^k\right)_{k\in\mathbb{N}}$ of the $\mathbf{k}$-module $\operatorname{Sym}(V)$. Therefore, these two $\mathbf{k}$-linear maps must be identical (because if two $\mathbf{k}$-linear maps from one and the same domain are equal to each other on a basis of this domain, then these $\mathbf{k}$-linear maps must be identical). In other words, $f^{(u)}f^{(v)} = \dbinom{u+v}{u} f^{(u+v)}$. Thus, (13.35.10) is proven.

Thus, we have shown that

$$f^{(i)}f^{(j)} = \binom{i+j}{i} f^{(i+j)} \qquad \text{for every } i \in \mathbb{N} \text{ and } j \in \mathbb{N}. \tag{13.35.11}$$

(In fact, (13.35.11) follows from (13.35.10) by renaming the variables u and v as i and j.)

We shall now show that

$$\Delta_{(\operatorname{Sym}(V))^o}\left(f^{(n)}\right) = \sum_{i+j=n} f^{(i)} \otimes f^{(j)} \qquad \text{for every } n \in \mathbb{N} \tag{13.35.12}$$

(where, again, "$\sum_{i+j=n}$" is shorthand for "$\sum\limits_{\substack{(i,j)\in\mathbb{N}^2;\\ i+j=n}}$").

Proof of (13.35.12): Let $n \in \mathbb{N}$. We identify $(\operatorname{Sym}(V))^o \otimes (\operatorname{Sym}(V))^o$ with $(\operatorname{Sym}(V) \otimes \operatorname{Sym}(V))^o$ (since $\operatorname{Sym}(V)$ is a graded $\mathbf{k}$-module of finite type); thus, an element of $(\operatorname{Sym}(V))^o \otimes (\operatorname{Sym}(V))^o$ can be regarded as a $\mathbf{k}$-linear map $\operatorname{Sym}(V) \otimes \operatorname{Sym}(V) \to \mathbf{k}$. In particular, $\Delta_{(\operatorname{Sym}(V))^o}\left(f^{(n)}\right)$ thus becomes a $\mathbf{k}$-linear map $\operatorname{Sym}(V) \otimes \operatorname{Sym}(V) \to \mathbf{k}$.

Fix $(k,\ell) \in \mathbb{N}^2$. Thus, $k \in \mathbb{N}$ and $\ell \in \mathbb{N}$. Then,

$$\begin{aligned}
&\left(\Delta_{(\operatorname{Sym}(V))^o}\left(f^{(n)}\right)\right)\left(x^k \otimes x^\ell\right) \\
&= \left(\Delta_{(\operatorname{Sym}(V))^o}\left(f^{(n)}\right), x^k \otimes x^\ell\right)_{\operatorname{Sym}(V)\otimes\operatorname{Sym}(V)} \\
&= \left(f^{(n)}, \underbrace{x^k x^\ell}_{=x^{k+\ell}}\right)_{\operatorname{Sym}(V)} \qquad \left(\text{by (13.35.7), applied to } a = f^{(n)},\ p = x^k \text{ and } q = x^\ell\right) \\
&= \left(f^{(n)}, x^{k+\ell}\right)_{\operatorname{Sym}(V)} = \delta_{n,k+\ell} \qquad \left(\text{by (13.35.3), applied to } i = n \text{ and } j = k+\ell\right).
\end{aligned}$$

Compared with

$$\begin{aligned}
\left(\sum_{i+j=n} f^{(i)} \otimes f^{(j)}\right)\left(x^k \otimes x^\ell\right) &= \sum_{i+j=n} \underbrace{\left(f^{(i)} \otimes f^{(j)}\right)\left(x^k \otimes x^\ell\right)}_{=f^{(i)}\left(x^k\right)f^{(j)}\left(x^\ell\right)} = \sum_{i+j=n} \underbrace{f^{(i)}\left(x^k\right)}_{\substack{=\delta_{i,k} \\ \text{(by (13.35.3), applied} \\ \text{to } k \text{ instead of } j)}} \underbrace{f^{(j)}\left(x^\ell\right)}_{\substack{=\delta_{j,\ell} \\ \text{(by (13.35.3), applied to } j \\ \text{and } \ell \text{ instead of } i \text{ and } j)}} \\
&= \sum_{i+j=n} \underbrace{\delta_{i,k}}_{=\delta_{k,i}} \underbrace{\delta_{j,\ell}}_{=\delta_{\ell,j}} = \sum_{i+j=n} \delta_{k,i}\delta_{\ell,j} = \sum_{i+j=n} 1\delta_{k,i}\delta_{\ell,j} = 1\delta_{k+\ell,n} \\
&\qquad \left(\text{by (13.35.9), applied to } k,\ \ell,\ \mathbf{k} \text{ and } 1 \text{ instead of } u,\ v,\ \mathfrak{A} \text{ and } t_{i,j}\right) \\
&= \delta_{k+\ell,n} = \delta_{n,k+\ell},
\end{aligned}$$

this yields

$$\left(\Delta_{(\operatorname{Sym}(V))^o}\left(f^{(n)}\right)\right)\left(x^k \otimes x^\ell\right) = \left(\sum_{i+j=n} f^{(i)} \otimes f^{(j)}\right)\left(x^k \otimes x^\ell\right). \tag{13.35.13}$$

Let us now forget that we fixed (k, ℓ). We thus have shown that (13.35.13) holds for every $(k, \ell) \in \mathbb{N}^2$. In other words, the two $\mathbf{k}$-linear maps $\Delta_{(\operatorname{Sym}(V))^o}\left(f^{(n)}\right)$ and $\sum_{i+j=n} f^{(i)} \otimes f^{(j)}$ (from $\operatorname{Sym}(V) \otimes \operatorname{Sym}(V)$ to $\mathbf{k}$) are equal to each other on the basis $\left(x^k \otimes x^\ell\right)_{(k,\ell)\in\mathbb{N}^2}$ of the $\mathbf{k}$-module $\operatorname{Sym}(V) \otimes \operatorname{Sym}(V)$. Therefore, these two $\mathbf{k}$-linear maps must be identical (because if two $\mathbf{k}$-linear maps from one and the same domain are equal to each other on a basis of this domain, then these $\mathbf{k}$-linear maps must be identical). In other words, $\Delta_{(\operatorname{Sym}(V))^o}\left(f^{(n)}\right) = \sum_{i+j=n} f^{(i)} \otimes f^{(j)}$. Thus, (13.35.12) is proven.

Finally, let us show that

$$S_{(\operatorname{Sym}(V))^o}\left(f^{(n)}\right) = (-1)^n f^{(n)} \qquad \text{for every } n \in \mathbb{N}. \tag{13.35.14}$$

Proof of (13.35.14): Let $n \in \mathbb{N}$. Let $k \in \mathbb{N}$. Then,

$$\begin{aligned}
\left(S_{(\operatorname{Sym}(V))^o}\left(f^{(n)}\right)\right)\left(x^k\right) &= \left(S_{(\operatorname{Sym}(V))^o}\left(f^{(n)}\right), x^k\right)_{\operatorname{Sym}(V)} = \left(f^{(n)}, \underbrace{S\left(x^k\right)}_{\substack{=(-1)^k x^k \\ \text{(by (13.35.1), applied} \\ \text{to } k \text{ instead of } n)}}\right)_{\operatorname{Sym}(V)} \\
&\qquad \left(\text{by (13.35.8), applied to } a = f^{(n)} \text{ and } b = x^k\right) \\
&= \left(f^{(n)}, (-1)^k x^k\right)_{\operatorname{Sym}(V)} = f^{(n)}\left((-1)^k x^k\right) \\
&= (-1)^k \underbrace{f^{(n)}\left(x^k\right)}_{\substack{=\delta_{n,k} \\ \text{(by (13.35.3), applied to} \\ i=n \text{ and } j=k)}} \qquad \left(\text{since the map } f^{(n)} \text{ is } \mathbf{k}\text{-linear}\right) \\
&= (-1)^k \delta_{n,k}.
\end{aligned}$$

Since $(-1)^k \delta_{n,k} = (-1)^n \delta_{n,k}$ [459], this becomes

$$\left(S_{(\operatorname{Sym}(V))^o}\left(f^{(n)}\right)\right)\left(x^k\right) = (-1)^n \delta_{n,k}.$$

Compared with

$$\left((-1)^n f^{(n)}\right)\left(x^k\right) = (-1)^n \underbrace{f^{(n)}\left(x^k\right)}_{\substack{=\delta_{n,k} \\ \text{(by (13.35.3), applied to} \\ i=n \text{ and } j=k)}} = (-1)^n \delta_{n,k},$$

this yields $\left(S_{(\operatorname{Sym}(V))^o}\left(f^{(n)}\right)\right)\left(x^k\right) = \left((-1)^n f^{(n)}\right)\left(x^k\right)$.

Let us now forget that we fixed k. We thus have shown that $\left(S_{(\operatorname{Sym}(V))^o}\left(f^{(n)}\right)\right)\left(x^k\right) = \left((-1)^n f^{(n)}\right)\left(x^k\right)$ for every $k \in \mathbb{N}$. In other words, the two $\mathbf{k}$-linear maps $S_{(\operatorname{Sym}(V))^o}\left(f^{(n)}\right)$ and $(-1)^n f^{(n)}$ are equal to each other on the basis $\left(x^k\right)_{k\in\mathbb{N}}$ of the $\mathbf{k}$-module $\operatorname{Sym}(V)$. Therefore, these two $\mathbf{k}$-linear maps must be identical (because if two $\mathbf{k}$-linear maps from one and the same domain are equal to each other on a basis of this domain, then these $\mathbf{k}$-linear maps must be identical). In other words, $S_{(\operatorname{Sym}(V))^o}\left(f^{(n)}\right) = (-1)^n f^{(n)}$. Thus, (13.35.14) is proven.

Now, all of the identities (13.35.11), (13.35.12) and (13.35.14) are proven; thus, Exercise 1.6.4(b) is solved.

Before we start the solution of Exercise 1.6.4(c), let us make a few more simple observations.

We have

$$\epsilon\left(x^k\right) = \delta_{k,0} \qquad \text{for every } k \in \mathbb{N}. \tag{13.35.15}$$

[459] This equality is obvious when $n = k$, and elsewise it follows from $\delta_{n,k} = 0$.

[460]

The definition of the **k**-algebra $(\operatorname{Sym}(V))^o$ shows that the unity map $u_{(\operatorname{Sym}(V))^o}$ of the **k**-algebra $(\operatorname{Sym}(V))^o$ is adjoint to the counit map ϵ of the **k**-coalgebra $\operatorname{Sym}(V)$. In other words, we have

$$\left(u_{(\operatorname{Sym}(V))^o}(a), b\right)_{\operatorname{Sym}(V)} = (a, \epsilon(b))_{\mathbf{k}} \tag{13.35.16}$$

for any $a \in \mathbf{k}$ and any $b \in \operatorname{Sym}(V)$. Thus,

$$u_{(\operatorname{Sym}(V))^o}(a) = af^{(0)} \qquad \text{for every } a \in \mathbf{k}. \tag{13.35.17}$$

[461]

The definition of the **k**-coalgebra $(\operatorname{Sym}(V))^o$ shows that the counit map $\epsilon_{(\operatorname{Sym}(V))^o}$ of the **k**-coalgebra $(\operatorname{Sym}(V))^o$ is adjoint to the unity map u of the **k**-algebra $\operatorname{Sym}(V)$. In other words, we have

$$\left(\epsilon_{(\operatorname{Sym}(V))^o}(a), b\right)_{\mathbf{k}} = (a, u(b))_{\operatorname{Sym}(V)} \tag{13.35.19}$$

for any $a \in (\operatorname{Sym}(V))^o$ and any $b \in \mathbf{k}$. Thus,

$$\epsilon_{(\operatorname{Sym}(V))^o}\left(f^{(k)}\right) = \delta_{k,0} \qquad \text{for every } k \in \mathbb{N}. \tag{13.35.20}$$

[462]

(c) For the time being, let us **not** assume that $\mathbb{Q}$ is a subring of **k**. Instead, we shall introduce a map and prove some of its properties which hold for every **k**.

We define a **k**-linear map $\Phi : \operatorname{Sym}(V) \to (\operatorname{Sym}(V))^o$ by

$$\left(\Phi\left(x^k\right) = k! f^{(k)} \qquad \text{for every } k \in \mathbb{N}\right). \tag{13.35.21}$$

[460] *Proof of (13.35.15):* Let $k \in \mathbb{N}$. Then, the counit ϵ of $\operatorname{Sym}(V)$ is a **k**-algebra homomorphism (by the axioms of a **k**-bialgebra (since $\operatorname{Sym}(V)$ is a **k**-bialgebra)). Hence, $\epsilon\left(x^k\right) = \left(\underbrace{\epsilon(x)}_{=0}\right)^k = 0^k = \begin{cases} 1, & \text{if } k = 0; \\ 0, & \text{if } k \neq 0 \end{cases} = \delta_{k,0}$. This proves (13.35.15).

[461] *Proof of (13.35.17):* Let $a \in \mathbf{k}$. Then, every $b \in \operatorname{Sym}(V)$ satisfies

$$\begin{aligned} \left(u_{(\operatorname{Sym}(V))^o}(a)\right)(b) &= \left(u_{(\operatorname{Sym}(V))^o}(a), b\right)_{\operatorname{Sym}(V)} = (a, \epsilon(b))_{\mathbf{k}} && \text{(by (13.35.16))} \\ &= a \cdot \epsilon(b) && \left(\text{by the definition of the form } (\cdot,\cdot)_{\mathbf{k}}\right). \end{aligned} \tag{13.35.18}$$

Now, every $k \in \mathbb{N}$ satisfies

$$\begin{aligned} \left(u_{(\operatorname{Sym}(V))^o}(a)\right)\left(x^k\right) &= a \cdot \underbrace{\epsilon\left(x^k\right)}_{\substack{=\delta_{k,0} \\ \text{(by (13.35.15))}}} && \left(\text{by (13.35.18), applied to } b = x^k\right) \\ &= a \cdot \underbrace{\delta_{k,0}}_{\substack{=f^{(0)}\left(x^k\right) \\ \text{(since (13.35.3) (applied to} \\ i=0 \text{ and } j=k\text{) yields } f^{(0)}\left(x^k\right)=\delta_{0,k}=\delta_{k,0}\text{)}}} = a \cdot f^{(0)}\left(x^k\right) = \left(af^{(0)}\right)\left(x^k\right). \end{aligned}$$

In other words, the two **k**-linear maps $u_{(\operatorname{Sym}(V))^o}(a)$ and $af^{(0)}$ are equal to each other on the basis $\left(x^k\right)_{k \in \mathbb{N}}$ of the **k**-module $\operatorname{Sym}(V)$. Therefore, these two **k**-linear maps must be identical (because if two **k**-linear maps from one and the same domain are equal to each other on a basis of this domain, then these **k**-linear maps must be identical). In other words, $u_{(\operatorname{Sym}(V))^o}(a) = af^{(0)}$. Thus, (13.35.17) is proven.

[462] *Proof of (13.35.20):* Let $k \in \mathbb{N}$. The map u is the unity map of the **k**-algebra $\operatorname{Sym}(V)$; thus, $u(1) = 1_{\operatorname{Sym}(V)} = x^0$.

The definition of the bilinear form $(\cdot,\cdot)_{\mathbf{k}}$ yields $\left(\epsilon_{(\operatorname{Sym}(V))^o}\left(f^{(k)}\right), 1\right)_{\mathbf{k}} = \epsilon_{(\operatorname{Sym}(V))^o}\left(f^{(k)}\right) \cdot 1 = \epsilon_{(\operatorname{Sym}(V))^o}\left(f^{(k)}\right)$, so that

$$\begin{aligned} \epsilon_{(\operatorname{Sym}(V))^o}\left(f^{(k)}\right) &= \left(\epsilon_{(\operatorname{Sym}(V))^o}\left(f^{(k)}\right), 1\right)_{\mathbf{k}} = \left(f^{(k)}, \underbrace{u(1)}_{=x^0}\right)_{\operatorname{Sym}(V)} && \left(\text{by (13.35.19), applied to } a = f^{(k)} \text{ and } b = 1\right) \\ &= \left(f^{(k)}, x^0\right)_{\operatorname{Sym}(V)} = \delta_{k,0} && \text{(by (13.35.3), applied to } i = k \text{ and } j = 0\text{)}. \end{aligned}$$

This proves (13.35.20).

(This is well-defined, since $\left(x^k\right)_{k\in\mathbb{N}}$ is a basis of the $\mathbf{k}$-module $\operatorname{Sym}(V)$.) We can now see that this $\mathbf{k}$-linear map Φ satisfies $\Phi \circ m = m_{(\operatorname{Sym}(V))^o} \circ (\Phi \otimes \Phi)$ [463] and $\Phi \circ u = u_{(\operatorname{Sym}(V))^o}$ [464]. Hence, Φ is a $\mathbf{k}$-algebra

[463] *Proof.* Fix $(k, \ell) \in \mathbb{N}^2$. Recall that m is the multiplication map of the $\mathbf{k}$-algebra $\operatorname{Sym}(V)$. Thus, $m\left(x^k \otimes x^\ell\right) = x^k x^\ell = x^{k+\ell}$. Also, $m_{(\operatorname{Sym}(V))^o}$ is the multiplication map of the $\mathbf{k}$-algebra $(\operatorname{Sym}(V))^o$. Thus, $m_{(\operatorname{Sym}(V))^o}\left(f^{(k)} \otimes f^{(\ell)}\right) = f^{(k)} f^{(\ell)} = \dbinom{k+\ell}{k} f^{(k+\ell)}$ (by (13.35.10), applied to $u = k$ and $v = \ell$). Now,

$$(\Phi \circ m)\left(x^k \otimes x^\ell\right) = \Phi\left(\underbrace{m\left(x^k \otimes x^\ell\right)}_{=x^{k+\ell}}\right) = \Phi\left(x^{k+\ell}\right) = (k+\ell)!\, f^{(k+\ell)} \qquad \left(\text{by (13.35.21), applied to } k+\ell \text{ instead of } k\right).$$

Compared with

$$\begin{aligned}
\left(m_{(\operatorname{Sym}(V))^o} \circ (\Phi \otimes \Phi)\right)\left(x^k \otimes x^\ell\right) &= m_{(\operatorname{Sym}(V))^o}\left(\underbrace{(\Phi \otimes \Phi)\left(x^k \otimes x^\ell\right)}_{=\Phi\left(x^k\right) \otimes \Phi\left(x^\ell\right)}\right) = m_{(\operatorname{Sym}(V))^o}\left(\underbrace{\Phi\left(x^k\right)}_{\substack{=k!f^{(k)} \\ \text{(by (13.35.21))}}} \otimes \underbrace{\Phi\left(x^\ell\right)}_{\substack{=\ell!f^{(\ell)} \\ \text{(by (13.35.21),} \\ \text{applied to } \ell \text{ instead of } k\text{)}}}\right) \\
&= m_{(\operatorname{Sym}(V))^o}\left(\underbrace{\left(k!f^{(k)}\right) \otimes \left(\ell! f^{(\ell)}\right)}_{=k!\ell! f^{(k)} \otimes f^{(\ell)}}\right) = m_{(\operatorname{Sym}(V))^o}\left(k!\ell! f^{(k)} \otimes f^{(\ell)}\right) \\
&= k!\ell!\, \underbrace{m_{(\operatorname{Sym}(V))^o}\left(f^{(k)} \otimes f^{(\ell)}\right)}_{=\binom{k+\ell}{k} f^{(k+\ell)}} \qquad \left(\text{since the map } m_{(\operatorname{Sym}(V))^o} \text{ is } \mathbf{k}\text{-linear}\right) \\
&= \underbrace{k!\ell!\binom{k+\ell}{k}}_{\substack{=(k+\ell)! \\ \left(\text{since } \binom{k+\ell}{k} = \frac{(k+\ell)!}{k!\ell!}\right)}} f^{(k+\ell)} = (k+\ell)!\, f^{(k+\ell)},
\end{aligned}$$

this yields $(\Phi \circ m)\left(x^k \otimes x^\ell\right) = \left(m_{(\operatorname{Sym}(V))^o} \circ (\Phi \otimes \Phi)\right)\left(x^k \otimes x^\ell\right)$.

Let us now forget that we fixed (k, ℓ). We thus have shown that $(\Phi \circ m)\left(x^k \otimes x^\ell\right) = \left(m_{(\operatorname{Sym}(V))^o} \circ (\Phi \otimes \Phi)\right)\left(x^k \otimes x^\ell\right)$ for every $(k, \ell) \in \mathbb{N}^2$. In other words, the two $\mathbf{k}$-linear maps $\Phi \circ m$ and $m_{(\operatorname{Sym}(V))^o} \circ (\Phi \otimes \Phi)$ (from $\operatorname{Sym}(V) \otimes \operatorname{Sym}(V)$ to $(\operatorname{Sym}(V))^o$) are equal to each other on the basis $\left(x^k \otimes x^\ell\right)_{(k,\ell)\in\mathbb{N}^2}$ of the $\mathbf{k}$-module $\operatorname{Sym}(V) \otimes \operatorname{Sym}(V)$. Therefore, these two $\mathbf{k}$-linear maps must be identical (because if two $\mathbf{k}$-linear maps from one and the same domain are equal to each other on a basis of this domain, then these $\mathbf{k}$-linear maps must be identical). In other words, $\Phi \circ m = m_{(\operatorname{Sym}(V))^o} \circ (\Phi \otimes \Phi)$. Qed.

[464] *Proof.* The map u is the unity map of the $\mathbf{k}$-algebra $\operatorname{Sym}(V)$; thus, $u(1) = 1_{\operatorname{Sym}(V)} = x^0$. Now, every $a \in \mathbf{k}$ satisfies

$$\begin{aligned}
(\Phi \circ u)\left(\underbrace{a}_{=a\cdot 1}\right) &= (\Phi \circ u)(a \cdot 1) = a \cdot \underbrace{(\Phi \circ u)(1)}_{=\Phi(u(1))} \qquad (\text{since the map } \Phi \circ u \text{ is } \mathbf{k}\text{-linear}) \\
&= a \cdot \Phi\left(\underbrace{u(1)}_{=x^0}\right) = a \cdot \underbrace{\Phi\left(x^0\right)}_{\substack{=0!f^{(0)} \\ \text{(by (13.35.21), applied to 0} \\ \text{instead of } k\text{)}}} = a \cdot \underbrace{0!}_{=1} f^{(0)} = af^{(0)} = u_{(\operatorname{Sym}(V))^o}(a) \qquad (\text{by (13.35.17)}).
\end{aligned}$$

Thus, $\Phi \circ u = u_{(\operatorname{Sym}(V))^o}$, qed.

homomorphism. Also, the **k**-linear map Φ satisfies $(\Phi \otimes \Phi) \circ \Delta = \Delta_{(\mathrm{Sym}(V))^o} \circ \Phi$ [465] and $\epsilon = \epsilon_{(\mathrm{Sym}(V))^o} \circ \Phi$ [466]. Hence, Φ is a **k**-coalgebra homomorphism. Thus, Φ is a **k**-bialgebra homomorphism (since Φ is both a **k**-algebra homomorphism and a **k**-coalgebra homomorphism), and therefore a Hopf algebra homomorphism (according to Corollary 1.4.27, applied to $H_1 = \mathrm{Sym}(V)$, $H_2 = (\mathrm{Sym}(V))^o$, $S_1 = S$ and $S_2 = S_{(\mathrm{Sym}(V))^o}$).

[465] *Proof.* Let $k \in \mathbb{N}$. We have

$$\begin{aligned}
((\Phi \otimes \Phi) \circ \Delta)\left(x^k\right) &= (\Phi \otimes \Phi)\left(\underbrace{\Delta\left(x^k\right)}_{\substack{=\sum_{i+j=k}\binom{k}{i}x^i\otimes x^j \\ \text{(by (13.35.2), applied to } n=k)}}\right) = (\Phi \otimes \Phi)\left(\sum_{i+j=k}\binom{k}{i}x^i \otimes x^j\right) \\
&= \sum_{i+j=k}\binom{k}{i}\underbrace{\Phi\left(x^i\right)}_{\substack{=i!f^{(i)} \\ \text{(by (13.35.21))}}} \otimes \underbrace{\Phi\left(x^j\right)}_{\substack{=j!f^{(j)} \\ \text{(by (13.35.21))}}} \qquad (\text{by the definition of } \Phi \otimes \Phi) \\
&= \sum_{i+j=k}\binom{k}{i} i! f^{(i)} \otimes j! f^{(j)} = \sum_{i+j=k}\binom{k}{i} i! \underbrace{j!}_{\substack{=(k-i)! \\ \text{(since } j=k-i \\ \text{(since } i+j=k))}} f^{(i)} \otimes f^{(j)} \\
&= \sum_{i+j=k}\underbrace{\binom{k}{i} i!\,(k-i)!}_{\substack{=k! \\ \left(\text{since } \binom{k}{i}=\dfrac{k!}{i!\,(k-i)!}\right)}} f^{(i)} \otimes f^{(j)} = k! \sum_{i+j=k} f^{(i)} \otimes f^{(j)}.
\end{aligned}$$

Compared with

$$\begin{aligned}
\left(\Delta_{(\mathrm{Sym}(V))^o} \circ \Phi\right)\left(x^k\right) &= \Delta_{(\mathrm{Sym}(V))^o}\left(\underbrace{\Phi\left(x^k\right)}_{\substack{=k!f^{(k)} \\ \text{(by (13.35.21))}}}\right) = \Delta_{(\mathrm{Sym}(V))^o}\left(k! f^{(k)}\right) = k! \underbrace{\Delta_{(\mathrm{Sym}(V))^o}\left(f^{(k)}\right)}_{\substack{=\sum_{i+j=k} f^{(i)}\otimes f^{(j)} \\ \text{(by (13.35.12), applied to } n=k)}} \\
&\qquad \left(\text{since the map } \Delta_{(\mathrm{Sym}(V))^o} \text{ is } \mathbf{k}\text{-linear}\right) \\
&= k! \sum_{i+j=k} f^{(i)} \otimes f^{(j)},
\end{aligned}$$

this yields $((\Phi \otimes \Phi) \circ \Delta)\left(x^k\right) = \left(\Delta_{(\mathrm{Sym}(V))^o} \circ \Phi\right)\left(x^k\right)$.

Let us now forget that we fixed k. We thus have shown that $((\Phi \otimes \Phi) \circ \Delta)\left(x^k\right) = \left(\Delta_{(\mathrm{Sym}(V))^o} \circ \Phi\right)\left(x^k\right)$ for every $k \in \mathbb{N}$. In other words, the two **k**-linear maps $(\Phi \otimes \Phi) \circ \Delta$ and $\Delta_{(\mathrm{Sym}(V))^o} \circ \Phi$ are equal to each other on the basis $\left(x^k\right)_{k\in\mathbb{N}}$ of the **k**-module $\mathrm{Sym}(V)$. Therefore, these two **k**-linear maps must be identical (because if two **k**-linear maps from one and the same domain are equal to each other on a basis of this domain, then these **k**-linear maps must be identical). In other words, $(\Phi \otimes \Phi) \circ \Delta = \Delta_{(\mathrm{Sym}(V))^o} \circ \Phi$. Qed.

[466] *Proof.* Let $k \in \mathbb{N}$. We have

$$\begin{aligned}
\left(\epsilon_{(\mathrm{Sym}(V))^o} \circ \Phi\right)\left(x^k\right) &= \epsilon_{(\mathrm{Sym}(V))^o}\left(\underbrace{\Phi\left(x^k\right)}_{\substack{=k!f^{(k)} \\ \text{(by (13.35.21))}}}\right) = \epsilon_{(\mathrm{Sym}(V))^o}\left(k! f^{(k)}\right) = k! \underbrace{\epsilon_{(\mathrm{Sym}(V))^o}\left(f^{(k)}\right)}_{\substack{=\delta_{k,0} \\ \text{(by (13.35.20))}}} \\
&\qquad \left(\text{since the map } \epsilon_{(\mathrm{Sym}(V))^o} \text{ is } \mathbf{k}\text{-linear}\right) \\
&= k!\delta_{k,0}.
\end{aligned}$$

But it is easy (by treating the cases $k = 0$ and $k \neq 0$ separately) to see that $k!\delta_{k,0} = \delta_{k,0}$. Hence,

$$\left(\epsilon_{(\mathrm{Sym}(V))^o} \circ \Phi\right)\left(x^k\right) = k!\delta_{k,0} = \delta_{k,0} = \epsilon\left(x^k\right) \qquad (\text{by (13.35.15)}).$$

In other words, $\epsilon\left(x^k\right) = \left(\epsilon_{(\mathrm{Sym}(V))^o} \circ \Phi\right)\left(x^k\right)$.

The map Φ is also graded[467]. Thus, Φ is a homomorphism of graded $\mathbf{k}$-Hopf algebras (since Φ is graded and a Hopf algebra homomorphism).

Now, let us assume that $\mathbb{Q}$ is a subring of $\mathbf{k}$. We define a $\mathbf{k}$-linear map $\Psi : (\operatorname{Sym}(V))^o \to \operatorname{Sym}(V)$ by

$$\left(\Psi\left(f^{(k)}\right) = \frac{x^k}{k!} \qquad \text{for every } k \in \mathbb{N}\right). \tag{13.35.22}$$

(This is well-defined, since $\left(f^{(k)}\right)_{k\in\mathbb{N}}$ is a basis of the $\mathbf{k}$-module $(\operatorname{Sym}(V))^o$.) It is clear that the maps Φ and Ψ are mutually inverse[468]. Thus, the map Φ is invertible, and its inverse is Ψ.

The map Φ is an invertible homomorphism of graded $\mathbf{k}$-Hopf algebras and therefore an isomorphism of graded $\mathbf{k}$-Hopf algebras (since every invertible homomorphism of graded $\mathbf{k}$-Hopf algebras is an isomorphism of graded $\mathbf{k}$-Hopf algebras). Hence, the inverse of Φ is also an isomorphism of graded $\mathbf{k}$-Hopf algebras. In other words, Ψ is an isomorphism of graded $\mathbf{k}$-Hopf algebras (since the inverse of Φ is Ψ).

But every $n \in \mathbb{N}$ satisfies $\Psi\left(f^{(n)}\right) = \dfrac{x^n}{n!}$ (according to (13.35.22), applied to $k = n$). Hence, Ψ is the $\mathbf{k}$-linear map $(\operatorname{Sym}(V))^o \to \operatorname{Sym}(V)$ sending $f^{(n)} \mapsto \frac{x^n}{n!}$. Thus, the $\mathbf{k}$-linear map $(\operatorname{Sym}(V))^o \to \operatorname{Sym}(V)$ sending $f^{(n)} \mapsto \frac{x^n}{n!}$ is an isomorphism of graded $\mathbf{k}$-Hopf algebras (since Ψ is an isomorphism of graded $\mathbf{k}$-Hopf algebras). This solves Exercise 1.6.4(c).

(d) For the time being, let us **not** assume that $\mathbf{k}$ is a field of characteristic $p > 0$. Instead, let us first prove a formula which does not require any restrictions on $\mathbf{k}$.

Namely, let us show that

$$\left(f^{(1)}\right)^m = m! f^{(m)} \qquad \text{for every } m \in \mathbb{N}. \tag{13.35.23}$$

Proof of (13.35.23): Let us consider the $\mathbf{k}$-linear map $\Phi : \operatorname{Sym}(V) \to (\operatorname{Sym}(V))^o$ defined in our solution to Exercise 1.6.4(c) above. Then, Φ is a $\mathbf{k}$-algebra homomorphism. (In fact, this was proven in our solution to Exercise 1.6.4(c) above.) Applying (13.35.21) to $k = 1$, we obtain $\Phi\left(x^1\right) = \underbrace{1!}_{=1} f^{(1)} = f^{(1)}$, so that

Let us now forget that we fixed k. We thus have shown that $\epsilon\left(x^k\right) = \left(\epsilon_{(\operatorname{Sym}(V))^o} \circ \Phi\right)\left(x^k\right)$ for every $k \in \mathbb{N}$. In other words, the two $\mathbf{k}$-linear maps ϵ and $\epsilon_{(\operatorname{Sym}(V))^o} \circ \Phi$ are equal to each other on the basis $\left(x^k\right)_{k\in\mathbb{N}}$ of the $\mathbf{k}$-module $\operatorname{Sym}(V)$. Therefore, these two $\mathbf{k}$-linear maps must be identical (because if two $\mathbf{k}$-linear maps from one and the same domain are equal to each other on a basis of this domain, then these $\mathbf{k}$-linear maps must be identical). In other words, $\epsilon = \epsilon_{(\operatorname{Sym}(V))^o} \circ \Phi$. Qed.

[467] *Proof.* Let $n \in \mathbb{N}$. For every graded $\mathbf{k}$-module A, let A_n denote the n-th homogeneous component of A. We shall show that $\Phi\left((\operatorname{Sym}(V))_n\right) \subset \left((\operatorname{Sym}(V))^o\right)_n$.

Indeed, recall that $\left(f^{(k)}\right)_{k\in\mathbb{N}}$ is a **graded** basis of the $\mathbf{k}$-module $(\operatorname{Sym}(V))^o$. Thus, the one-element family $\left(f^{(n)}\right)$ is a basis of the n-th homogeneous component $\left((\operatorname{Sym}(V))^o\right)_n$. Thus, this family $\left(f^{(n)}\right)$ spans the $\mathbf{k}$-module $\left((\operatorname{Sym}(V))^o\right)_n$. In other words, $\left((\operatorname{Sym}(V))^o\right)_n = \mathbf{k} \cdot f^{(n)}$.

On the other hand, recall that $\left(x^k\right)_{k\in\mathbb{N}}$ is a **graded** basis of the $\mathbf{k}$-module $\operatorname{Sym}(V)$. Thus, the one-element family (x^n) is a basis of the n-th homogeneous component $(\operatorname{Sym}(V))_n$. Thus, this family (x^n) spans the $\mathbf{k}$-module $(\operatorname{Sym}(V))_n$. In other words, $(\operatorname{Sym}(V))_n = \mathbf{k} \cdot x^n$. Hence,

$$\begin{aligned}
\Phi\left(\underbrace{(\operatorname{Sym}(V))_n}_{=\mathbf{k}\cdot x^n}\right) &= \Phi\left(\mathbf{k} \cdot x^n\right) \subset \mathbf{k} \cdot \underbrace{\Phi\left(x^n\right)}_{\substack{=n! f^{(n)} \\ \text{(by (13.35.21),} \\ \text{applied to } k=n)}} \qquad \text{(since the map } \Phi \text{ is } \mathbf{k}\text{-linear)} \\
&= \underbrace{\mathbf{k} \cdot n!}_{\subset \mathbf{k}} f^{(n)} \subset \mathbf{k} \cdot f^{(n)} = \left((\operatorname{Sym}(V))^o\right)_n \qquad \left(\text{since } \left((\operatorname{Sym}(V))^o\right)_n = \mathbf{k} \cdot f^{(n)}\right).
\end{aligned}$$

Now, let us forget that we fixed n. We thus have shown that $\Phi\left((\operatorname{Sym}(V))_n\right) \subset \left((\operatorname{Sym}(V))^o\right)_n$ for every $n \in \mathbb{N}$. In other words, the map Φ is graded, qed.

[468] Indeed, the map $\Phi \circ \Psi$ sends every $f^{(k)}$ to $f^{(k)}$ and thus is the identity map, whereas the map $\Psi \circ \Phi$ sends every x^k to x^k and therefore is the identity map as well.

$f^{(1)} = \Phi\left(x^1\right) = \Phi\left(x\right)$. Thus, for every $m \in \mathbb{N}$, we have

$$\left(\underbrace{f^{(1)}}_{=\Phi(x)}\right)^m = \left(\Phi\left(x\right)\right)^m = \Phi\left(x^m\right) \qquad \text{(since } \Phi \text{ is a } \mathbf{k}\text{-algebra homomorphism)}$$
$$= m! f^{(m)} \qquad \text{(by (13.35.21), applied to } k = m\text{)}.$$

This proves (13.35.23).

Now, let us assume that $\mathbf{k}$ is a field of characteristic $p > 0$. Then, $p \cdot 1_{\mathbf{k}} = 0$. But $\underbrace{p!}_{=(p-1)!p} \cdot 1_{\mathbf{k}} = (p-1)! \underbrace{p \cdot 1_{\mathbf{k}}}_{=0} = 0$. Applying (13.35.23) to $m = p$, we obtain

$$\left(f^{(1)}\right)^p = p! f^{(p)} = \underbrace{p! \cdot 1_{\mathbf{k}}}_{=0} \cdot f^{(p)} = 0.$$

Now, it remains to show that there can be no Hopf isomorphism $\left(\operatorname{Sym}\left(V\right)\right)^o \to \operatorname{Sym}\left(V\right)$. In fact, we can prove something stronger: Namely, there exists no algebra isomorphism $\left(\operatorname{Sym}\left(V\right)\right)^o \to \operatorname{Sym}\left(V\right)$. This is because the algebra $\left(\operatorname{Sym}\left(V\right)\right)^o$ has a nonzero nilpotent element (namely, $f^{(1)}$ is nonzero and satisfies $\left(f^{(1)}\right)^p = 0$), whereas the algebra $\operatorname{Sym}\left(V\right)$ contains no nonzero nilpotent elements (in fact, $\operatorname{Sym}\left(V\right) \cong \mathbf{k}\left[x\right]$ is isomorphic to a polynomial ring over the field $\mathbf{k}$, and therefore is an integral domain, which yields that it contains no nonzero nilpotent elements). This concludes the solution of Exercise 1.6.4(d).

13.36. **Solution to Exercise 1.6.5.** *Solution to Exercise 1.6.5.* (a) Let $\Delta' : \mathbf{k}\left[\mathbf{x}\right] \to \mathbf{k}\left[\mathbf{x}, \mathbf{y}\right]$ be the map sending every polynomial $f\left(x_1, x_2, ..., x_n\right) \in \mathbf{k}\left[\mathbf{x}\right]$ to $f\left(x_1 + y_1, x_2 + y_2, ..., x_n + y_n\right) \in \mathbf{k}\left[\mathbf{x}, \mathbf{y}\right]$. Our goal is to prove that $\Delta' = \Delta_{\operatorname{Sym}(V)}$, where $\Delta_{\operatorname{Sym}(V)}$ is the usual coproduct on $\operatorname{Sym}\left(V\right)$ (part of the coalgebra structure obtained by regarding $\operatorname{Sym}\left(V\right)$ as a quotient of $T\left(V\right)$ as in Exercise 1.3.14), and where we are identifying $\operatorname{Sym}\left(V\right)$ with $\mathbf{k}\left[\mathbf{x}\right]$ and $\operatorname{Sym}\left(V\right) \otimes \operatorname{Sym}\left(V\right)$ with $\mathbf{k}\left[\mathbf{x}, \mathbf{y}\right]$ along the isomorphisms given at the beginning of the exercise.

Notice first that $\Delta_{\operatorname{Sym}(V)}$ is a $\mathbf{k}$-algebra homomorphism $\operatorname{Sym}\left(V\right) \to \operatorname{Sym}\left(V\right) \otimes \operatorname{Sym}\left(V\right)$ (because the axioms of a $\mathbf{k}$-bialgebra require that the coproduct of any $\mathbf{k}$-bialgebra A is a $\mathbf{k}$-algebra homomorphism $A \to A \otimes A$). On the other hand, the map Δ' is a $\mathbf{k}$-algebra homomorphism[469]. Hence, the equality that we are trying to prove, namely $\Delta' = \Delta_{\operatorname{Sym}(V)}$, is an equality between two $\mathbf{k}$-algebra homomorphisms. It is well-known that in order to prove such an equality, it is enough to verify it on a generating set of the domain of these homomorphisms[470]; i.e., it is enough to pick out a generating set of its domain, and check that for every element s of the generating set, the images of s under the two sides of the equality are equal to each other. In our case, the $\mathbf{k}$-algebra homomorphisms Δ' and $\Delta_{\operatorname{Sym}(V)}$ have domain $\mathbf{k}\left[\mathbf{x}\right]$, and as a generating set of this $\mathbf{k}$-algebra $\mathbf{k}\left[\mathbf{x}\right]$ we can pick the set $\left\{x_1, x_2, ..., x_n\right\}$. We then have to check that for every element s of this generating set, the images of s under Δ' and $\Delta_{\operatorname{Sym}(V)}$ are equal to each other.

So let $s \in \left\{x_1, x_2, ..., x_n\right\}$ be arbitrary. Then, $s = x_i$ for some $i \in \left\{1, 2, ..., n\right\}$. Consider this i. We have $x_i \in V$, and thus x_i is a primitive element of $T\left(V\right)$ (by the definition of the coalgebra structure on $T\left(V\right)$). This shows that $\Delta_{T(V)}\left(x_i\right) = 1 \otimes x_i + x_i \otimes 1$. Since $\operatorname{Sym}\left(V\right)$ is a quotient bialgebra of $T\left(V\right)$, this shows that $\Delta_{\operatorname{Sym}(V)}\left(x_i\right) = 1 \otimes x_i + x_i \otimes 1$ as well. Under our identification of $\operatorname{Sym}\left(V\right) \otimes \operatorname{Sym}\left(V\right)$ with $\mathbf{k}\left[\mathbf{x}, \mathbf{y}\right]$, the element $1 \otimes x_i$ of $\operatorname{Sym}\left(V\right) \otimes \operatorname{Sym}\left(V\right)$ equals the element y_i of $\mathbf{k}\left[\mathbf{x}, \mathbf{y}\right]$, and the element $x_i \otimes 1$ of $\operatorname{Sym}\left(V\right) \otimes \operatorname{Sym}\left(V\right)$ equals the element x_i of $\mathbf{k}\left[\mathbf{x}, \mathbf{y}\right]$. Hence, $\Delta_{\operatorname{Sym}(V)}\left(x_i\right) = 1 \otimes x_i + x_i \otimes 1$ rewrites as $\Delta_{\operatorname{Sym}(V)}\left(x_i\right) = y_i + x_i$.

But by the definition of Δ', we have $\Delta'\left(x_i\right) = x_i\left(x_1 + y_1, x_2 + y_2, ..., x_n + y_n\right) = x_i + y_i = y_i + x_i = \Delta_{\operatorname{Sym}(V)}\left(x_i\right)$. In other words, $\Delta'\left(s\right) = \Delta_{\operatorname{Sym}(V)}\left(s\right)$ (since $s = x_i$). Hence, the images of s under Δ' and

[469] In fact, by its very definition, it is an *evaluation homomorphism*, meaning a map from a polynomial ring to a commutative $\mathbf{k}$-algebra A which sends every polynomial f to the evaluation $f\left(a_1, a_2, ..., a_\ell\right)$ for some given tuple $\left(a_1, a_2, ..., a_\ell\right)$ of elements of A. Such maps are always $\mathbf{k}$-algebra homomorphisms.

[470] In our case, the domain is $\mathbf{k}\left[\mathbf{x}\right]$.

$\Delta_{\operatorname{Sym}(V)}$ are equal to each other. This is exactly what we needed to prove, and so the solution to Exercise 1.6.5(a) is complete.

We will solve the remaining parts of the exercise starting with (c), then continuing with (d) and finally solving (b). This order has the advantage of requiring the least amount of work.

(c) We introduce a few notations. As long as we are using the topologist's sign convention, an element a of a graded $\mathbf{k}$-module is said to be *odd* if a is a sum of homogeneous elements of odd degree, and an element b of a graded $\mathbf{k}$-module is said to be *even* if b is a sum of homogeneous elements of even degree. By assumption, every element of V is odd. Notice that 0 is both even and odd, and a graded $\mathbf{k}$-module can (in general) contain vectors which are neither even nor odd.

It is easy to see that if P and Q are two graded $\mathbf{k}$-modules, and $p \in P$ and $q \in Q$ are two elements, then

$$T(p \otimes q) = q \otimes p \qquad \text{if } p \text{ is even and } q \text{ is even;} \tag{13.36.1}$$

$$T(p \otimes q) = q \otimes p \qquad \text{if } p \text{ is even and } q \text{ is odd;} \tag{13.36.2}$$

$$T(p \otimes q) = q \otimes p \qquad \text{if } p \text{ is odd and } q \text{ is even;} \tag{13.36.3}$$

$$T(p \otimes q) = -q \otimes p \qquad \text{if } p \text{ is odd and } q \text{ is odd.} \tag{13.36.4}$$

[471]

Now, let $x \in V$. Then, x is odd (since every element of V is odd). The definition of Δ enforces that Δ is a $\mathbf{k}$-bialgebra homomorphism $T(V) \to T(V) \otimes T(V)$, but the meaning of this depends on the $\mathbf{k}$-algebra $T(V) \otimes T(V)$, and therefore on the twist map T, because the multiplication map of the $\mathbf{k}$-algebra $T(V) \otimes T(V)$ is

$$m_{T(V) \otimes T(V)} = \left(m_{T(V)} \otimes m_{T(V)}\right) \circ (\operatorname{id} \otimes T \otimes \operatorname{id}).$$

The fact that we used the topologist's sign convention (1.3.3) in this definition means that this twist map T is as in (1.3.3). Thus, (13.36.1), (13.36.2), (13.36.3) and (13.36.4) apply. Since x is odd, we can apply (13.36.4) to $P = T(V)$, $Q = T(V)$, $p = x$ and $q = x$, and obtain $T(x \otimes x) = -x \otimes x$. Thus, in the $\mathbf{k}$-algebra

[471] *Proof.* We will only prove (13.36.4); the other three identities are analogous.

Let P and Q be two graded $\mathbf{k}$-modules. Let $p \in P$ and $q \in Q$. Assume that p is odd and q is odd. Since p is odd, we know that p is a sum of homogeneous elements of P of odd degree; in other words, we can write p as $p = \sum_{i \in I} p_i$, where I is a finite set and each $i \in I$ has $p_i \in P$ homogeneous of odd degree. Similarly, we can write q as $q = \sum_{j \in J} q_j$, where J is a finite set and each $j \in J$ has $q_j \in Q$ homogeneous of odd degree. Using these p_i and q_j, we now find

$$\begin{aligned} T\left(\underbrace{p}_{=\sum_{i\in I} p_i} \otimes \underbrace{q}_{=\sum_{j\in J} q_j}\right) &= T\left(\left(\sum_{i \in I} p_i\right) \otimes \left(\sum_{j \in J} q_j\right)\right) = T\left(\sum_{i \in I} \sum_{j \in J} p_i \otimes q_j\right) \\ &= \sum_{i \in I} \sum_{j \in J} \underbrace{T(p_i \otimes q_j)}_{\substack{=(-1)^{\deg(p_i)\cdot\deg(q_j)} q_j \otimes p_i \\ \text{(by the definition of the} \\ \text{topologist's twist map } T\text{)}}} = \sum_{i \in I} \sum_{j \in J} \underbrace{(-1)^{\deg(p_i)\cdot\deg(q_j)}}_{\substack{=-1 \\ \text{(since } \deg(p_i) \text{ and} \\ \deg(q_j) \text{ are odd} \\ \text{(because } p_i \text{ and } q_j \text{ are} \\ \text{homogeneous of odd degree))}}} q_j \otimes p_i \\ &= \sum_{i \in I} \sum_{j \in J} (-1) q_j \otimes p_i = -\sum_{i \in I} \sum_{j \in J} q_j \otimes p_i = -\underbrace{\left(\sum_{j \in J} q_j\right)}_{=q} \otimes \underbrace{\left(\sum_{i \in I} p_i\right)}_{=p} = -q \otimes p. \end{aligned}$$

This proves (13.36.4).

$T(V) \otimes T(V)$, we have

$$\begin{aligned}
&(1 \otimes x) \cdot (x \otimes 1) \\
&= \underbrace{m_{T(V)\otimes T(V)}}_{=\left(m_{T(V)}\otimes m_{T(V)}\right)\circ(\mathrm{id}\otimes T\otimes\mathrm{id})} \underbrace{((1 \otimes x) \otimes (x \otimes 1))}_{=1\otimes x\otimes x\otimes 1} \\
&= \left(\left(m_{T(V)} \otimes m_{T(V)}\right) \circ (\mathrm{id} \otimes T \otimes \mathrm{id})\right) (1 \otimes x \otimes x \otimes 1) \\
&= \left(m_{T(V)} \otimes m_{T(V)}\right) \left(\underbrace{(\mathrm{id} \otimes T \otimes \mathrm{id}) (1 \otimes x \otimes x \otimes 1)}_{\substack{=1\otimes T(x\otimes x)\otimes 1 \\ =-1\otimes x\otimes x\otimes 1 \\ (\text{since } T(x\otimes x)=-x\otimes x)}} \right) \\
&= \left(m_{T(V)} \otimes m_{T(V)}\right) (-1 \otimes x \otimes x \otimes 1) = -\underbrace{m_{T(V)} (1 \otimes x)}_{=1\cdot x = x} \otimes \underbrace{m_{T(V)} (x \otimes 1)}_{=x\cdot 1 = x} \\
&= -x \otimes x.
\end{aligned}$$

Similarly, we can compute

$$\begin{aligned}
(1 \otimes x) \cdot (1 \otimes x) &= 1 \otimes x^2, \\
(x \otimes 1) \cdot (1 \otimes x) &= x \otimes x, \\
(x \otimes 1) \cdot (x \otimes 1) &= x^2 \otimes 1.
\end{aligned}$$

(Since $1 \in T(V)$ is even, we have to use (13.36.3), (13.36.1) and respectively (13.36.2) instead of (13.36.4) here; thus, we incur no negative signs.)

Since Δ is a $\mathbf{k}$-algebra homomorphism, we have

$$\begin{aligned}
\Delta\left(x^2\right) &= \left(\underbrace{\Delta(x)}_{\substack{=1\otimes x + x\otimes 1 \\ (\text{since } x \text{ is primitive in } T(V))}} \right)^2 = (1 \otimes x + x \otimes 1)^2 \\
&= (1 \otimes x + x \otimes 1) \cdot (1 \otimes x + x \otimes 1) \\
&= \underbrace{(1 \otimes x) \cdot (1 \otimes x)}_{=1\otimes x^2} + \underbrace{(1 \otimes x) \cdot (x \otimes 1)}_{=-x\otimes x} + \underbrace{(x \otimes 1) \cdot (1 \otimes x)}_{=x\otimes x} + \underbrace{(x \otimes 1) \cdot (x \otimes 1)}_{=x^2\otimes 1} \\
&= 1 \otimes x^2 + (-x \otimes x) + x \otimes x + x^2 \otimes 1 = 1 \otimes x^2 + x^2 \otimes 1.
\end{aligned}$$

This solves Exercise 1.6.5(c).

(d) Let us first check that the ideal[472] J of $T(V)$ is a graded $\mathbf{k}$-submodule of $T(V)$. This is **not** obvious! We know that

$$\left(\begin{array}{c} \text{if an ideal } I \text{ of a graded } \mathbf{k}\text{-algebra } A \text{ is generated by homogeneous} \\ \text{elements, then } I \text{ is a graded } \mathbf{k}\text{-submodule of } A \end{array} \right). \tag{13.36.5}$$

[473] So we should concentrate on showing that the ideal J is generated by homogeneous elements.

In fact, we know that J is the ideal of $T(V)$ generated by $\left\{x^2\right\}_{x\in V}$. Thus,

$$x^2 \in J \qquad \text{for every } x \in V. \tag{13.36.6}$$

[472] By "ideal", we always mean "two-sided ideal", unless we explicitly say "left ideal" or "right ideal".

[473] This is a well-known fact in the case when A is commutative (and the topologist's sign convention is not used), but it is proven exactly the same way in the general case.

Also, any two elements x and y of V satisfy

$$xy + yx = \underbrace{(x+y)^2}_{\substack{\in J \\ \text{(by (13.36.6), applied} \\ \text{to } x+y \text{ instead of } x)}} - \underbrace{x^2}_{\substack{\in J \\ \text{(by (13.36.6))}}} - \underbrace{y^2}_{\substack{\in J \\ \text{(by (13.36.6), applied} \\ \text{to } y \text{ instead of } x)}} \qquad \left(\text{since } (x+y)^2 = x^2 + xy + yx + y^2\right)$$
$$\in J - J - J \subset J \qquad (\text{since } J \text{ is a } \mathbf{k}\text{-module}). \tag{13.36.7}$$

Now, define a subset G of $T(V)$ by

$$G = \left\{x^2 \mid x \in V;\ x \text{ is homogeneous}\right\} \cup \left\{xy + yx \mid x \in V;\ y \in V;\ x \text{ and } y \text{ are homogeneous}\right\}. \tag{13.36.8}$$

It is clear that all elements of G are homogeneous elements of $T(V)$. By the definition of G, we have

$$x^2 \in G \qquad \text{for every homogeneous } x \in V, \tag{13.36.9}$$

and for the same reason we have

$$xy + yx \in G \qquad \text{for any two homogeneous elements } x \text{ and } y \text{ of } V. \tag{13.36.10}$$

We shall now show that the ideal generated by G is J.

Indeed, let $g \in G$ be arbitrary. We will now show that $g \in J$. We have

$$g \in G = \left\{x^2 \mid x \in V;\ x \text{ is homogeneous}\right\} \cup \left\{xy + yx \mid x \in V;\ y \in V;\ x \text{ and } y \text{ are homogeneous}\right\}.$$

Hence, either g has the form $g = x^2$ for some homogeneous $x \in V$, or g has the form $g = xy + yx$ for two homogeneous elements x and y of V. In the first of these two cases, it is clear that $g = x^2 \in J$ (by (13.36.6)). In the second of these cases, we have $g = xy + yx \in J$ (by (13.36.7)). Hence, we have shown that $g \in J$ in either case. This proves that $g \in J$.

Now, forget that we fixed g. We have thus proven that $g \in J$ for every $g \in G$. Thus, $G \subset J$. In other words, J contains G as a subset. Since J is an ideal of $T(V)$, we thus have

$$J \supset (\text{the smallest ideal of } T(V) \text{ containing } G \text{ as a subset}) = (\text{the ideal of } T(V) \text{ generated by } G). \tag{13.36.11}$$

Let us now show the reverse inclusion. Let $x \in V$. Then, x is a sum of homogeneous elements of V (because every element of a graded $\mathbf{k}$-module is a sum of homogeneous elements). In other words, we can write x in the form $x = \sum_{i=1}^{\ell} x_i$, where $\ell \in \mathbb{N}$, and where $x_1, x_2, \ldots, x_\ell$ are homogeneous elements of V. Consider this ℓ and these $x_1, x_2, \ldots, x_\ell$. Set $I = \{1, 2, \ldots, \ell\}$. Thus, we have the following equality of summation signs: $\sum_{i \in I} = \sum_{i \in \{1,2,\ldots,\ell\}} = \sum_{i=1}^{\ell}$. Hence, the equality $x = \sum_{i=1}^{\ell} x_i$ rewrites as $x = \sum_{i \in I} x_i$.

Squaring both sides of the equality $x = \sum_{i \in I} x_i$, we obtain

$$\begin{aligned}
x^2 &= \left(\sum_{i\in I} x_i\right)^2 = \sum_{(i,j)\in I^2} x_i x_j = \underbrace{\sum_{\substack{(i,j)\in I^2;\\ i<j}} x_i x_j}_{\substack{= \sum_{\substack{(i,j)\in I^2;\\ i>j}} x_j x_i \\ \text{(here, we substituted (j,i)} \\ \text{for (i,j) in the sum)}}} + \sum_{\substack{(i,j)\in I^2;\\ i=j}} \underbrace{x_i x_j}_{\substack{=x_j^2 \\ \text{(since $i=j$)}}} + \sum_{\substack{(i,j)\in I^2;\\ i>j}} x_i x_j \\
&= \sum_{\substack{(i,j)\in I^2;\\ i>j}} x_j x_i + \sum_{\substack{(i,j)\in I^2;\\ i=j}} x_j^2 + \sum_{\substack{(i,j)\in I^2;\\ i>j}} x_i x_j = \sum_{\substack{(i,j)\in I^2;\\ i=j}} x_j^2 + \underbrace{\sum_{\substack{(i,j)\in I^2;\\ i>j}} x_i x_j + \sum_{\substack{(i,j)\in I^2;\\ i>j}} x_j x_i}_{= \sum_{\substack{(i,j)\in I^2;\\ i>j}} (x_i x_j + x_j x_i)} \\
&= \sum_{\substack{(i,j)\in I^2;\\ i=j}} \underbrace{x_j^2}_{\substack{\in G \\ \text{(by (13.36.9),} \\ \text{applied to x_j instead of x)}}} + \sum_{\substack{(i,j)\in I^2;\\ i>j}} \underbrace{(x_i x_j + x_j x_i)}_{\substack{\in G \\ \text{(by (13.36.10), applied} \\ \text{to x_i and x_j instead of x and y)}}} \\
&\in \sum_{\substack{(i,j)\in I^2;\\ i=j}} G + \sum_{\substack{(i,j)\in I^2;\\ i>j}} G \subset (\text{the ideal of } T(V) \text{ generated by } G).
\end{aligned}$$

Now, forget that we fixed x. We thus have shown that $x^2 \in (\text{the ideal of } T(V) \text{ generated by } G)$ for every $x \in V$. Hence, $(\text{the ideal of } T(V) \text{ generated by } G)$ contains the elements x^2 for all $x \in V$. Since $(\text{the ideal of } T(V) \text{ generated by } G)$ is an ideal, this yields that

$$\begin{aligned}
&(\text{the ideal of } T(V) \text{ generated by } G) \\
&\supset \left(\text{the smallest ideal of } T(V) \text{ containing the elements } x^2 \text{ for all } x \in V\right) \\
&= \left(\text{the ideal generated by } \left\{x^2\right\}_{x\in V}\right) = J.
\end{aligned}$$

Combined with (13.36.11), this yields

$$J = (\text{the ideal of } T(V) \text{ generated by } G).$$

The ideal J is thus generated by G. Thus, the ideal J of $T(V)$ is generated by homogeneous elements of $T(V)$ (since all elements of G are homogeneous elements of $T(V)$). Therefore, (13.36.5) (applied to J and $T(V)$ instead of I and A) yields that J is a graded **k**-submodule of $T(V)$. Hence, the quotient $T(V)/J$ is a graded **k**-module.

Now, let us show that $\Delta(J) \subset J \otimes T(V) + T(V) \otimes J$.

If A and B are two graded **k**-algebras (in the topologist's sense) and P and Q are two ideals of A and B which are graded **k**-submodules of A and B, then $P \otimes Q$ is an ideal of $A \otimes B$ [474]. Applying this to $A = T(V)$, $B = T(V)$, $P = J$ and $Q = T(V)$, we conclude that $J \otimes T(V)$ is an ideal of $T(V) \otimes T(V)$. Similarly, $T(V) \otimes J$ is an ideal of $T(V) \otimes T(V)$ as well. The sum $J \otimes T(V) + T(V) \otimes J$ of these two ideals therefore is an ideal of $T(V) \otimes T(V)$, too. As a consequence, $\Delta^{-1}(J \otimes T(V) + T(V) \otimes J)$ is an

[474] *Proof.* Since P is an ideal of A, we see that P is a **k**-submodule of A satisfying $m_A(P \otimes A) \subset P$ and $m_A(A \otimes P) \subset P$. Since Q is an ideal of B, we see that Q is a **k**-submodule of B satisfying $m_B(Q \otimes B) \subset Q$ and $m_B(B \otimes Q) \subset Q$.

ideal of $T(V)$ (because Δ is a $\mathbf{k}$-algebra homomorphism[475], and the preimage of an ideal under a $\mathbf{k}$-algebra homomorphism is always an ideal).

Now, every $x \in V$ satisfies

$$\Delta\left(x^2\right) = \underbrace{1}_{\in T(V)} \otimes \underbrace{x^2}_{\substack{\in J \\ (\text{by } (13.36.6))}} + \underbrace{x^2}_{\substack{\in J \\ (\text{by } (13.36.6))}} \otimes \underbrace{1}_{\in T(V)} \qquad (\text{by Exercise 1.6.5(c)})$$
$$\in T(V) \otimes J + J \otimes T(V) = J \otimes T(V) + T(V) \otimes J,$$

so that $x^2 \in \Delta^{-1}(J \otimes T(V) + T(V) \otimes J)$. So the ideal $\Delta^{-1}(J \otimes T(V) + T(V) \otimes J)$ contains the elements x^2 for all $x \in V$. Hence,

$$\Delta^{-1}(J \otimes T(V) + T(V) \otimes J) \supset \left(\text{the smallest ideal which contains the elements } x^2 \text{ for all } x \in V\right)$$
$$= \left(\text{the ideal generated by } \left\{x^2\right\}_{x \in V}\right) = J.$$

Thus, $\Delta(J) \subset J \otimes T(V) + T(V) \otimes J$.

It remains to prove that $\epsilon(J) = 0$. This is similar to the above argument but much simpler. Since ϵ is a $\mathbf{k}$-algebra homomorphism, its kernel $\ker \epsilon$ is an ideal of $T(V)$. Since ϵ is a $\mathbf{k}$-algebra homomorphism, every $x \in V$ satisfies $\epsilon\left(x^2\right) = \left(\underbrace{\epsilon(x)}_{=0}\right)^2 = 0$ and thus $x^2 \in \ker \epsilon$. Thus, the ideal $\ker \epsilon$ contains the elements x^2 for all $x \in V$. Hence,

$$\ker \epsilon \supset \left(\text{the smallest ideal which contains the elements } x^2 \text{ for all } x \in V\right)$$
$$= \left(\text{the ideal generated by } \left\{x^2\right\}_{x \in V}\right) = J.$$

This yields $\epsilon(J) = 0$. Combined with $\Delta(J) \subset J \otimes T(V) + T(V) \otimes J$, this shows that J is a two-sided coideal of $T(V)$. Since J is also a two-sided ideal and a graded $\mathbf{k}$-submodule, this shows that the quotient $T(V)/J$ inherits a graded $\mathbf{k}$-bialgebra structure from $T(V)$. This quotient $T(V)/J$ is $\wedge V$, and so we obtain a graded $\mathbf{k}$-bialgebra structure on $\wedge V$. The $\mathbf{k}$-bialgebra $\wedge V$ obtained this way is graded (because it is the quotient of the graded $\mathbf{k}$-bialgebra $T(V)$ by the graded ideal J) and connected (since its 0-th graded component is $\wedge^0 V = \mathbf{k}$ [476]), therefore a Hopf algebra (by Proposition 1.4.16[477]). Thus, $\wedge V$ is a

But since Q is a graded $\mathbf{k}$-submodule of B, there is a topologist's twist map $T : Q \otimes A \to A \otimes Q$, which is the restriction of the twist map $T : B \otimes A \to A \otimes B$. Thus, $T(Q \otimes A) \subset A \otimes Q$. Now,

$$\underbrace{m_{A \otimes B}}_{=(m_A \otimes m_B) \circ (\mathrm{id} \otimes T \otimes \mathrm{id})} \left(\underbrace{(P \otimes Q) \otimes (A \otimes B)}_{=P \otimes Q \otimes A \otimes B}\right)$$
$$= ((m_A \otimes m_B) \circ (\mathrm{id} \otimes T \otimes \mathrm{id}))(P \otimes Q \otimes A \otimes B)$$
$$= (m_A \otimes m_B)\left(\underbrace{(\mathrm{id} \otimes T \otimes \mathrm{id})(P \otimes Q \otimes A \otimes B)}_{\substack{=P \otimes T(Q \otimes A) \otimes B \\ \subset P \otimes A \otimes Q \otimes B \\ (\text{since } T(Q \otimes A) \subset A \otimes Q)}}\right)$$
$$\subset (m_A \otimes m_B)(P \otimes A \otimes Q \otimes B) = \underbrace{m_A(P \otimes A)}_{\subset P} \otimes \underbrace{m_B(Q \otimes B)}_{\subset Q} \subset P \otimes Q.$$

Similarly, $m_{A \otimes B}((A \otimes B) \otimes (P \otimes Q)) \subset P \otimes Q$ (but here, we need to use $T(B \otimes P) \subset P \otimes B$ instead of $T(Q \otimes A) \subset A \otimes Q$). These two inclusions prove that $P \otimes Q$ is an ideal of $A \otimes B$, qed.

We could have also proven this by working with elements, but that way we would have to take care of the fact that the twist T is the topologist's one and comes with signs. The way we have done it, we were almost entirely untroubled by this fact.

[475] by the definition of Δ

[476] Here we are using the fact that $V_0 = 0$ (which is a consequence of the fact that V is concentrated in odd degrees).

[477] Here it helps to notice that our use of the topologist's sign convention does not invalidate the proof of Proposition 1.4.16; in fact, no changes are necessary to that proof! (This might not be too surprising, given that said proof made no use of the bialgebra axioms (1.3.4).)

connected graded Hopf algebra, and therefore the comultiplication $\Delta_{\wedge V}$ of $\wedge V$ is part of a connected graded Hopf algebra structure on $\wedge V$. In other words, the comultiplication $\Delta_{\wedge V}$ makes the $\mathbf{k}$-algebra $\wedge V$ into a connected graded Hopf algebra.

Let us recall that the comultiplication $\Delta_{T(V)}$ of the $\mathbf{k}$-bialgebra $T(V)$ satisfies

$$\Delta_{T(V)}(x) = 1 \otimes x + x \otimes 1 \qquad \text{for every } x \in V$$

(by the definition of $\Delta_{T(V)}$). Since the $\mathbf{k}$-bialgebra $\wedge V$ was defined as a quotient of $T(V)$, we can project this equality down on $(\wedge V) \otimes (\wedge V)$, and thus conclude that the comultiplication $\Delta_{\wedge V}$ of the $\mathbf{k}$-bialgebra $\wedge V$ satisfies

$$\Delta_{\wedge V}(x) = 1 \otimes x + x \otimes 1 \qquad \text{for every } x \in V. \tag{13.36.12}$$

Thus, every $i \in \{1, 2, ..., n\}$ satisfies

$$\Delta_{\wedge V}(x_i) = 1 \otimes x_i + x_i \otimes 1 \tag{13.36.13}$$

(by (13.36.12), applied to $x = x_i$). But our identification of $(\wedge V) \otimes (\wedge V)$ with $\wedge(V \oplus V)$ equates $1 \otimes x_i$ with y_i, and equates $x_i \otimes 1$ with x_i for every $i \in \{1, 2, ..., n\}$. Hence, for every $i \in \{1, 2, ..., n\}$, the equality (13.36.13) rewrites as

$$\Delta_{\wedge V}(x_i) = \underbrace{1 \otimes x_i}_{=y_i} + \underbrace{x_i \otimes 1}_{=x_i} = y_i + x_i = x_i + y_i \tag{13.36.14}$$

in $(\wedge V) \otimes (\wedge V) = \wedge(V \oplus V)$.

Recall that we made $\wedge V$ into a $\mathbf{k}$-bialgebra by viewing it as the quotient $T(V)/J$. Now, it remains to prove that the coproduct on $\wedge V$ which is part of this $\mathbf{k}$-bialgebra structure on $T(V)$ is the same as the one defined in Exercise 1.6.5(b)[478]. In order to do so, we shall prove the following claim: If $\Delta_{\wedge V}$ denotes the coproduct on $\wedge V$ obtained by regarding $\wedge V$ as the quotient $\mathbf{k}$-bialgebra $T(V)/J$, then

$$\Delta_{\wedge V}\left(\sum_{i_1<\cdots<i_d} c_{i_1,...,i_d} x_{i_1} \wedge \cdots \wedge x_{i_d}\right) = \sum_{i_1<\cdots<i_d} c_{i_1,...,i_d} (x_{i_1} + y_{i_1}) \wedge \cdots \wedge (x_{i_d} + y_{i_d}) \tag{13.36.15}$$

for every family $(c_{i_1,...,i_d})_{i_1<\cdots<i_d}$ of elements of $\mathbf{k}$ indexed by the strictly increasing sequences $(i_1 < \cdots < i_d)$ of elements of $\{1, 2, ..., n\}$.

Proof of (13.36.15): Let $(c_{i_1,...,i_d})_{i_1<\cdots<i_d}$ be a family of elements of $\mathbf{k}$ indexed by the strictly increasing sequences $(i_1 < \cdots < i_d)$ of elements of $\{1, 2, ..., n\}$. The multiplication in $\wedge V$ is given by the wedge product, so that we have $x_{i_1} \cdots x_{i_d} = x_{i_1} \wedge \cdots \wedge x_{i_d}$ for every strictly increasing sequence $(i_1 < \cdots < i_d)$ of elements of $\{1, 2, ..., n\}$.

But the comultiplication $\Delta_{\wedge V}$ is a $\mathbf{k}$-algebra homomorphism (by the axioms of a $\mathbf{k}$-bialgebra, since $\wedge V$ is a $\mathbf{k}$-bialgebra), and thus we have

$$\begin{aligned}
\Delta_{\wedge V}\left(\sum_{i_1<\cdots<i_d} c_{i_1,...,i_d} x_{i_1} \cdots x_{i_d}\right) &= \sum_{i_1<\cdots<i_d} c_{i_1,...,i_d} \underbrace{\Delta_{\wedge V}(x_{i_1})}_{\substack{=x_{i_1}+y_{i_1} \\ \text{(by (13.36.14),} \\ \text{applied to } i=i_1)}} \cdots \underbrace{\Delta_{\wedge V}(x_{i_d})}_{\substack{=x_{i_d}+y_{i_d} \\ \text{(by (13.36.14),} \\ \text{applied to } i=i_d)}} \\
&= \sum_{i_1<\cdots<i_d} c_{i_1,...,i_d} \underbrace{(x_{i_1} + y_{i_1}) \cdots (x_{i_d} + y_{i_d})}_{\substack{=(x_{i_1}+y_{i_1})\wedge\cdots\wedge(x_{i_d}+y_{i_d}) \\ \text{(since multiplication in } \wedge(V\oplus V) \\ \text{is given by the wedge product)}}} \\
&= \sum_{i_1<\cdots<i_d} c_{i_1,...,i_d} (x_{i_1} + y_{i_1}) \wedge \cdots \wedge (x_{i_d} + y_{i_d}).
\end{aligned}$$

Since $x_{i_1} \cdots x_{i_d} = x_{i_1} \wedge \cdots \wedge x_{i_d}$ for every strictly increasing sequence $(i_1 < \cdots < i_d)$ of elements of $\{1, 2, ..., n\}$, this rewrites as follows:

$$\Delta_{\wedge V}\left(\sum_{i_1<\cdots<i_d} c_{i_1,...,i_d} x_{i_1} \wedge \cdots \wedge x_{i_d}\right) = \sum_{i_1<\cdots<i_d} c_{i_1,...,i_d} (x_{i_1} + y_{i_1}) \wedge \cdots \wedge (x_{i_d} + y_{i_d}).$$

[478] We should be careful because we have not yet proven that the latter coproduct satisfies the axioms for a coproduct (we have not solved Exercise 1.6.5(b) yet).

Thus, (13.36.15) is proven.

Now, the equality (13.36.15) shows that the map $\Delta_{\wedge V}$ satisfies the defining property of the map (1.6.6) (namely, mapping every $\sum_{i_1<\cdots<i_d} c_{i_1,\ldots,i_d} x_{i_1} \wedge \cdots \wedge x_{i_d}$ to $\sum_{i_1<\cdots<i_d} c_{i_1,\ldots,i_d} (x_{i_1}+y_{i_1}) \wedge \cdots \wedge (x_{i_d}+y_{i_d})$). Hence, there exists a map satisfying this property. Such a map is furthermore unique (because the defining property determines its value on every element of the form $\sum_{i_1<\cdots<i_d} c_{i_1,\ldots,i_d} x_{i_1} \wedge \cdots \wedge x_{i_d}$, but every element of $\wedge V$ can be written in this form), and therefore the map (1.6.6) is well-defined. This map is our map $\Delta_{\wedge V}$ (because our map $\Delta_{\wedge V}$ satisfies the defining property of the map (1.6.6)), and therefore makes $\wedge V$ into a connected graded Hopf algebra (since we know that the comultiplication $\Delta_{\wedge V}$ makes the $\mathbf{k}$-algebra $\wedge V$ into a connected graded Hopf algebra). This solves Exercise 1.6.5(b).

But the coproduct on $\wedge V$ inherited from $T(V)$ is $\Delta_{\wedge V}$, and as we know, this $\Delta_{\wedge V}$ is exactly the map (1.6.6), i.e., the coproduct defined in Exercise 1.6.5(b). Hence, the coproduct on $\wedge V$ inherited from $T(V)$ is the coproduct defined in Exercise 1.6.5(b). Thus, Exercise 1.6.5(d) is also solved. The solution to Exercise 1.6.5 is thus complete.

13.37. **Solution to Exercise 1.6.6.** *Solution to Exercise 1.6.6.* Define a $\mathbf{k}$-linear map $\rho_{U,V} : U^* \otimes V^* \to (U \otimes V)^*$ for any two $\mathbf{k}$-modules U and V as in Exercise 1.6.1. Recall that this $\rho_{U,V}$ is a $\mathbf{k}$-module isomorphism if both U and V are finite free. Hence, $\rho_{A,A}$ is a $\mathbf{k}$-module isomorphism.

Also, basic linear algebra shows that if U, V, U' and V' are four $\mathbf{k}$-modules and if $\alpha : U \to U'$ and $\beta : V \to V'$ are two $\mathbf{k}$-linear maps, then

$$\rho_{U,V} \circ (\alpha^* \otimes \beta^*) = (\alpha \otimes \beta)^* \circ \rho_{U',V'}.$$

This (applied to $U = C$, $V = C$, $U' = A$, $V' = A$, $\alpha = f$ and $\beta = g$) yields

$$\rho_{C,C} \circ (f^* \otimes g^*) = (f \otimes g)^* \circ \rho_{A,A}.$$

The definition of convolution yields both

$$f \star g = m_A \circ (f \otimes g) \circ \Delta_C \tag{13.37.1}$$

and

$$\begin{aligned} f^* \star g^* &= \underbrace{m_{C^*}}_{\substack{=\Delta_C^* \circ \rho_{C,C} \\ \text{(by the definition of } m_{C^*}\text{)}}} \circ (f^* \otimes g^*) \circ \underbrace{\Delta_{A^*}}_{\substack{=\rho_{A,A}^{-1} \circ m_A^* \\ \text{(by the definition of } \Delta_{A^*}\text{)}}} \\ &= \Delta_C^* \circ \underbrace{\rho_{C,C} \circ (f^* \otimes g^*)}_{=(f \otimes g)^* \circ \rho_{A,A}} \circ \rho_{A,A}^{-1} \circ m_A^* = \Delta_C^* \circ (f \otimes g)^* \circ \underbrace{\rho_{A,A} \circ \rho_{A,A}^{-1}}_{=\mathrm{id}_{(A \otimes A)^*}} \circ m_A^* \\ &= \Delta_C^* \circ (f \otimes g)^* \circ m_A^* = \left(\underbrace{m_A \circ (f \otimes g) \circ \Delta_C}_{\substack{=f \star g \\ \text{(by (13.37.1))}}} \right)^* = (f \star g)^* . \end{aligned}$$

This solves Exercise 1.6.6.

13.38. **Solution to Exercise 1.6.8.** *Solution to Exercise 1.6.8.* Our goal is to prove Proposition 1.6.7.

Let $m_{⧢}$ denote the $\mathbf{k}$-linear map $T(V) \otimes T(V) \to T(V)$ which sends every $a \otimes b$ to $a ⧢ b$. Let u denote the unit map of $T(V)$ (that is, the $\mathbf{k}$-linear map $\mathbf{k} \to T(V)$ sending $1_{\mathbf{k}}$ to $1_{T(V)}$). Let S denote the antipode of the Hopf algebra $T(V)$. Then, the result that we have to prove boils down to the statement that the $\mathbf{k}$-module $T(V)$, endowed with the multiplication $m_{⧢}$, the unit u, the comultiplication $\Delta_{⧢}$ and the counit ϵ, becomes a commutative Hopf algebra with the antipode S. Since we already know that $m_{⧢}$, u, $\Delta_{⧢}$, ϵ and S are $\mathbf{k}$-linear, this latter statement will immediately follow once we can show that the following diagrams commute:

- the diagrams (1.1.1) and (1.1.2), with A and m replaced by $T(V)$ and $m_{⧢}$;
- the diagrams (1.2.1) and (1.2.2), with C and Δ replaced by $T(V)$ and $\Delta_{⧢}$;

- the diagrams (1.3.4), with A, m and Δ replaced by $T(V)$, $m_{\sqcup\!\sqcup}$ and $\Delta_{\sqcup\!\sqcup}$;
- the diagram (1.4.3), with A, m and Δ replaced by $T(V)$, $m_{\sqcup\!\sqcup}$ and $\Delta_{\sqcup\!\sqcup}$;
- the diagram (1.5.1), with A and m replaced by $T(V)$ and $m_{\sqcup\!\sqcup}$.

Out of all these statements, we will only prove the commutativity of the first diagram in (1.3.4), since all other diagrams are similar but only easier.

So we must prove the commutativity of the first diagram in (1.3.4), with A, m and Δ replaced by $T(V)$, $m_{\sqcup\!\sqcup}$ and $\Delta_{\sqcup\!\sqcup}$. In other words, we must show that the diagram

(13.38.1)
$$\begin{tikzcd}
 & & T(V)\otimes T(V) \arrow[dll, "\Delta_{\sqcup\!\sqcup}\otimes\Delta_{\sqcup\!\sqcup}"'] \arrow[ddr, "m_{\sqcup\!\sqcup}"] & \\
T(V)\otimes T(V)\otimes T(V)\otimes T(V) \arrow[dd, "\mathrm{id}\otimes T\otimes\mathrm{id}"'] & & & \\
 & & & T(V) \arrow[ddl, "\Delta_{\sqcup\!\sqcup}"] \\
T(V)\otimes T(V)\otimes T(V)\otimes T(V) \arrow[drr, "m_{\sqcup\!\sqcup}\otimes m_{\sqcup\!\sqcup}"'] & & & \\
 & & T(V)\otimes T(V) &
\end{tikzcd}$$

commutes, where $T : T(V)\otimes T(V) \to T(V)\otimes T(V)$ is the twist map $T_{T(V),T(V)}$ sending every $a\otimes b$ to $b\otimes a$ (and being $\mathbf{k}$-linear at that). Let us prove this now.

We need to prove that the diagram (13.38.1) commutes. In other words, we need to prove the identity

$$(m_{\sqcup\!\sqcup}\otimes m_{\sqcup\!\sqcup})\circ(\mathrm{id}\otimes T\otimes\mathrm{id})\circ(\Delta_{\sqcup\!\sqcup}\otimes\Delta_{\sqcup\!\sqcup}) = \Delta_{\sqcup\!\sqcup}\circ m_{\sqcup\!\sqcup}. \tag{13.38.2}$$

By linearity, it is clearly enough to verify this identity only on the pure tensors in $T(V)\otimes T(V)$; that is, it is enough to check that every $a\in T(V)$ and $b\in T(V)$ satisfy

$$((m_{\sqcup\!\sqcup}\otimes m_{\sqcup\!\sqcup})\circ(\mathrm{id}\otimes T\otimes\mathrm{id})\circ(\Delta_{\sqcup\!\sqcup}\otimes\Delta_{\sqcup\!\sqcup}))(a\otimes b) = (\Delta_{\sqcup\!\sqcup}\circ m_{\sqcup\!\sqcup})(a\otimes b). \tag{13.38.3}$$

So let $a\in T(V)$ and $b\in T(V)$ be arbitrary. All we need now is to prove (13.38.3). By linearity again, we can WLOG assume that a and b have the form $a = v_1v_2\cdots v_p$ and $b = v_{p+1}v_{p+2}\cdots v_{p+q}$ for some $p\in\mathbb{N}$, $q\in\mathbb{N}$ and $v_1,v_2,\ldots,v_{p+q}\in V$ (since $T(V)$ is spanned as a $\mathbf{k}$-module by pure tensors). Assume this, and define W to be the free $\mathbf{k}$-module with basis $\{x_1,x_2,\ldots,x_{p+q}\}$. Let A be the tensor algebra $T(W)$ of this $\mathbf{k}$-module W. Then, W is a finite free $\mathbf{k}$-module, and so we know from Example 1.6.3 (applied to W instead of V) that the graded dual A^o of its tensor algebra $A = T(W)$ is a Hopf algebra whose basis $\left\{y_{(i_1,i_2,\ldots,i_\ell)}\right\}$ is indexed by words in the alphabet $I := \{1,2,\ldots,p+q\}$.

Notice that $\left\{y_{(i_1,i_2,\ldots,i_\ell)}\otimes y_{(j_1,j_2,\ldots,j_m)}\right\}_{\ell\in\mathbb{N},\ m\in\mathbb{N},\ (i_1,i_2,\ldots,i_\ell)\in I^\ell,\ (j_1,j_2,\ldots,j_m)\in I^m}$ is a $\mathbf{k}$-module basis of $A^o\otimes A^o$ (since $\left\{y_{(i_1,i_2,\ldots,i_\ell)}\right\}$ is a basis of A^o). Relabelling this basis, we see that

$$\left\{y_{(i_1,i_2,\ldots,i_\ell)}\otimes y_{(i_{\ell+1},i_{\ell+2},\ldots,i_{\ell+m})}\right\}_{\ell\in\mathbb{N},\ m\in\mathbb{N},\ (i_1,i_2,\ldots,i_{\ell+m})\in I^{\ell+m}} \text{ is a } \mathbf{k}\text{-module basis of } A^o\otimes A^o.$$

We can define a $\mathbf{k}$-linear map $\phi : A^o \to T(V)$ by setting

$$\phi\left(y_{(i_1,i_2,\ldots,i_\ell)}\right) = v_{i_1}v_{i_2}\cdots v_{i_\ell} \qquad \text{for every } \ell\in\mathbb{N} \text{ and } (i_1,i_2,\ldots,i_\ell)\in I^\ell$$

(because $\left\{y_{(i_1,i_2,\ldots,i_\ell)}\right\}$ is a basis of A^o). Consider this map ϕ.

Notice that the definition of ϕ yields $\phi\left(y_{(1,2,\ldots,p)}\right) = v_1 v_2 \cdots v_p = a$, and similarly $\phi\left(y_{(p+1,p+2,\ldots,p+q)}\right) = b$. Thus,

$$\begin{aligned}\underbrace{a}_{=\phi\left(y_{(1,2,\ldots,p)}\right)} \otimes \underbrace{b}_{=\phi\left(y_{(p+1,p+2,\ldots,p+q)}\right)} &= \phi\left(y_{(1,2,\ldots,p)}\right) \otimes \phi\left(y_{(p+1,p+2,\ldots,p+q)}\right) \\ &= (\phi \otimes \phi)\left(\underbrace{y_{(1,2,\ldots,p)} \otimes y_{(p+1,p+2,\ldots,p+q)}}_{\in A^o \otimes A^o}\right) \in (\phi \otimes \phi)\left(A^o \otimes A^o\right).\end{aligned}$$

In other words, $a \otimes b$ lies in the image of the map $\phi \otimes \phi$.

Notice that we already know that A^o is a $\mathbf{k}$-bialgebra, and thus it satisfies all the axioms of a bialgebra; in particular, the diagrams (1.3.4), with A, m and Δ replaced by A^o, m_{A^o} and Δ_{A^o}, commute. In particular, the first of these diagrams commutes. In other words, the diagram

(13.38.4)
$$\begin{tikzcd}
& A^o \otimes A^o \arrow[ld, "\Delta_{A^o} \otimes \Delta_{A^o}"'] \arrow[rdd, "m_{A^o}"] & \\
A^o \otimes A^o \otimes A^o \otimes A^o \arrow[dd, "\operatorname{id} \otimes T \otimes \operatorname{id}"'] & & \\
& & A^o \arrow[ldd, "\Delta_{A^o}"] \\
A^o \otimes A^o \otimes A^o \otimes A^o \arrow[rd, "m_{A^o} \otimes m_{A^o}"'] & & \\
& A^o \otimes A^o &
\end{tikzcd}$$

(with T now denoting the twist map $T_{A^o,A^o} : A^o \otimes A^o \to A^o \otimes A^o$) commutes.

Now, we claim that

$$\phi \circ m_{A^o} = m_{\sqcup\!\sqcup} \circ (\phi \otimes \phi), \tag{13.38.5}$$

$$\phi \circ u_{A^o} = u, \tag{13.38.6}$$

$$(\phi \otimes \phi) \circ \Delta_{A^o} = \Delta_{\sqcup\!\sqcup} \circ \phi, \tag{13.38.7}$$

$$\epsilon_{A^o} = \epsilon \circ \phi, \tag{13.38.8}$$

$$\phi \circ S_{A^o} = S \circ \phi. \tag{13.38.9}$$

[479]

We are going to only prove the two equalities (13.38.5) and (13.38.7), leaving the (simpler!) proofs of the other three equalities (13.38.6), (13.38.8) and (13.38.9) to the reader. (Only the equalities (13.38.5) and (13.38.7) will be used in the proof of the commutativity of the first diagram in (1.3.4); the other are used for the other diagrams.)

Proof of (13.38.5)*:* We need to prove the equality (13.38.5). Since both sides of this equality (13.38.5) are $\mathbf{k}$-linear maps, it is enough to prove this equality on a $\mathbf{k}$-basis of $A^o \otimes A^o$. Let us pick the basis $\left\{y_{(i_1,i_2,\ldots,i_\ell)} \otimes y_{(i_{\ell+1},i_{\ell+2},\ldots,i_{\ell+m})}\right\}_{\ell\in\mathbb{N},\ m\in\mathbb{N},\ (i_1,i_2,\ldots,i_{\ell+m})\in I^{\ell+m}}$; it thus is enough to prove the equality (13.38.5) on this basis, i.e., to prove that

$$\begin{aligned}&\left(\phi \circ m_{A^o}\right)\left(y_{(i_1,i_2,\ldots,i_\ell)} \otimes y_{(i_{\ell+1},i_{\ell+2},\ldots,i_{\ell+m})}\right) \\ &= \left(m_{\sqcup\!\sqcup} \circ (\phi \otimes \phi)\right)\left(y_{(i_1,i_2,\ldots,i_\ell)} \otimes y_{(i_{\ell+1},i_{\ell+2},\ldots,i_{\ell+m})}\right)\end{aligned} \tag{13.38.10}$$

[479] If we knew that the $\mathbf{k}$-module $T(V)$, endowed with the multiplication $m_{\sqcup\!\sqcup}$, the unit u, the comultiplication $\Delta_{\sqcup\!\sqcup}$ and the counit ϵ, becomes a Hopf algebra with the antipode S, then these five equalities would be saying that $\phi : A^o \to T(V)$ is a Hopf algebra homomorphism.

for every $\ell \in \mathbb{N}$, $m \in \mathbb{N}$ and $(i_1, i_2, \ldots, i_{\ell+m}) \in I^{\ell+m}$. So let us fix $\ell \in \mathbb{N}$, $m \in \mathbb{N}$ and $(i_1, i_2, \ldots, i_{\ell+m}) \in I^{\ell+m}$. Comparing

$$\begin{aligned}
&\left(\phi \circ m_{A^o}\right)\left(y_{(i_1,i_2,\ldots,i_\ell)} \otimes y_{(i_{\ell+1},i_{\ell+2},\ldots,i_{\ell+m})}\right) \\
&= \phi\left(\underbrace{m_{A^o}\left(y_{(i_1,i_2,\ldots,i_\ell)} \otimes y_{(i_{\ell+1},i_{\ell+2},\ldots,i_{\ell+m})}\right)}_{\substack{=y_{(i_1,i_2,\ldots,i_\ell)}y_{(i_{\ell+1},i_{\ell+2},\ldots,i_{\ell+m})} \\ =\sum_{\sigma\in\mathrm{Sh}_{\ell,m}} y_{\left(i_{\sigma(1)},i_{\sigma(2)},\ldots,i_{\sigma(\ell+m)}\right)} \\ \text{(by (1.6.4))}}}\right) = \phi\left(\sum_{\sigma\in\mathrm{Sh}_{\ell,m}} y_{\left(i_{\sigma(1)},i_{\sigma(2)},\ldots,i_{\sigma(\ell+m)}\right)}\right) \\
&= \sum_{\sigma\in\mathrm{Sh}_{\ell,m}} \underbrace{\phi\left(y_{\left(i_{\sigma(1)},i_{\sigma(2)},\ldots,i_{\sigma(\ell+m)}\right)}\right)}_{\substack{=v_{i_{\sigma(1)}}v_{i_{\sigma(2)}}\cdots v_{i_{\sigma(\ell+m)}} \\ \text{(by the definition of } \phi\text{)}}} = \sum_{\sigma\in\mathrm{Sh}_{\ell,m}} v_{i_{\sigma(1)}}v_{i_{\sigma(2)}}\cdots v_{i_{\sigma(\ell+m)}}
\end{aligned}$$

with

$$\begin{aligned}
&\left(m_{\shuffle} \circ (\phi \otimes \phi)\right)\left(y_{(i_1,i_2,\ldots,i_\ell)} \otimes y_{(i_{\ell+1},i_{\ell+2},\ldots,i_{\ell+m})}\right) \\
&= m_{\shuffle}\left(\underbrace{(\phi\otimes\phi)\left(y_{(i_1,i_2,\ldots,i_\ell)} \otimes y_{(i_{\ell+1},i_{\ell+2},\ldots,i_{\ell+m})}\right)}_{=\phi\left(y_{(i_1,i_2,\ldots,i_\ell)}\right)\otimes\phi\left(y_{(i_{\ell+1},i_{\ell+2},\ldots,i_{\ell+m})}\right)}\right) = m_{\shuffle}\left(\underbrace{\phi\left(y_{(i_1,i_2,\ldots,i_\ell)}\right)}_{=v_{i_1}v_{i_2}\cdots v_{i_\ell}} \otimes \underbrace{\phi\left(y_{(i_{\ell+1},i_{\ell+2},\ldots,i_{\ell+m})}\right)}_{=v_{i_{\ell+1}}v_{i_{\ell+2}}\cdots v_{i_{\ell+m}}}\right) \\
&= m_{\shuffle}\left(\left(v_{i_1}v_{i_2}\cdots v_{i_\ell}\right)\otimes\left(v_{i_{\ell+1}}v_{i_{\ell+2}}\cdots v_{i_{\ell+m}}\right)\right) = \left(v_{i_1}v_{i_2}\cdots v_{i_\ell}\right) \shuffle \left(v_{i_{\ell+1}}v_{i_{\ell+2}}\cdots v_{i_{\ell+m}}\right) \\
&= \sum_{\sigma\in\mathrm{Sh}_{\ell,m}} v_{i_{\sigma(1)}}v_{i_{\sigma(2)}}\cdots v_{i_{\sigma(\ell+m)}} \qquad \text{(by the definition of } \shuffle\text{)},
\end{aligned}$$

we obtain $\left(\phi \circ m_{A^o}\right)\left(y_{(i_1,i_2,\ldots,i_\ell)} \otimes y_{(i_{\ell+1},i_{\ell+2},\ldots,i_{\ell+m})}\right) = \left(m_{\shuffle} \circ (\phi\otimes\phi)\right)\left(y_{(i_1,i_2,\ldots,i_\ell)} \otimes y_{(i_{\ell+1},i_{\ell+2},\ldots,i_{\ell+m})}\right)$. Thus, (13.38.10) is proven, and this establishes the equality (13.38.5).

Proof of (13.38.7)*:* Now we must prove the equality (13.38.7). By linearity, it is enough to verify this equality on a **k**-basis of A^o. We will use the basis $\left\{y_{(i_1,i_2,\ldots,i_\ell)}\right\}$; then, we need to check the equality

$$\left((\phi\otimes\phi)\circ\Delta_{A^o}\right)\left(y_{(i_1,i_2,\ldots,i_\ell)}\right) = \left(\Delta_{\shuffle}\circ\phi\right)\left(y_{(i_1,i_2,\ldots,i_\ell)}\right)$$

holds for every $\ell \in \mathbb{N}$ and $(i_1, i_2, \ldots, i_\ell) \in I^\ell$. This equality follows by comparing

$$\begin{aligned}
\left((\phi\otimes\phi)\circ\Delta_{A^o}\right)\left(y_{(i_1,i_2,\ldots,i_\ell)}\right) &= (\phi\otimes\phi)\left(\underbrace{\Delta_{A^o}\left(y_{(i_1,i_2,\ldots,i_\ell)}\right)}_{\substack{=\sum_{j=0}^{\ell} y_{(i_1,\ldots,i_j)}\otimes y_{(i_{j+1},i_{j+2},\ldots,i_\ell)} \\ \text{(by (1.6.1))}}}\right) \\
&= (\phi\otimes\phi)\left(\sum_{j=0}^{\ell} y_{(i_1,\ldots,i_j)}\otimes y_{(i_{j+1},i_{j+2},\ldots,i_\ell)}\right) \\
&= \sum_{j=0}^{\ell} \underbrace{\phi\left(y_{(i_1,\ldots,i_j)}\right)}_{=v_{i_1}v_{i_2}\cdots v_{i_j}} \otimes \underbrace{\phi\left(y_{(i_{j+1},i_{j+2},\ldots,i_\ell)}\right)}_{=v_{i_{j+1}}v_{i_{j+2}}\cdots v_{i_\ell}} = \sum_{j=0}^{\ell}\left(v_{i_1}v_{i_2}\cdots v_{i_j}\right)\otimes\left(v_{i_{j+1}}v_{i_{j+2}}\cdots v_{i_\ell}\right)
\end{aligned}$$

with

$$\begin{aligned}\left(\Delta_{\shuffle}\circ\phi\right)\left(y_{(i_1,i_2,\ldots,i_\ell)}\right) &= \Delta_{\shuffle}\left(\underbrace{\phi\left(y_{(i_1,i_2,\ldots,i_\ell)}\right)}_{=v_{i_1}v_{i_2}\cdots v_{i_\ell}}\right) = \Delta_{\shuffle}\left(v_{i_1}v_{i_2}\cdots v_{i_\ell}\right)\\ &= \sum_{k=0}^{\ell}\left(v_{i_1}v_{i_2}\cdots v_{i_k}\right)\otimes\left(v_{i_{k+1}}v_{i_{k+2}}\cdots v_{i_\ell}\right) \qquad \text{(by the definition of } \Delta_{\shuffle}\text{)}\\ &= \sum_{j=0}^{\ell}\left(v_{i_1}v_{i_2}\cdots v_{i_j}\right)\otimes\left(v_{i_{j+1}}v_{i_{j+2}}\cdots v_{i_\ell}\right).\end{aligned}$$

Thus, (13.38.7) is proven.

Now, we can derive (13.38.3) in a very straightforward way from the commutativity of (13.38.4) using the equalities (13.38.7) and (13.38.5): Consider the diagram

$$\begin{tikzcd}
 & & A^o\otimes A^o \arrow[d, "\phi\otimes\phi"] \arrow[ddll, "\Delta_{A^o}\otimes\Delta_{A^o}"'] \arrow[dddrr, "m_{A^o}"] & & \\
 & & \mathcal{T}\otimes\mathcal{T} \arrow[dl, "\Delta_{\shuffle}\otimes\Delta_{\shuffle}"'] \arrow[ddr, "m_{\shuffle}"'] & & \\
A^o\otimes A^o\otimes A^o\otimes A^o \arrow[r, "\phi\otimes\phi\otimes\phi\otimes\phi"'] \arrow[dd, "\operatorname{id}\otimes T\otimes\operatorname{id}"'] & \mathcal{T}\otimes\mathcal{T}\otimes\mathcal{T}\otimes\mathcal{T} \arrow[dd, "\operatorname{id}\otimes T\otimes\operatorname{id}"'] & & & \\
 & & & \mathcal{T} \arrow[ddl, "\Delta_{\shuffle}"'] & A^o \arrow[l, "\phi"] \arrow[dddll, "\Delta_{A^o}"] \\
A^o\otimes A^o\otimes A^o\otimes A^o \arrow[r, "\phi\otimes\phi\otimes\phi\otimes\phi"] \arrow[ddrr, "m_{A^o}\otimes m_{A^o}"'] & \mathcal{T}\otimes\mathcal{T}\otimes\mathcal{T}\otimes\mathcal{T} \arrow[dr, "m_{\shuffle}\otimes m_{\shuffle}"'] & & & \\
 & & \mathcal{T}\otimes\mathcal{T} & & \\
 & & A^o\otimes A^o \arrow[u, "\phi\otimes\phi"] & &
\end{tikzcd},$$

where $\mathcal{T}$ is shorthand for $T(V)$. The large pentagon in this diagram is commutative (because it is the diagram (13.38.4), which is known to commute), and so are all five quadrilaterals[480]. This does not automatically yield the commutativity of the small pentagon, but it yields that all paths from the $A^o\otimes A^o$ at the top of the diagram to the $\mathcal{T}\otimes\mathcal{T}$ one row above the very bottom give the same map; in particular, we have

$$\left(m_{\shuffle}\otimes m_{\shuffle}\right)\circ\left(\operatorname{id}\otimes T\otimes\operatorname{id}\right)\circ\left(\Delta_{\shuffle}\otimes\Delta_{\shuffle}\right)\circ\left(\phi\otimes\phi\right) = \left(\Delta_{\shuffle}\circ m_{\shuffle}\right)\circ\left(\phi\otimes\phi\right).$$

Thus, the two maps $\left(m_{\shuffle}\otimes m_{\shuffle}\right)\circ\left(\operatorname{id}\otimes T\otimes\operatorname{id}\right)\circ\left(\Delta_{\shuffle}\otimes\Delta_{\shuffle}\right)$ and $\Delta_{\shuffle}\circ m_{\shuffle}$ are equal to each other on the image of the map $\phi\otimes\phi$. Since $a\otimes b$ lies in the image of the map $\phi\otimes\phi$, this yields that these two maps are

[480] In fact:

- the northeastern quadrilateral commutes because of (13.38.5);
- the southeastern quadrilateral commutes because of (13.38.7);
- the northwestern quadrilateral commutes because

$$\begin{aligned}\left(\phi\otimes\phi\otimes\phi\otimes\phi\right)\circ\left(\Delta_{A^o}\otimes\Delta_{A^o}\right) &= \underbrace{\left(\left(\phi\otimes\phi\right)\circ\Delta_{A^o}\right)}_{\substack{=\Delta_{\shuffle}\circ\phi\\ \text{(by (13.38.7))}}}\otimes\underbrace{\left(\left(\phi\otimes\phi\right)\circ\Delta_{A^o}\right)}_{\substack{=\Delta_{\shuffle}\circ\phi\\ \text{(by (13.38.7))}}}\\ &= \left(\Delta_{\shuffle}\circ\phi\right)\otimes\left(\Delta_{\shuffle}\circ\phi\right) = \left(\Delta_{\shuffle}\otimes\Delta_{\shuffle}\right)\circ\left(\phi\otimes\phi\right);\end{aligned}$$

- the southwestern quadrilateral commutes for a similar reason (but using (13.38.5) instead of (13.38.7));
- the western quadrilateral commutes as a consequence of simple linear algebra.

equal to each other on $a \otimes b$. In other words, (13.38.3) holds. As we have said above, this completes our proof of Proposition 1.6.7.

13.39. **Solution to Exercise 1.7.9.** *Solution to Exercise 1.7.9.* We shall use the following simple fact:

Fact A.0: Let α and β be two maps in $\operatorname{Hom}(C, A)$. Let $x \in C$. Let some $k \in \mathbb{N}$, some elements $y_1, y_2, \ldots, y_k \in C$ and some elements $z_1, z_2, \ldots, z_k \in C$ be chosen such that $\Delta(x) = \sum_{p=1}^{k} y_p \otimes z_p$. Then,

$$(\alpha \star \beta)(x) = \sum_{p=1}^{k} \alpha(y_p) \beta(z_p).$$

[*Proof of Fact A.0:* Let m denote the multiplication map $A \otimes A \to A$ of the $\mathbf{k}$-algebra A. The definition of convolution yields $\alpha \star \beta = m \circ (\alpha \otimes \beta) \circ \Delta$. Hence,

$$\begin{aligned}(\alpha \star \beta)(x) &= (m \circ (\alpha \otimes \beta) \circ \Delta)(x) = m\left((\alpha \otimes \beta)\left(\underbrace{\Delta(x)}_{=\sum_{p=1}^{k} y_p \otimes z_p}\right)\right) \\ &= m\left(\underbrace{(\alpha \otimes \beta)\left(\sum_{p=1}^{k} y_p \otimes z_p\right)}_{=\sum_{p=1}^{k} \alpha(y_p) \otimes \beta(z_p)}\right) = m\left(\sum_{p=1}^{k} \alpha(y_p) \otimes \beta(z_p)\right) \\ &= \sum_{p=1}^{k} \alpha(y_p) \beta(z_p) \qquad \text{(by the definition of the map } m\text{)}.\end{aligned}$$

This proves Fact A.0.]

Proof of Proposition 1.7.4. Let g be the map $\sum_{q \in Q} f_q$.[481] Thus, g is a map $C \to A$. Moreover, $g = \sum_{q \in Q} f_q$. Thus, each $x \in C$ satisfies

$$g(x) = \left(\sum_{q \in Q} f_q\right)(x) = \sum_{q \in Q} f_q(x) \tag{13.39.1}$$

(by the definition of $\sum_{q \in Q} f_q$).

Now, let $c \in C$, $d \in C$, $\lambda \in \mathbf{k}$ and $\mu \in \mathbf{k}$ be arbitrary. Then, (13.39.1) (applied to $x = c$) yields

$$g(c) = \sum_{q \in Q} f_q(c).$$

Moreover, the family $(f_q)_{q \in Q} \in (\operatorname{Hom}(C, A))^Q$ is pointwise finitely supported. In other words, for each $x \in C$, the family $(f_q(x))_{q \in Q} \in A^Q$ of elements of A is finitely supported (by the definition of "pointwise finitely supported"). Applying this to $x = c$, we conclude that the family $(f_q(c))_{q \in Q} \in A^Q$ of elements of A is finitely supported. In other words, all but finitely many $q \in Q$ satisfy $f_q(c) = 0$. Hence, all but finitely many $q \in Q$ satisfy $\lambda f_q(c) = 0$ (since every $q \in Q$ that satisfies $f_q(c) = 0$ must also satisfy $\lambda \underbrace{f_q(c)}_{=0} = 0$). In other words, the family $(\lambda f_q(c))_{q \in Q} \in A^Q$ is finitely supported. The same argument (applied to d and μ instead of c and λ) shows that the family $(\mu f_q(d))_{q \in Q} \in A^Q$ of elements of A is finitely supported.

[481] This map is well-defined, since the family $(f_q)_{q \in Q}$ is pointwise finitely supported.

Recall that sums of finitely supported families satisfy the same rules as finite sums. Since the families $\left(\lambda f_q\left(c\right)\right)_{q\in Q}$ and $\left(\mu f_q\left(d\right)\right)_{q\in Q}$ are finitely supported, we thus have

$$\sum_{q\in Q}\lambda f_q\left(c\right)+\sum_{q\in Q}\mu f_q\left(d\right)=\sum_{q\in Q}\left(\lambda f_q\left(c\right)+\mu f_q\left(d\right)\right) \tag{13.39.2}$$

(and in particular, the family $\left(\lambda f_q\left(c\right)+\mu f_q\left(d\right)\right)_{q\in Q}$ is finitely supported). For the same reason, we have

$$\lambda\sum_{q\in Q}f_q\left(c\right)=\sum_{q\in Q}\lambda f_q\left(c\right) \tag{13.39.3}$$

(since the family $\left(f_q\left(c\right)\right)_{q\in Q}$ is finitely supported) and

$$\mu\sum_{q\in Q}f_q\left(d\right)=\sum_{q\in Q}\mu f_q\left(d\right) \tag{13.39.4}$$

(for similar reasons).

Now, (13.39.1) (applied to $x=\lambda c+\mu d$) yields

$$\begin{aligned} g\left(\lambda c+\mu d\right)&=\sum_{q\in Q}\underbrace{f_q\left(\lambda c+\mu d\right)}_{\substack{=\lambda f_q(c)+\mu f_q(d)\\ (\text{since } f_q\in\operatorname{Hom}(C,A))}}=\sum_{q\in Q}\left(\lambda f_q\left(c\right)+\mu f_q\left(d\right)\right)\\ &=\underbrace{\sum_{q\in Q}\lambda f_q\left(c\right)}_{\substack{=\lambda\sum_{q\in Q}f_q(c)\\ (\text{by (13.39.3)})}}+\underbrace{\sum_{q\in Q}\mu f_q\left(d\right)}_{\substack{=\mu\sum_{q\in Q}f_q(d)\\ (\text{by (13.39.4)})}}\qquad(\text{by (13.39.2)})\\ &=\lambda\sum_{q\in Q}f_q\left(c\right)+\mu\sum_{q\in Q}f_q\left(d\right). \end{aligned}$$

Comparing this with

$$\lambda\underbrace{g\left(c\right)}_{=\sum_{q\in Q}f_q(c)}+\mu\underbrace{g\left(d\right)}_{\substack{=\sum_{q\in Q}f_q(d)\\ (\text{by (13.39.1) (applied to } x=d))}}=\lambda\sum_{q\in Q}f_q\left(c\right)+\mu\sum_{q\in Q}f_q\left(d\right),$$

we obtain $g\left(\lambda c+\mu d\right)=\lambda g\left(c\right)+\mu g\left(d\right)$.

Now, forget that we fixed c, d, λ and μ. We thus have proven that $g\left(\lambda c+\mu d\right)=\lambda g\left(c\right)+\mu g\left(d\right)$ for every $c\in C$, $d\in C$, $\lambda\in\mathbf{k}$ and $\mu\in\mathbf{k}$. In other words, the map g is $\mathbf{k}$-linear. In other words, $g\in\operatorname{Hom}\left(C,A\right)$. Thus, $\sum_{q\in Q}f_q=g\in\operatorname{Hom}\left(C,A\right)$. This proves Proposition 1.7.4. □

Proof of Proposition 1.7.5. Fix $x\in C$. The family $\left(f_q\left(x\right)\right)_{q\in Q}\in A^Q$ of elements of A is finitely supported (since the family $\left(f_q\right)_{q\in Q}\in\left(\operatorname{Hom}\left(C,A\right)\right)^Q$ is pointwise finitely supported). In other words, all but finitely many $q\in Q$ satisfy $f_q\left(x\right)=0$. In other words, there exists a finite subset Q_1 of Q such that

$$\text{every } q\in Q\setminus Q_1 \text{ satisfies } f_q\left(x\right)=0. \tag{13.39.5}$$

Similarly, there exists a finite subset Q_2 of Q such that

$$\text{every } q\in Q\setminus Q_2 \text{ satisfies } g_q\left(x\right)=0. \tag{13.39.6}$$

Consider these two finite subsets Q_1 and Q_2.

The set $Q_1\cup Q_2$ is finite (since it is the union of the two finite sets Q_1 and Q_2). Thus, all but finitely many $q\in Q$ satisfy $q\in Q\setminus\left(Q_1\cup Q_2\right)$.

We have $Q\setminus\underbrace{\left(Q_1\cup Q_2\right)}_{\supset Q_1}\subset Q\setminus Q_1$ and $Q\setminus\underbrace{\left(Q_1\cup Q_2\right)}_{\supset Q_2}\subset Q\setminus Q_2$.

Also, $Q_1\cup Q_2$ is a subset of Q (since both Q_1 and Q_2 are subsets of Q). In other words, $Q_1\cup Q_2\subset Q$.

Every $q\in Q\setminus\left(Q_1\cup Q_2\right)$ satisfies

$$\left(f_q+g_q\right)\left(x\right)=\underbrace{f_q\left(x\right)}_{\substack{=0\\ (\text{by (13.39.5)}\\ (\text{since } q\in Q\setminus(Q_1\cup Q_2)\subset Q\setminus Q_1))}}+\underbrace{g_q\left(x\right)}_{\substack{=0\\ (\text{by (13.39.6)}\\ (\text{since } q\in Q\setminus(Q_1\cup Q_2)\subset Q\setminus Q_2))}}=0. \tag{13.39.7}$$

Hence, all but finitely many $q \in Q$ satisfy $(f_q + g_q)(x) = 0$ (since all but finitely many $q \in Q$ satisfy $q \in Q \setminus (Q_1 \cup Q_2)$). In other words,

$$\text{the family } \left((f_q + g_q)(x)\right)_{q \in Q} \in A^Q \text{ is finitely supported.} \tag{13.39.8}$$

The set Q is the union of its two disjoint subsets $Q_1 \cup Q_2$ and $Q \setminus (Q_1 \cup Q_2)$ (since $Q_1 \cup Q_2 \subset Q$). Thus, the sum $\sum_{q \in Q} f_q(x)$ can be split as follows:

$$\begin{aligned}\sum_{q \in Q} f_q(x) &= \sum_{q \in Q_1 \cup Q_2} f_q(x) + \sum_{q \in Q \setminus (Q_1 \cup Q_2)} \underbrace{f_q(x)}_{\substack{=0 \\ \text{(by (13.39.5)} \\ \text{(since } q \in Q \setminus (Q_1 \cup Q_2) \subset Q \setminus Q_1))}} \\ &= \sum_{q \in Q_1 \cup Q_2} f_q(x) + \underbrace{\sum_{q \in Q \setminus (Q_1 \cup Q_2)} 0}_{=0} = \sum_{q \in Q_1 \cup Q_2} f_q(x). \end{aligned} \tag{13.39.9}$$

A similar argument (using (13.39.6) instead of (13.39.5)) yields

$$\sum_{q \in Q} g_q(x) = \sum_{q \in Q_1 \cup Q_2} g_q(x).$$

Adding this equality to (13.39.9), we obtain

$$\sum_{q \in Q} f_q(x) + \sum_{q \in Q} g_q(x) = \sum_{q \in Q_1 \cup Q_2} f_q(x) + \sum_{q \in Q_1 \cup Q_2} g_q(x). \tag{13.39.10}$$

But recall again that the set Q is the union of its two disjoint subsets $Q_1 \cup Q_2$ and $Q \setminus (Q_1 \cup Q_2)$. Hence,

$$\begin{aligned}\sum_{q \in Q} (f_q + g_q)(x) &= \sum_{q \in Q_1 \cup Q_2} \underbrace{(f_q + g_q)(x)}_{=f_q(x)+g_q(x)} + \sum_{q \in Q \setminus (Q_1 \cup Q_2)} \underbrace{(f_q + g_q)(x)}_{\substack{=0 \\ \text{(by (13.39.7))}}} \\ &= \sum_{q \in Q_1 \cup Q_2} (f_q(x) + g_q(x)) + \underbrace{\sum_{q \in Q \setminus (Q_1 \cup Q_2)} 0}_{=0} = \sum_{q \in Q_1 \cup Q_2} (f_q(x) + g_q(x)) \\ &= \sum_{q \in Q_1 \cup Q_2} f_q(x) + \sum_{q \in Q_1 \cup Q_2} g_q(x) \qquad \left(\begin{array}{c}\text{here, we have manipulated a finite sum} \\ \text{(since } Q_1 \cup Q_2 \text{ is a finite set)}\end{array}\right) \\ &= \sum_{q \in Q} f_q(x) + \sum_{q \in Q} g_q(x) \qquad \text{(by (13.39.10))}. \end{aligned} \tag{13.39.11}$$

Now, let us forget that we fixed x. We thus have proven that every $x \in C$ satisfies (13.39.8) and (13.39.11).

In particular, every $x \in C$ satisfies (13.39.8). In other words, for each $x \in C$, the family $\left((f_q + g_q)(x)\right)_{q \in Q} \in A^Q$ of elements of A is finitely supported. In other words, the family $(f_q + g_q)_{q \in Q} \in (\operatorname{Hom}(C, A))^Q$ is pointwise finitely supported (by the definition of "pointwise finitely supported").

Hence, the sum $\sum_{q \in Q} (f_q + g_q)$ is well-defined. Also, the sum $\sum_{q \in Q} f_q$ is well-defined (since the family $(f_q)_{q \in Q}$ is pointwise finitely supported). Similarly, the sum $\sum_{q \in Q} g_q$ is well-defined.

Moreover, each $x \in C$ satisfies

$$\begin{aligned}\left(\sum_{q \in Q} (f_q + g_q)\right)(x) &= \sum_{q \in Q} (f_q + g_q)(x) = \underbrace{\sum_{q \in Q} f_q(x)}_{=\left(\sum_{q \in Q} f_q\right)(x)} + \underbrace{\sum_{q \in Q} g_q(x)}_{=\left(\sum_{q \in Q} g_q\right)(x)} \qquad \text{(by (13.39.11))} \\ &= \left(\sum_{q \in Q} f_q\right)(x) + \left(\sum_{q \in Q} g_q\right)(x) = \left(\sum_{q \in Q} f_q + \sum_{q \in Q} g_q\right)(x). \end{aligned}$$

In other words, $\sum_{q \in Q} (f_q + g_q) = \sum_{q \in Q} f_q + \sum_{q \in Q} g_q$. In other words, $\sum_{q \in Q} f_q + \sum_{q \in Q} g_q = \sum_{q \in Q} (f_q + g_q)$. This completes the proof of Proposition 1.7.5. □

Proof of Proposition 1.7.6. Let $x \in C$. Write the element $\Delta(x) \in C \otimes C$ in the form $\Delta(x) = \sum_{p=1}^{k} y_p \otimes z_p$ for some $k \in \mathbb{N}$, some elements $y_1, y_2, \ldots, y_k \in C$ and some elements $z_1, z_2, \ldots, z_k \in C$. (This is possible, because $\Delta(x)$ can be written as a sum of pure tensors[482].)

For each $p \in \{1, 2, \ldots, k\}$, there exists a finite subset Q_p of Q such that

$$\text{every } q \in Q \setminus Q_p \text{ satisfies } f_q(y_p) = 0 \tag{13.39.12}$$

[483]. Consider this Q_p.

Define a subset Q' of Q by $Q' = Q_1 \cup Q_2 \cup \cdots \cup Q_k$. Thus, Q' is the union of the k finite sets $Q_1, Q_2, \ldots, Q_k$. Therefore, Q' itself is a finite set (since a union of k finite sets is always finite).

Every $q \in Q \setminus Q'$ and $p \in \{1, 2, \ldots, k\}$ satisfy $f_q(y_p) = 0$ [484].

We have thus constructed a finite subset Q' of Q with the property that every $q \in Q \setminus Q'$ and $p \in \{1, 2, \ldots, k\}$ satisfy

$$f_q(y_p) = 0. \tag{13.39.13}$$

Similarly, we can construct a finite subset R' of R with the property that every $r \in R \setminus R'$ and $p \in \{1, 2, \ldots, k\}$ satisfy

$$g_r(z_p) = 0. \tag{13.39.14}$$

Consider this R'.

The set $Q' \times R'$ is a Cartesian product of two finite sets (since Q' and R' are finite sets), and thus is itself finite. Hence, all but finitely many $(q, r) \in Q \times R$ satisfy $(q, r) \in (Q \times R) \setminus (Q' \times R')$.

We shall now show that each $(q, r) \in (Q \times R) \setminus (Q' \times R')$ satisfies $(f_q \star g_r)(x) = 0$.

Indeed, fix $(q, r) \in (Q \times R) \setminus (Q' \times R')$. Thus, $(q, r) \in Q \times R$ and $(q, r) \notin Q' \times R'$.

We are in one of the following two cases:

Case 1: We have $q \in Q'$.

Case 2: We have $q \notin Q'$.

Let us first consider Case 1. In this case, we have $q \in Q'$. If we had $r \in R'$, we thus would have $(q, r) \in Q' \times R'$ (since $q \in Q'$ and $r \in R'$), which would contradict $(q, r) \notin Q' \times R'$. Hence, we cannot have $r \in R'$. In other words, we have $r \notin R'$. Combining $r \in R$ with $r \notin R'$, we obtain $r \in R \setminus R'$. Now, Fact A.0 (applied to $\alpha = f_q$ and $\beta = g_r$) yields

$$(f_q \star g_r)(x) = \sum_{p=1}^{k} f_q(y_p) \underbrace{g_r(z_p)}_{\substack{=0 \\ \text{(by (13.39.14))}}} = \sum_{p=1}^{k} f_q(y_p) 0 = 0.$$

Thus, $(f_q \star g_r)(x) = 0$ is proven in Case 1.

Let us now consider Case 2. In this case, we have $q \notin Q'$. Combining $q \in Q$ with $q \notin Q'$, we obtain $q \in Q \setminus Q'$. Now, Fact A.0 (applied to $\alpha = f_q$ and $\beta = g_r$) yields

$$(f_q \star g_r)(x) = \sum_{p=1}^{k} \underbrace{f_q(y_p)}_{\substack{=0 \\ \text{(by (13.39.13))}}} g_r(z_p) = \sum_{p=1}^{k} 0 g_r(z_p) = 0.$$

Thus, $(f_q \star g_r)(x) = 0$ is proven in Case 2.

We have thus proven $(f_q \star g_r)(x) = 0$ in both Cases 1 and 2. Hence, $(f_q \star g_r)(x) = 0$ always holds.

Now, forget that we fixed (q, r). We thus have proven that

$$\text{each } (q, r) \in (Q \times R) \setminus (Q' \times R') \text{ satisfies } (f_q \star g_r)(x) = 0. \tag{13.39.15}$$

[482] This is because any tensor in $C \otimes C$ can be written as a sum of pure tensors.

[483] *Proof:* Let $p \in \{1, 2, \ldots, k\}$. The family $(f_q)_{q \in Q} \in (\operatorname{Hom}(C, A))^Q$ is pointwise finitely supported. Hence, the family $(f_q(y_p))_{q \in Q} \in A^Q$ is finitely supported. In other words, all but finitely many $q \in Q$ satisfy $f_q(y_p) = 0$. In other words, there exists a finite subset Q_p of Q such that every $q \in Q \setminus Q_p$ satisfies $f_q(y_p) = 0$. Qed.

[484] *Proof of (13.39.13):* Let $q \in Q \setminus Q'$ and $p \in \{1, 2, \ldots, k\}$.

We have $Q_p \subset Q_1 \cup Q_2 \cup \cdots \cup Q_k = Q'$ (since $Q' = Q_1 \cup Q_2 \cup \cdots \cup Q_k$), so that $Q \setminus \underbrace{Q_p}_{\subset Q'} \supset Q \setminus Q'$. Hence, $Q \setminus Q' \subset Q \setminus Q_p$. Now, (13.39.12) yields $f_q(y_p) = 0$ (since $q \in Q \setminus Q' \subset Q \setminus Q_p$). This proves (13.39.13).

Thus, all but finitely many $(q,r) \in Q \times R$ satisfy $(f_q \star g_r)(x) = 0$ (since all but finitely many $(q,r) \in Q \times R$ satisfy $(q,r) \in (Q \times R) \setminus (Q' \times R')$). In other words,

$$\text{the family } \left((f_q \star g_r)(x)\right)_{(q,r)\in Q\times R} \in A^{Q\times R} \text{ is finitely supported.} \tag{13.39.16}$$

Define a map $F : C \to A$ by $F = \sum_{q\in Q} f_q$. (This is well-defined, since the family $(f_q)_{q\in Q} \in (\operatorname{Hom}(C,A))^Q$ is pointwise finitely supported.)

Define a map $G : C \to A$ by $G = \sum_{r\in R} g_r$. (This is well-defined, since the family $(g_r)_{r\in R} \in (\operatorname{Hom}(C,A))^R$ is pointwise finitely supported.)

We have $F = \sum_{q\in Q} f_q \in \operatorname{Hom}(C,A)$ (by Proposition 1.7.4) and $G \in \operatorname{Hom}(C,A)$ (for similar reasons). Hence, the map $F \star G \in \operatorname{Hom}(C,A)$ is well-defined.

Define a further map $F' \in \operatorname{Hom}(C,A)$ by $F' = \sum_{q\in Q'} f_q$. Notice that this is a finite sum, since Q' is a finite set.

Define a further map $G' \in \operatorname{Hom}(C,A)$ by $G' = \sum_{r\in R'} g_r$. Notice that this is a finite sum, since R' is a finite set.

From $F' = \sum_{q\in Q'} f_q$ and $G' = \sum_{r\in R'} g_r$, we obtain[485]

$$\begin{aligned} F' \star G' &= \left(\sum_{q\in Q'} f_q\right) \star \left(\sum_{r\in R'} g_r\right) \\ &= \underbrace{\sum_{q\in Q'}\sum_{r\in R'}}_{=\sum_{(q,r)\in Q'\times R'}} (f_q \star g_r) = \sum_{(q,r)\in Q'\times R'} (f_q \star g_r) . \end{aligned} \tag{13.39.17}$$

But every $p \in \{1,2,\ldots,k\}$ satisfies

$$F'(y_p) = F(y_p) \tag{13.39.18}$$

[486] and

$$G'(z_p) = G(z_p) \tag{13.39.19}$$

(for similar reasons). Hence,

$$(F' \star G')(x) = (F \star G)(x) \tag{13.39.20}$$

[485] The following manipulations of sums are legitimate, since all the sums involved are finite (because Q' and R' are finite sets).

[486] *Proof of (13.39.18):* Let $p \in \{1,2,\ldots,k\}$. Recall that $Q' \subset Q$; thus, the set Q is the union of its two disjoint subsets Q' and $Q \setminus Q'$.

From $F = \sum_{q\in Q} f_q$, we obtain

$$\begin{aligned} F(y_p) &= \left(\sum_{q\in Q} f_q\right)(y_p) = \sum_{q\in Q} f_q(y_p) = \sum_{q\in Q'} f_q(y_p) + \sum_{q\in Q\setminus Q'} \underbrace{f_q(y_p)}_{\substack{=0 \\ \text{(by (13.39.13))}}} \\ &\qquad \left(\text{since the set } Q \text{ is the union of its disjoint subsets } Q' \text{ and } Q\setminus Q'\right) \\ &= \sum_{q\in Q'} f_q(y_p) + \underbrace{\sum_{q\in Q\setminus Q'} 0}_{=0} = \sum_{q\in Q'} f_q(y_p) . \end{aligned}$$

Comparing this with $\underbrace{F'}_{=\sum_{q\in Q'} f_q}(y_p) = \left(\sum_{q\in Q'} f_q\right)(y_p) = \sum_{q\in Q'} f_q(y_p)$, we obtain $F'(y_p) = F(y_p)$. This proves (13.39.18).

[487]. Thus,

$$(F \star G)(x) = \underbrace{(F' \star G')}_{\substack{=\sum_{(q,r)\in Q'\times R'}(f_q \star g_r) \\ \text{(by (13.39.17))}}}(x) = \left(\sum_{(q,r)\in Q'\times R'}(f_q \star g_r)\right)(x)$$

$$= \sum_{(q,r)\in Q'\times R'}(f_q \star g_r)(x). \tag{13.39.22}$$

On the other hand, $\underbrace{Q'}_{\subset Q} \times \underbrace{R'}_{\subset R} \subset Q \times R$. Hence, the set $Q \times R$ is the union of its two disjoint subsets $Q' \times R'$ and $(Q \times R) \setminus (Q' \times R')$.

But the sum $\sum_{(q,r)\in Q\times R}(f_q \star g_r)(x)$ is well-defined (since the family $((f_q \star g_r)(x))_{(q,r)\in Q\times R} \in A^{Q\times R}$ is finitely supported). Since the set $Q\times R$ is the union of its two disjoint subsets $Q'\times R'$ and $(Q \times R)\setminus(Q' \times R')$, we can split this sum as follows:

$$\sum_{(q,r)\in Q\times R}(f_q \star g_r)(x)$$

$$= \sum_{(q,r)\in Q'\times R'}(f_q \star g_r)(x) + \sum_{(q,r)\in (Q\times R)\setminus(Q'\times R')}\underbrace{(f_q \star g_r)(x)}_{\substack{=0 \\ \text{(by (13.39.15))}}}$$

$$= \sum_{(q,r)\in Q'\times R'}(f_q \star g_r)(x) + \underbrace{\sum_{(q,r)\in (Q\times R)\setminus(Q'\times R')} 0}_{=0} = \sum_{(q,r)\in Q'\times R'}(f_q \star g_r)(x)$$

$$= \left(\underbrace{F}_{=\sum_{q\in Q} f_q} \star \underbrace{G}_{=\sum_{r\in R} g_r}\right)(x) \qquad \text{(by (13.39.22))}$$

$$= \left(\left(\sum_{q\in Q} f_q\right) \star \left(\sum_{r\in R} g_r\right)\right)(x). \tag{13.39.23}$$

Now, forget that we fixed x. We thus have shown that each $x \in C$ satisfies (13.39.16) and (13.39.23).

In particular, each $x \in C$ satisfies (13.39.16). In other words, for each $x \in C$, the family $((f_q \star g_r)(x))_{(q,r)\in Q\times R} \in A^{Q\times R}$ is finitely supported. In other words, the family $(f_q \star g_r)_{(q,r)\in Q\times R} \in (\operatorname{Hom}(C, A))^{Q\times R}$ is pointwise finitely supported (by the definition of "pointwise finitely supported"). Hence, the sum $\sum_{(q,r)\in Q\times R}(f_q \star g_r)$ is well-defined. For each $x \in C$, we have

$$\left(\sum_{(q,r)\in Q\times R}(f_q \star g_r)\right)(x) = \sum_{(q,r)\in Q\times R}(f_q \star g_r)(x) = \left(\left(\sum_{q\in Q} f_q\right) \star \left(\sum_{r\in R} g_r\right)\right)(x)$$

[487] *Proof of (13.39.20):* Fact A.0 (applied to $\alpha = F$ and $\beta = G$) yields

$$(F \star G)(x) = \sum_{p=1}^{k} F(y_p) G(z_p). \tag{13.39.21}$$

Fact A.0 (applied to $\alpha = F'$ and $\beta = G'$) yields

$$(F' \star G')(x) = \sum_{p=1}^{k} \underbrace{F'(y_p)}_{\substack{=F(y_p) \\ \text{(by (13.39.18))}}} \underbrace{G'(z_p)}_{\substack{=G(z_p) \\ \text{(by (13.39.19))}}} = \sum_{p=1}^{k} F(y_p) G(z_p).$$

Comparing this with (13.39.21), we obtain $(F' \star G')(x) = (F \star G)(x)$.

(by (13.39.23)). In other words, we have

$$\sum_{(q,r)\in Q\times R} (f_q \star g_r) = \left(\sum_{q\in Q} f_q\right) \star \left(\sum_{r\in R} g_r\right).$$

This completes the proof of Proposition 1.7.6. □

We shall delay the proof of Proposition 1.7.7 until after Proposition 1.7.8 is proven; the reason is that Proposition 1.7.8 yields a quick shortcut to Proposition 1.7.7.

Proof of Proposition 1.7.8. Let $x \in C$. Write the element $\Delta(x) \in C \otimes C$ in the form $\Delta(x) = \sum_{p=1}^{k} y_p \otimes z_p$ for some $k \in \mathbb{N}$, some elements $y_1, y_2, \ldots, y_k \in C$ and some elements $z_1, z_2, \ldots, z_k \in C$. (This is possible, because $\Delta(x)$ can be written as a sum of pure tensors[488].)

For each $p \in \{1, 2, \ldots, k\}$, there exists a finite subset Q_p of Q such that

$$\text{every } q \in Q \setminus Q_p \text{ satisfies } f_q(y_p) = 0 \tag{13.39.24}$$

[489]. Consider this Q_p.

Define a subset Q' of Q by $Q' = Q_1 \cup Q_2 \cup \cdots \cup Q_k$. Thus, Q' is the union of the k finite sets $Q_1, Q_2, \ldots, Q_k$. Hence, Q' itself a finite set. Hence, all but finitely many $q \in Q$ satisfy $q \in Q \setminus Q'$.

Every $q \in Q \setminus Q'$ and $p \in \{1, 2, \ldots, k\}$ satisfy

$$f_q(y_p) = 0 \tag{13.39.25}$$

[490].

Now, fix $q \in Q \setminus Q'$. Fact A.0 (applied to $\alpha = f_q$ and $\beta = g_q$) yields

$$(f_q \star g_q)(x) = \sum_{p=1}^{k} \underbrace{f_q(y_p)}_{\substack{=0 \\ \text{(by (13.39.25))}}} g_q(z_p) = \sum_{p=1}^{k} 0 g_q(z_p) = 0.$$

Now, forget that we fixed q. We thus have proven that each $q \in Q \setminus Q'$ satisfies $(f_q \star g_q)(x) = 0$. Hence, all but finitely many $q \in Q$ satisfy $(f_q \star g_q)(x) = 0$ (since all but finitely many $q \in Q$ satisfy $q \in Q \setminus Q'$). In other words, the family $((f_q \star g_q)(x))_{q\in Q} \in A^Q$ is finitely supported.

Now, forget that we fixed x. We thus have shown that for each $x \in C$, the family $((f_q \star g_q)(x))_{q\in Q} \in A^Q$ is finitely supported. In other words, the family $(f_q \star g_q)_{q\in Q} \in (\operatorname{Hom}(C, A))^Q$ is pointwise finitely supported. This proves Proposition 1.7.8. □

Proof of Proposition 1.7.7. Let $\mathfrak{i}$ be the unity of the $\mathbf{k}$-algebra $(\operatorname{Hom}(C, A), \star)$. (This $\mathfrak{i}$ is the map $u_A \circ \epsilon_C : C \to A$; but this does not matter to us.)

Applying Proposition 1.7.8 to the family $(g_q)_{q\in Q} = (\lambda_q \mathfrak{i})_{q\in Q}$, we conclude that the family $(f_q \star (\lambda_q \mathfrak{i}))_{q\in Q} \in (\operatorname{Hom}(C, A))^Q$ is pointwise finitely supported. Since each $q \in Q$ satisfies

$$f_q \star (\lambda_q \mathfrak{i}) = \lambda_q \cdot \underbrace{(f_q \star \mathfrak{i})}_{\substack{=f_q \\ \text{(since } \mathfrak{i} \text{ is the unity of} \\ \text{the } \mathbf{k}\text{-algebra } (\operatorname{Hom}(C,A),\star))}} = \lambda_q f_q,$$

this rewrites as follows: The family $(\lambda_q f_q)_{q\in Q} \in (\operatorname{Hom}(C, A))^Q$ is pointwise finitely supported. This proves Proposition 1.7.7. □

We have now proven all five Propositions 1.7.4, 1.7.5, 1.7.6, 1.7.7 and 1.7.8. Thus, Exercise 1.7.9 is solved.

[*Remark:* We can re-interpret the concept of "pointwise finitely supported" families $(f_q)_{q\in Q} \in (\operatorname{Hom}(C, A))^Q$ and their sums $\sum_{q\in Q} f_q$ in topological terms. To that end, we shall use the concept of a "net" (see, e.g., https://en.wikipedia.org/wiki/Net_(mathematics) or [219, §4] for an introduction).

[488] This is because any tensor in $C \otimes C$ can be written as a sum of pure tensors.

[489] This can be shown in the same way as we did it during the proof of Proposition 1.7.6.

[490] This can be shown in the same way as we did it during the proof of Proposition 1.7.6.

Recall that a *preordered set* means a set Z equipped with a preorder relation (i.e., a binary relation on Z that is both reflexive and transitive[491]). This preorder relation is commonly denoted by $\leq$. A nonempty preordered set Z is said to be a *directed set* if its preorder relation $\leq$ has the property that every two elements $x \in Z$ and $y \in Z$ have an upper bound (i.e., some $z \in Z$ satisfying $x \leq z$ and $y \leq z$). If Z is a preordered set, and if $\mathcal{A}(z)$ is a logical statement for each $z \in Z$, then we say that "$\mathcal{A}(z)$ holds *for all sufficiently high* $z \in Z$" if and only if there exists a $w \in Z$ such that every $z \in Z$ satisfying $w \leq z$ satisfies $\mathcal{A}(z)$.

Two important examples of directed sets are the following:

- The set $\mathbb{N}$, equipped with the usual less-or-equal relation $\leq$, is a directed set. This directed set will simply be called $\mathbb{N}$.
- If Q is any set, then the set $\mathcal{P}_{\text{fin}}(Q)$ of all finite subsets of Q is naturally a directed set: Its preorder relation $\leq$ is defined to be the subset relation $\subset$. Every two elements $x \in \mathcal{P}_{\text{fin}}(Q)$ and $y \in \mathcal{P}_{\text{fin}}(Q)$ clearly have an upper bound (for example, $x \cup y$). This directed set will simply be called $\mathcal{P}_{\text{fin}}(Q)$.

A *net* in a set X is defined to be a family $(x_z)_{z\in Z} \in X^Z$, where Z is some directed set.

If X is a topological space, if $x \in X$, and if $(x_z)_{z\in Z} \in X^Z$ is a net in X, then the net $(x_z)_{z\in Z}$ is said to *converge* to x if and only if for each neighborhood U of x, we have

$$(x_z \in U \text{ for all sufficiently high } z \in Z).$$

Thus, in any topological space, we have defined the notion of a convergent net. This notion generalizes the notion of a convergent sequence (indeed, convergent sequences are precisely the same as convergent nets whose indexing set Z is the directed set $\mathbb{N}$). But it is, in a sense, a more natural notion than the latter: Unlike the latter, it characterizes the topological space. That is, we can define a topological space on a set Y by specifying which nets in Y converge to which elements of Y (provided that this specification satisfies certain axioms); but we cannot (in general) define a topological space on a set Y by specifying which sequences in Y converge to which elements of Y.

If a net $(x_z)_{z\in Z} \in X^Z$ in a topological space X converges to an element $x \in X$, then x is called a *limit* of $(x_z)_{z\in Z}$. If X is Hausdorff, then any convergent net $(x_z)_{z\in Z} \in X^Z$ has only one limit, and so we can call this limit "*the* limit" of $(x_z)_{z\in Z}$.

If X is a topological space with the discrete topology, then convergence of nets can be described very simply: A net $(x_z)_{z\in Z} \in X^Z$ in a discrete topological space X converges to an element $x \in X$ if and only if we have ($x_z = x$ for all sufficiently high $z \in Z$). This behavior is also called *stabilization*: i.e., we say that a net $(x_z)_{z\in Z} \in X^Z$ *stabilizes* to an element $x \in X$ if and only if we have ($x_z = x$ for all sufficiently high $z \in Z$).

Let us equip the set A with the discrete topology. Thus, A becomes a topological $\mathbf{k}$-algebra (because equipping any $\mathbf{k}$-algebra with the discrete topology results in a topological $\mathbf{k}$-algebra).

Now, we equip the set $\operatorname{Hom}(C, A)$ with a topology, which can be defined in any of the following two ways:

- It is the unique topology on the set $\operatorname{Hom}(C, A)$ that has the following property: A net $(f_z)_{z\in Z}$ of maps $f_z \in \operatorname{Hom}(C, A)$ converges to a map $f \in \operatorname{Hom}(C, A)$ in this topology if and only if for each $c \in C$, the net $(f_z(c))_{z\in Z}$ in A stabilizes to $f(c)$.
- Alternatively, we can define the topology on $\operatorname{Hom}(C, A)$ in the usual way (i.e., via open sets): The set A^C of all maps from C to A is equipped with a product topology (since it is the product $\prod_{c\in C} A$). The set $\operatorname{Hom}(C, A)$ thus also gets a topology, being a subset of A^C.

These two definitions give rise to the same topology. This topology is called the *topology of pointwise convergence*. We consider $\operatorname{Hom}(C, A)$ to be equipped with this topology from now on. This topology allows us to work with limits of nets in $\operatorname{Hom}(C, A)$ (as long as these nets converge), since the topological space $\operatorname{Hom}(C, A)$ is Hausdorff.

Proposition 13.39.1. *The* $\mathbf{k}$*-algebra* $(\operatorname{Hom}(C, A), \star)$ *is a topological* $\mathbf{k}$*-algebra. That is, the maps*

$$\begin{aligned} \operatorname{Hom}(C, A) \times \operatorname{Hom}(C, A) &\to \operatorname{Hom}(C, A), && (f, g) \mapsto f + g, \\ \operatorname{Hom}(C, A) &\to \operatorname{Hom}(C, A), && f \mapsto -f, \\ \operatorname{Hom}(C, A) \times \operatorname{Hom}(C, A) &\to \operatorname{Hom}(C, A), && (f, g) \mapsto f \star g, \\ \mathbf{k} \times \operatorname{Hom}(C, A) &\to \operatorname{Hom}(C, A), && (\lambda, f) \mapsto \lambda f \end{aligned}$$

are continuous (where the topology on $\mathbf{k}$ *is the discrete topology).*

[491] but (unlike a partial order) not necessarily antisymmetric

We omit the proof of Proposition 13.39.1, since we shall not use it; it is not hard to prove with some standard techniques from point-set topology. But let us see how it allows us to re-interpret pointwise finitely supported families:

- Any family $\left(a_q\right)_{q\in Q} \in A^Q$ of elements of A gives rise to a net $\left(\sum_{q\in K} a_q\right)_{K\in\mathcal{P}_{\operatorname{fin}}(Q)} \in A^{\mathcal{P}_{\operatorname{fin}}(Q)}$ in A (which consists of all sums of finite subfamilies of $\left(a_q\right)_{q\in Q}$). It is easy to see that the family $\left(a_q\right)_{q\in Q}$ is finitely supported if and only if the net $\left(\sum_{q\in K} a_q\right)_{K\in\mathcal{P}_{\operatorname{fin}}(Q)} \in A^{\mathcal{P}_{\operatorname{fin}}(Q)}$ converges in the discrete space A. In this case, the limit of the net is precisely $\sum_{q\in Q} a_q$.
- Any family $\left(f_q\right)_{q\in Q} \in \left(\operatorname{Hom}(C,A)\right)^Q$ of elements of $\operatorname{Hom}(C,A)$ gives rise to a net $\left(\sum_{q\in K} f_q\right)_{K\in\mathcal{P}_{\operatorname{fin}}(Q)} \in \left(\operatorname{Hom}(C,A)\right)^{\mathcal{P}_{\operatorname{fin}}(Q)}$ in $\operatorname{Hom}(C,A)$ (which consists of all sums of finite subfamilies of $\left(f_q\right)_{q\in Q}$). It is easy to see that the family $\left(f_q\right)_{q\in Q}$ is pointwise finitely supported if and only if the net $\left(\sum_{q\in K} f_q\right)_{K\in\mathcal{P}_{\operatorname{fin}}(Q)} \in \left(\operatorname{Hom}(C,A)\right)^{\mathcal{P}_{\operatorname{fin}}(Q)}$ converges in $\operatorname{Hom}(C,A)$. In this case, the limit of the net is precisely $\sum_{q\in Q} f_q$.

It is clear that this line of reasoning allows us to generalize the notion of "pointwise finitely supported families" to families in any topological $\mathbf{k}$-module, and to define the sum of any such family.]

13.40. **Solution to Exercise 1.7.13.**

Solution to Exercise 1.7.13. Before we start proving Proposition 1.7.11, let us prove some facts which will be useful on several occasions:

> *Fact B.1:* If $f \in \operatorname{Hom}(C,A)$ is a pointwise $\star$-nilpotent map, and if $\left(\lambda_n\right)_{n\in\mathbb{N}} \in \mathbf{k}^{\mathbb{N}}$ is any family of scalars, then the family $\left(\lambda_n f^{\star n}\right)_{n\in\mathbb{N}} \in \left(\operatorname{Hom}(C,A)\right)^{\mathbb{N}}$ is pointwise finitely supported.

[*Proof of Fact B.1:* Fact B.1 has been stated in Definition 1.7.10(b); it was already proven in a footnote.]

> *Fact B.2:* Let $f \in \mathfrak{n}(C,A)$. Let $\left(\lambda_n\right)_{n\in\mathbb{N}} \in \mathbf{k}^{\mathbb{N}}$ be any family of scalars. Then, the family $\left(\lambda_n f^{\star n}\right)_{n\in\mathbb{N}} \in \left(\operatorname{Hom}(C,A)\right)^{\mathbb{N}}$ is pointwise finitely supported, and its sum $\sum_{n\geq 0} \lambda_n f^{\star n}$ belongs to $\operatorname{Hom}(C,A)$.

[*Proof of Fact B.2:* We have $f \in \mathfrak{n}(C,A)$. In other words, f is a pointwise $\star$-nilpotent map in $\operatorname{Hom}(C,A)$ (since $\mathfrak{n}(C,A)$ is the set of all pointwise $\star$-nilpotent maps in $\operatorname{Hom}(C,A)$). Thus, Fact B.1 shows that the family $\left(\lambda_n f^{\star n}\right)_{n\in\mathbb{N}} \in \left(\operatorname{Hom}(C,A)\right)^{\mathbb{N}}$ is pointwise finitely supported. Hence, the sum $\sum_{n\in\mathbb{N}} \lambda_n f^{\star n}$ is well-defined. In other words, the sum $\sum_{n\geq 0} \lambda_n f^{\star n}$ is well-defined (since $\sum_{n\geq 0} = \sum_{n\in\mathbb{N}}$).

Now, Proposition 1.7.4 (applied to $\mathbb{N}$ and $\left(\lambda_n f^{\star n}\right)_{n\in\mathbb{N}}$ instead of Q and $\left(f_q\right)_{q\in Q}$) shows that the map $\sum_{n\in\mathbb{N}} \lambda_n f^{\star n}$ belongs to $\operatorname{Hom}(C,A)$. In other words, the map $\sum_{n\geq 0} \lambda_n f^{\star n}$ belongs to $\operatorname{Hom}(C,A)$ (since $\sum_{n\geq 0} = \sum_{n\in\mathbb{N}}$). Thus, Fact B.2 is proven.]

> *Fact B.3:* Let $\left(f_q\right)_{q\in Q}$ be a pointwise finitely supported family in $\left(\operatorname{Hom}(C,A)\right)^Q$. Let $\lambda \in \mathbf{k}$. Then, the family $\left(\lambda f_q\right)_{q\in Q} \in \left(\operatorname{Hom}(C,A)\right)^Q$ is also pointwise finitely supported, and satisfies
> $$\lambda \sum_{q\in Q} f_q = \sum_{q\in Q} \lambda f_q.$$

[*Proof of Fact B.3:* Fact B.3 is similar to Proposition 1.7.5, and its proof is analogous to the proof of the latter (but simpler).]

Before we state the next fact, let us recall how the convergence of a (possibly infinite) sum of power series in $\mathbf{k}[[T]]$ is defined:

> *Definition:* Let us introduce a notation: If $u \in \mathbf{k}[[T]]$ is any power series, and if $n \in \mathbb{N}$, then $\left[T^n\right] u$ will mean the coefficient of T^n in u. Thus, $u = \sum_{n\geq 0} \left(\left[T^n\right] u\right) \cdot T^n$ for each $u \in \mathbf{k}[[T]]$.
>
> Let $\left(r_q\right)_{q\in Q} \in \left(\mathbf{k}[[T]]\right)^Q$ be a family of power series in $\mathbf{k}[[T]]$. We say that the sum $\sum_{q\in Q} r_q$ *converges* in $\mathbf{k}[[T]]$ if and only if for each $n \in \mathbb{N}$,
> $$\text{all but finitely many } q \in Q \text{ satisfy } \left[T^n\right] r_q = 0.$$

[492] In this case, the sum $\sum_{q\in Q} r_q$ is defined to be the power series in $\mathbf{k}[[T]]$ whose coefficients are given by the rule

$$\left([T^n] \left(\sum_{q\in Q} r_q \right) = \sum_{q\in Q} [T^n] r_q \qquad \text{for all } n \in \mathbb{N} \right).$$

Now, we can state a useful fact that relates this notion of convergence to manipulations of maps in $\operatorname{Hom}(C, A)$:

Fact B.7: Let $(r_q)_{q\in Q} \in (\mathbf{k}[[T]])^Q$ be a family of power series such that the (possibly infinite) sum $\sum_{q\in Q} r_q$ converges in $\mathbf{k}[[T]]$. Let $f \in \mathfrak{n}(C, A)$. Then, the family $\left((r_q)^\star (f)\right)_{q\in Q} \in (\operatorname{Hom}(C, A))^Q$ is pointwise finitely supported and satisfies

$$\left(\sum_{q\in Q} r_q \right)^\star (f) = \sum_{q\in Q} (r_q)^\star (f).$$

[*Proof of Fact B.7:* Define a power series $s \in \mathbf{k}[[T]]$ by $s = \sum_{q\in Q} r_q$. (This is well-defined, since the sum $\sum_{q\in Q} r_q$ converges.)

Let us introduce a notation: If $u \in \mathbf{k}[[T]]$ is any power series, and if $n \in \mathbb{N}$, then $[T^n] u$ will mean the coefficient of T^n in u. Thus, $u = \sum_{n\geq 0} ([T^n] u) \cdot T^n$ for each $u \in \mathbf{k}[[T]]$. Applying this to $u = s$, we find $s = \sum_{n\geq 0} ([T^n] s) \cdot T^n$.

We know that the sum $\sum_{q\in Q} r_q$ converges in $\mathbf{k}[[T]]$. In other words, for each $n \in \mathbb{N}$,

$$\text{all but finitely many } q \in Q \text{ satisfy } [T^n] r_q = 0 \tag{13.40.1}$$

(by the definition of convergence for an infinite sum in $\mathbf{k}[[T]]$). (Of course, what precisely "all but finitely many $q \in Q$" means here – i.e., which q are excluded – depends on n.)

For every $q \in Q$, we have $r_q = \sum_{n\geq 0} ([T^n] r_q) \cdot T^n$ (since $u = \sum_{n\geq 0} ([T^n] u) \cdot T^n$ for each $u \in \mathbf{k}[[T]]$) and therefore

$$(r_q)^\star (f) = \sum_{n\geq 0} ([T^n] r_q) f^{\star n} \tag{13.40.2}$$

(by the definition of $(r_q)^\star (f)$).

We have $f \in \mathfrak{n}(C, A)$. In other words, f is a pointwise $\star$-nilpotent map in $\operatorname{Hom}(C, A)$ (since $\mathfrak{n}(C, A)$ is the set of all pointwise $\star$-nilpotent maps in $\operatorname{Hom}(C, A)$). Thus, the family $(f^{\star n})_{n\in\mathbb{N}} \in (\operatorname{Hom}(C, A))^{\mathbb{N}}$ is pointwise finitely supported (since f is pointwise $\star$-nilpotent). In other words, for each $x \in C$,

$$\text{the family } (f^{\star n}(x))_{n\in\mathbb{N}} \in A^{\mathbb{N}} \text{ is finitely supported.} \tag{13.40.3}$$

Now, let $x \in C$. Then, there exists an $N \in \mathbb{N}$ such that

$$\text{every } n \geq N \text{ satisfies } f^{\star n}(x) = 0 \tag{13.40.4}$$

[493]. Consider this N.

[492] Of course, what precisely "all but finitely many $q \in Q$" means here – i.e., which q are excluded – depends on n.

[493] *Proof.* The family $(f^{\star n}(x))_{n\in\mathbb{N}} \in A^{\mathbb{N}}$ is finitely supported (by (13.40.3)). In other words, all but finitely many $n \in \mathbb{N}$ satisfy $f^{\star n}(x) = 0$. In other words, there exists a finite subset Z of $\mathbb{N}$ such that

$$\text{each } n \in \mathbb{N} \setminus Z \text{ satisfies } f^{\star n}(x) = 0. \tag{13.40.5}$$

Consider this Z.

The set Z is a finite subset of $\mathbb{N}$, and thus has an upper bound (since any finite subset of $\mathbb{N}$ has an upper bound). In other words, there exists some $w \in \mathbb{N}$ such that each $z \in Z$ satisfies $z \leq w$. Consider this w.

Now, let $n \in \mathbb{N}$ be such that $n \geq w + 1$. Assume (for the sake of contradiction) that $n \in Z$. Recall that each $z \in Z$ satisfies $z \leq w$. Applying this to $z = n$, we obtain $n \leq w$ (since $n \in Z$), so that $n \leq w < w + 1$. This contradicts $n \geq w + 1$. This contradiction shows that our assumption (that $n \in Z$) was wrong. Hence, we have $n \notin Z$. Combining $n \in \mathbb{N}$ with $n \notin Z$, we obtain $n \in \mathbb{N} \setminus Z$. Hence, (13.40.5) yields $f^{\star n}(x) = 0$.

Now, forget that we fixed n. We thus have shown that every $n \geq w + 1$ satisfies $f^{\star n}(x) = 0$. Hence, there exists an $N \in \mathbb{N}$ such that every $n \geq N$ satisfies $f^{\star n}(x) = 0$ (namely, $N = w + 1$).

Every $q \in Q$ satisfies

$$\begin{aligned}
\underbrace{\left(\left(r_q\right)^{\star}(f)\right)}_{\substack{=\sum_{n\geq 0}([T^n]r_q)f^{\star n} \\ \text{(by (13.40.2))}}}(x) &= \left(\sum_{n\geq 0}\left(\left[T^n\right]r_q\right)f^{\star n}\right)(x) = \sum_{n\geq 0}\left(\left[T^n\right]r_q\right)f^{\star n}(x) \\
&= \underbrace{\sum_{\substack{n\geq 0;\\ n<N}}}_{=\sum_{n=0}^{N-1}}\left(\left[T^n\right]r_q\right)f^{\star n}(x) + \sum_{\substack{n\geq 0;\\ n\geq N}}\left(\left[T^n\right]r_q\right)\underbrace{f^{\star n}(x)}_{\substack{=0\\ \text{(by (13.40.4))}}} \\
&= \sum_{n=0}^{N-1}\left(\left[T^n\right]r_q\right)f^{\star n}(x) + \underbrace{\sum_{\substack{n\geq 0;\\ n\geq N}}\left(\left[T^n\right]r_q\right)0}_{=0} \\
&= \sum_{n=0}^{N-1}\left(\left[T^n\right]r_q\right)f^{\star n}(x). \qquad (13.40.6)
\end{aligned}$$

For each $n \in \mathbb{N}$, there exists a finite subset Q_n of Q such that

$$\text{all } q \in Q \setminus Q_n \text{ satisfy } \left[T^n\right]r_q = 0 \qquad (13.40.7)$$

(by (13.40.1)). Consider this Q_n.

Let Q' be the subset $Q_0 \cup Q_1 \cup \cdots \cup Q_{N-1}$ of Q. Then, Q' is the union of the N finite sets $Q_0, Q_1, \ldots, Q_{N-1}$. Hence, Q' itself is a finite set. Thus, all but finitely many $q \in Q$ satisfy $q \in Q \setminus Q'$. Notice that the set Q is the union of its two disjoint subsets Q' and $Q \setminus Q'$ (since Q' is a subset of Q).

Moreover, if $n \in \{0, 1, \ldots, N-1\}$, then

$$\text{every } q \in Q \setminus Q' \text{ satisfies } \left[T^n\right]r_q = 0 \qquad (13.40.8)$$

[494]. Hence, every $n \in \{0, 1, \ldots, N-1\}$ satisfies

$$\begin{aligned}
\left[T^n\right]\underbrace{s}_{=\sum_{q\in Q} r_q} &= \left[T^n\right]\left(\sum_{q\in Q} r_q\right) = \sum_{q\in Q}\left[T^n\right]r_q = \sum_{q\in Q'}\left[T^n\right]r_q + \sum_{q\in Q\setminus Q'}\underbrace{\left[T^n\right]r_q}_{\substack{=0\\ \text{(by (13.40.8))}}} \\
&\qquad \left(\text{since the set } Q \text{ is the union of its two disjoint subsets } Q' \text{ and } Q \setminus Q'\right) \\
&= \sum_{q\in Q'}\left[T^n\right]r_q + \underbrace{\sum_{q\in Q\setminus Q'} 0}_{=0} = \sum_{q\in Q'}\left[T^n\right]r_q. \qquad (13.40.9)
\end{aligned}$$

Now, each $q \in Q \setminus Q'$ satisfies

$$\begin{aligned}
\left(\left(r_q\right)^{\star}(f)\right)(x) &= \sum_{n=0}^{N-1}\underbrace{\left(\left[T^n\right]r_q\right)}_{\substack{=0\\ \text{(by (13.40.8))}}}f^{\star n}(x) \qquad \text{(by (13.40.6))} \\
&= \sum_{n=0}^{N-1} 0 f^{\star n}(x) = 0. \qquad (13.40.10)
\end{aligned}$$

[494] *Proof of (13.40.8):* Let $n \in \{0, 1, \ldots, N-1\}$. Let $q \in Q \setminus Q'$.
From $n \in \{0, 1, \ldots, N-1\}$, we obtain $Q_n \subset Q_0 \cup Q_1 \cup \cdots \cup Q_{N-1} = Q'$ (since $Q' = Q_0 \cup Q_1 \cup \cdots \cup Q_{N-1}$). Hence, $Q \setminus \underbrace{Q_n}_{\subset Q'} \supset Q \setminus Q'$, so that $Q \setminus Q' \subset Q \setminus Q_n$. Hence, $q \in Q \setminus Q' \subset Q \setminus Q_n$. Therefore, (13.40.7) yields $\left[T^n\right]r_q = 0$. This proves (13.40.8).

Hence, all but finitely many $q \in Q$ satisfy $\left(\left(r_q\right)^{\star}(f)\right)(x) = 0$ (since all but finitely many $q \in Q$ satisfy $q \in Q \setminus Q'$). In other words,

$$\text{the family } \left(\left(\left(r_q\right)^{\star}(f)\right)(x)\right)_{q\in Q} \in A^Q \text{ is finitely supported.} \tag{13.40.11}$$

Recall again that the set Q is the union of its two disjoint subsets Q' and $Q \setminus Q'$. Hence,

$$\begin{aligned}
\sum_{q\in Q}\left(\left(r_q\right)^{\star}(f)\right)(x) &= \sum_{q\in Q'} \underbrace{\left(\left(r_q\right)^{\star}(f)\right)(x)}_{\substack{=\sum_{n=0}^{N-1}([T^n]r_q)f^{\star n}(x) \\ \text{(by (13.40.6))}}} + \sum_{q\in Q\setminus Q'} \underbrace{\left(\left(r_q\right)^{\star}(f)\right)(x)}_{\substack{=0 \\ \text{(by (13.40.10))}}} \\
&= \underbrace{\sum_{q\in Q'}\sum_{n=0}^{N-1}}_{\substack{=\sum_{n=0}^{N-1}\sum_{q\in Q'} \\ \text{(here, we are interchanging} \\ \text{two } \textbf{finite} \text{ sums)}}} \left(\left[T^n\right] r_q\right) f^{\star n}(x) + \underbrace{\sum_{q\in Q\setminus Q'} 0}_{=0} \\
&= \sum_{n=0}^{N-1}\sum_{q\in Q'} \left(\left[T^n\right] r_q\right) f^{\star n}(x) .
\end{aligned} \tag{13.40.12}$$

On the other hand, recall that $s = \sum_{n\geq 0}\left(\left[T^n\right] s\right) \cdot T^n$. Therefore,

$$s^{\star}(f) = \sum_{n\geq 0}\left(\left[T^n\right] s\right) f^{\star n} \qquad \left(\text{by the definition of } s^{\star}(f)\right).$$

Applying both sides of this equality to x, we obtain

$$\begin{aligned}
\left(s^{\star}(f)\right)(x) &= \left(\sum_{n\geq 0}\left(\left[T^n\right] s\right) f^{\star n}\right)(x) = \sum_{n\geq 0}\left(\left[T^n\right] s\right) f^{\star n}(x) \\
&= \underbrace{\sum_{\substack{n\geq 0; \\ n<N}}}_{=\sum_{n=0}^{N-1}} \left(\left[T^n\right] s\right) f^{\star n}(x) + \sum_{\substack{n\geq 0; \\ n\geq N}} \left(\left[T^n\right] s\right) \underbrace{f^{\star n}(x)}_{\substack{=0 \\ \text{(by (13.40.4))}}} \\
&= \sum_{n=0}^{N-1} \underbrace{\left(\left[T^n\right] s\right)}_{\substack{=\sum_{q\in Q'}[T^n]r_q \\ \text{(by (13.40.9))}}} f^{\star n}(x) + \underbrace{\sum_{\substack{n\geq 0; \\ n\geq N}} \left(\left[T^n\right] s\right) 0}_{=0} \\
&= \sum_{n=0}^{N-1}\left(\sum_{q\in Q'}\left(\left[T^n\right] r_q\right)\right) f^{\star n}(x) = \sum_{n=0}^{N-1}\sum_{q\in Q'}\left(\left[T^n\right] r_q\right) f^{\star n}(x) \\
&= \sum_{q\in Q}\left(\left(r_q\right)^{\star}(f)\right)(x) \qquad \left(\text{by (13.40.12)}\right).
\end{aligned} \tag{13.40.13}$$

Now, forget that we fixed x. We thus have shown that each $x \in C$ satisfies (13.40.11) and (13.40.13).

In particular, for each $x \in C$, the family $\left(\left(\left(r_q\right)^{\star}(f)\right)(x)\right)_{q\in Q} \in A^Q$ is finitely supported (since each $x \in C$ satisfies (13.40.11)). In other words, the family $\left(\left(r_q\right)^{\star}(f)\right)_{q\in Q} \in \left(\operatorname{Hom}(C, A)\right)^Q$ is pointwise finitely supported. Hence, the sum $\sum_{q\in Q}\left(r_q\right)^{\star}(f)$ is well-defined.

Furthermore, (13.40.13) shows that each $x \in C$ satisfies

$$\left(s^{\star}(f)\right)(x) = \sum_{q\in Q}\left(\left(r_q\right)^{\star}(f)\right)(x) = \left(\sum_{q\in Q}\left(r_q\right)^{\star}(f)\right)(x).$$

In other words, we have $s^{\star}(f) = \sum_{q\in Q}\left(r_q\right)^{\star}(f)$. Since $s = \sum_{q\in Q} r_q$, this rewrites as $\left(\sum_{q\in Q} r_q\right)^{\star}(f) = \sum_{q\in Q}\left(r_q\right)^{\star}(f)$. This completes the proof of Fact B.7.]

Proof of Proposition 1.7.11. (a) Let $f \in \mathfrak{n}(C, A)$ and $k \in \mathbb{N}$. The power series $T^k \in \mathbf{k}[[T]]$ can be written in the form $T^k = \sum_{n\geq 0} \delta_{n,k} T^n$ (since all addends in the sum $\sum_{n\geq 0} \delta_{n,k} T^n$ are zero except for the addend for $n = k$). Hence, the definition of $\left(T^k\right)^\star(f)$ yields

$$\left(T^k\right)^\star(f) = \sum_{n\geq 0} \delta_{n,k} f^{\star n} = f^{\star k}$$

(since all addends in the sum $\sum_{n\geq 0} \delta_{n,k} f^{\star n}$ are zero except for the addend for $n = k$). This proves Proposition 1.7.11(a).

(b) We are going to prove the formulas (1.7.2), (1.7.4), (1.7.3), (1.7.5), and (1.7.6) in this order.

[*Proof of (1.7.2):* Let $f \in \mathfrak{n}(C, A)$ and $u, v \in \mathbf{k}[[T]]$. We must prove the equality (1.7.2).

Write the power series u in the form $u = \sum_{n\geq 0} u_n T^n$ with $(u_n)_{n\geq 0} \in \mathbf{k}^{\mathbb{N}}$. Thus, $u^\star(f) = \sum_{n\geq 0} u_n f^{\star n}$ (by the definition of $u^\star(f)$).

Write the power series v in the form $v = \sum_{n\geq 0} v_n T^n$ with $(v_n)_{n\geq 0} \in \mathbf{k}^{\mathbb{N}}$. Thus, $v^\star(f) = \sum_{n\geq 0} v_n f^{\star n}$ (by the definition of $v^\star(f)$).

Fact B.2 (applied to $(\lambda_n)_{n\in\mathbb{N}} = (u_n)_{n\in\mathbb{N}}$) shows that the family $(u_n f^{\star n})_{n\in\mathbb{N}} \in (\operatorname{Hom}(C, A))^{\mathbb{N}}$ is pointwise finitely supported, and that its sum $\sum_{n\geq 0} u_n f^{\star n}$ belongs to $\operatorname{Hom}(C, A)$.

Fact B.2 (applied to $(\lambda_n)_{n\in\mathbb{N}} = (v_n)_{n\in\mathbb{N}}$) shows that the family $(v_n f^{\star n})_{n\in\mathbb{N}} \in (\operatorname{Hom}(C, A))^{\mathbb{N}}$ is pointwise finitely supported, and that its sum $\sum_{n\geq 0} v_n f^{\star n}$ belongs to $\operatorname{Hom}(C, A)$.

Proposition 1.7.5 (applied to $Q = \mathbb{N}$, $(f_q)_{q\in Q} = (u_n f^{\star n})_{n\in\mathbb{N}}$ and $(g_q)_{q\in Q} = (v_n f^{\star n})_{n\in\mathbb{N}}$) now shows that the family $(u_n f^{\star n} + v_n f^{\star n})_{n\in\mathbb{N}} \in (\operatorname{Hom}(C, A))^{\mathbb{N}}$ is also pointwise finitely supported, and satisfies

$$\sum_{n\in\mathbb{N}} u_n f^{\star n} + \sum_{n\in\mathbb{N}} v_n f^{\star n} = \sum_{n\in\mathbb{N}} (u_n f^{\star n} + v_n f^{\star n}).$$

Since $\sum_{n\in\mathbb{N}} = \sum_{n\geq 0}$, this rewrites as

$$\sum_{n\geq 0} u_n f^{\star n} + \sum_{n\geq 0} v_n f^{\star n} = \sum_{n\geq 0} \underbrace{(u_n f^{\star n} + v_n f^{\star n})}_{=(u_n+v_n) f^{\star n}} = \sum_{n\geq 0} (u_n + v_n) f^{\star n}. \tag{13.40.14}$$

Adding the equalities $u = \sum_{n\geq 0} u_n T^n$ and $v = \sum_{n\geq 0} v_n T^n$, we obtain

$$u + v = \sum_{n\geq 0} u_n T^n + \sum_{n\geq 0} v_n T^n = \sum_{n\geq 0} (u_n + v_n) T^n.$$

Hence, the definition of $(u + v)^\star(f)$ yields

$$\begin{aligned}(u + v)^\star(f) &= \sum_{n\geq 0} (u_n + v_n) f^{\star n} = \underbrace{\sum_{n\geq 0} u_n f^{\star n}}_{=u^\star(f)} + \underbrace{\sum_{n\geq 0} v_n f^{\star n}}_{=v^\star(f)} \qquad \text{(by (13.40.14))} \\ &= u^\star(f) + v^\star(f).\end{aligned}$$

This proves (1.7.2).]

[*Proof of (1.7.4):* This is similar to the proof of (1.7.2), but this time we need to apply Fact B.3 (instead of applying Proposition 1.7.5). The straightforward details are left to the reader.]

[*Proof of (1.7.3):* Let $f \in \mathfrak{n}(C, A)$ and $u, v \in \mathbf{k}[[T]]$. We must prove the equality (1.7.3).

Write the power series u in the form $u = \sum_{n\geq 0} u_n T^n$ with $(u_n)_{n\geq 0} \in \mathbf{k}^{\mathbb{N}}$. Thus, $u^\star(f) = \sum_{n\geq 0} u_n f^{\star n}$ (by the definition of $u^\star(f)$).

Write the power series v in the form $v = \sum_{n\geq 0} v_n T^n$ with $(v_n)_{n\geq 0} \in \mathbf{k}^{\mathbb{N}}$. Thus, $v^\star(f) = \sum_{n\geq 0} v_n f^{\star n}$ (by the definition of $v^\star(f)$).

Fact B.2 (applied to $(\lambda_n)_{n\in\mathbb{N}} = (u_n)_{n\in\mathbb{N}}$) shows that the family $(u_n f^{\star n})_{n\in\mathbb{N}} \in (\operatorname{Hom}(C, A))^{\mathbb{N}}$ is pointwise finitely supported, and that its sum $\sum_{n\geq 0} u_n f^{\star n}$ belongs to $\operatorname{Hom}(C, A)$. Renaming the index n as q in this statement, we obtain the following: The family $(u_q f^{\star q})_{q\in\mathbb{N}} \in (\operatorname{Hom}(C, A))^{\mathbb{N}}$ is pointwise finitely supported, and its sum $\sum_{q\geq 0} u_q f^{\star q}$ belongs to $\operatorname{Hom}(C, A)$.

Similarly, the family $(v_r f^{\star r})_{r\in\mathbb{N}} \in (\operatorname{Hom}(C, A))^{\mathbb{N}}$ is pointwise finitely supported, and its sum $\sum_{r\geq 0} v_r f^{\star r}$ belongs to $\operatorname{Hom}(C, A)$.

Proposition 1.7.6 (applied to $Q=\mathbb{N}$, $R=\mathbb{N}$, $(f_q)_{q\in Q}=(u_qf^{\star q})_{q\in\mathbb{N}}$ and $(g_r)_{r\in R}=(v_rf^{\star r})_{r\in\mathbb{N}}$) thus shows that the family $((u_qf^{\star q})\star(v_rf^{\star r}))_{(q,r)\in\mathbb{N}\times\mathbb{N}}\in(\operatorname{Hom}(C,A))^{\mathbb{N}\times\mathbb{N}}$ is pointwise finitely supported, and satisfies

$$\sum_{(q,r)\in\mathbb{N}\times\mathbb{N}}((u_qf^{\star q})\star(v_rf^{\star r}))=\left(\sum_{q\in\mathbb{N}}u_qf^{\star q}\right)\star\left(\sum_{r\in\mathbb{N}}v_rf^{\star r}\right).$$

Hence,

$$\begin{aligned}\left(\sum_{q\in\mathbb{N}}u_qf^{\star q}\right)\star\left(\sum_{r\in\mathbb{N}}v_rf^{\star r}\right)&=\sum_{(q,r)\in\mathbb{N}\times\mathbb{N}}\underbrace{((u_qf^{\star q})\star(v_rf^{\star r}))}_{\substack{=u_qv_rf^{\star q}\star f^{\star r}\\=u_qv_rf^{\star(q+r)}}}\\&=\sum_{(q,r)\in\mathbb{N}\times\mathbb{N}}u_qv_rf^{\star(q+r)}.\end{aligned}\tag{13.40.15}$$

Multiplying the equalities

$$u=\sum_{n\geq0}u_nT^n=\sum_{n\in\mathbb{N}}u_nT^n=\sum_{q\in\mathbb{N}}u_qT^q\qquad\text{(here, we have renamed the summation index n as q)}$$

and

$$v=\sum_{n\geq0}v_nT^n=\sum_{n\in\mathbb{N}}v_nT^n=\sum_{r\in\mathbb{N}}v_rT^r\qquad\text{(here, we have renamed the summation index n as r)},$$

we obtain

$$\begin{aligned}uv=\left(\sum_{q\in\mathbb{N}}u_qT^q\right)\left(\sum_{r\in\mathbb{N}}v_rT^r\right)&=\underbrace{\sum_{q\in\mathbb{N}}\sum_{r\in\mathbb{N}}}_{=\sum_{(q,r)\in\mathbb{N}\times\mathbb{N}}}\underbrace{u_qT^qv_rT^r}_{=u_qv_rT^{q+r}}\\&=\sum_{(q,r)\in\mathbb{N}\times\mathbb{N}}u_qv_rT^{q+r}.\end{aligned}\tag{13.40.16}$$

Thus, in particular, the sum $\sum_{(q,r)\in\mathbb{N}\times\mathbb{N}}u_qv_rT^{q+r}$ converges in $\mathbf{k}[[T]]$. Hence, Fact B.7 (applied to $\mathbb{N}\times\mathbb{N}$ and $(u_qv_rT^{q+r})_{(q,r)\in\mathbb{N}\times\mathbb{N}}$ instead of Q and $(r_q)_{q\in Q}$) shows that the family $\left((u_qv_rT^{q+r})^\star(f)\right)_{(q,r)\in\mathbb{N}\times\mathbb{N}}\in(\operatorname{Hom}(C,A))^{\mathbb{N}\times\mathbb{N}}$ is pointwise finitely supported and satisfies

$$\left(\sum_{(q,r)\in\mathbb{N}\times\mathbb{N}}u_qv_rT^{q+r}\right)^\star(f)=\sum_{(q,r)\in\mathbb{N}\times\mathbb{N}}\left(u_qv_rT^{q+r}\right)^\star(f).\tag{13.40.17}$$

In light of (13.40.16), the equality (13.40.17) rewrites as

$$\begin{aligned}(uv)^\star(f)&=\sum_{(q,r)\in\mathbb{N}\times\mathbb{N}}\underbrace{\left(u_qv_rT^{q+r}\right)^\star(f)}_{\substack{=u_qv_r\left(T^{q+r}\right)^\star(f)\\\text{(by (1.7.4) (applied}\\\text{to u_qv_r and T^{q+r} instead of λ and u))}}}=\sum_{(q,r)\in\mathbb{N}\times\mathbb{N}}u_qv_r\underbrace{\left(T^{q+r}\right)^\star(f)}_{\substack{=f^{\star(q+r)}\\\text{(by (1.7.1) (applied to $k=q+r$))}}}\\&=\sum_{(q,r)\in\mathbb{N}\times\mathbb{N}}u_qv_rf^{\star(q+r)}=\left(\sum_{q\in\mathbb{N}}u_qf^{\star q}\right)\star\left(\sum_{r\in\mathbb{N}}v_rf^{\star r}\right)\qquad\text{(by (13.40.15))}.\end{aligned}$$

Comparing this with

$$\underbrace{u^{\star}\left(f\right)}_{=\sum_{n\geq 0} u_n f^{\star n}} \star \underbrace{v^{\star}\left(f\right)}_{=\sum_{n\geq 0} v_n f^{\star n}} = \left(\underbrace{\sum_{n\geq 0}}_{=\sum_{n\in\mathbb{N}}} u_n f^{\star n} \right) \star \left(\underbrace{\sum_{n\geq 0}}_{=\sum_{n\in\mathbb{N}}} v_n f^{\star n} \right) = \underbrace{\left(\sum_{n\in\mathbb{N}} u_n f^{\star n}\right)}_{=\sum_{q\in\mathbb{N}} u_q f^{\star q}} \star \underbrace{\left(\sum_{n\in\mathbb{N}} v_n f^{\star n}\right)}_{=\sum_{r\in\mathbb{N}} v_r f^{\star r}}$$
$$= \left(\sum_{q\in\mathbb{N}} u_q f^{\star q}\right) \star \left(\sum_{r\in\mathbb{N}} v_r f^{\star r}\right),$$

we obtain $\left(uv\right)^{\star}\left(f\right) = u^{\star}\left(f\right) \star v^{\star}\left(f\right)$. This proves (1.7.3).]

[*Proof of (1.7.5):* Let $f \in \mathfrak{n}\left(C, A\right)$. Applying (1.7.4) to $u = 0$ and $\lambda = 0$, we find $\left(0 \cdot 0\right)^{\star}\left(f\right) = 0 \cdot 0^{\star}\left(f\right) = 0$. In other words, $0^{\star}\left(f\right) = 0$. This proves (1.7.5).]

[*Proof of (1.7.6):* Let $f \in \mathfrak{n}\left(C, A\right)$. Applying (1.7.1) to $k = 0$, we find

$$\left(T^0\right)^{\star}\left(f\right) = f^{\star 0} = \left(\text{the unity of the } \mathbf{k}\text{-algebra } \left(\operatorname{Hom}\left(C, A\right), \star\right)\right) = u_A \epsilon_C$$

(since the unity of the $\mathbf{k}$-algebra $\left(\operatorname{Hom}\left(C, A\right), \star\right)$ is $u_A \epsilon_C$). In view of $T^0 = 1$, this rewrites as $1^{\star}\left(f\right) = u_A \epsilon_C$. This proves (1.7.6).]

We have now proven all the equalities (1.7.2), (1.7.3), (1.7.4), (1.7.5) and (1.7.6). Thus, Proposition 1.7.11(b) is proven.

(c) Let $f, g \in \mathfrak{n}\left(C, A\right)$ be such that $f \star g = g \star f$. We must prove that $f + g \in \mathfrak{n}\left(C, A\right)$.

From $f \star g = g \star f$, we conclude that the elements f and g of the $\mathbf{k}$-algebra $\left(\operatorname{Hom}\left(C, A\right), \star\right)$ commute.

Let $\mathfrak{G}$ be the $\mathbf{k}$-subalgebra of $\left(\operatorname{Hom}\left(C, A\right), \star\right)$ generated by the two elements f and g. Thus, the $\mathbf{k}$-algebra $\mathfrak{G}$ is generated by commuting elements (because the elements f and g of the $\mathbf{k}$-algebra $\left(\operatorname{Hom}\left(C, A\right), \star\right)$ commute), and therefore is commutative (since any $\mathbf{k}$-algebra generated by commuting elements must be commutative). Hence, the binomial formula holds in this $\mathbf{k}$-algebra $\mathfrak{G}$. Thus, we have

$$\left(f + g\right)^{\star n} = \sum_{i=0}^{n} \binom{n}{i} f^{\star i} \star g^{\star (n-i)} \qquad \text{for each } n \in \mathbb{N} \tag{13.40.18}$$

(since the multiplication in the $\mathbf{k}$-algebra $\mathfrak{G}$ is $\star$).

We have $f \in \mathfrak{n}\left(C, A\right)$. In other words, f is a pointwise $\star$-nilpotent map in $\operatorname{Hom}\left(C, A\right)$ (since $\mathfrak{n}\left(C, A\right)$ is the set of all pointwise $\star$-nilpotent maps in $\operatorname{Hom}\left(C, A\right)$). Thus, the family $\left(f^{\star n}\right)_{n\in\mathbb{N}}$ is pointwise finitely supported. Renaming the index n as q in this statement, we thus conclude that the family $\left(f^{\star q}\right)_{q\in\mathbb{N}}$ is pointwise finitely supported. Similarly, the family $\left(g^{\star r}\right)_{r\in\mathbb{N}}$ is pointwise finitely supported.

Thus, Proposition 1.7.6 (applied to $Q = \mathbb{N}$, $R = \mathbb{N}$, $\left(f_q\right)_{q\in Q} = \left(f^{\star q}\right)_{q\in\mathbb{N}}$ and $\left(g_r\right)_{r\in R} = \left(g^{\star r}\right)_{r\in\mathbb{N}}$) shows that the family $\left(f^{\star q} \star g^{\star r}\right)_{(q,r)\in\mathbb{N}\times\mathbb{N}} \in \left(\operatorname{Hom}\left(C, A\right)\right)^{\mathbb{N}\times\mathbb{N}}$ is pointwise finitely supported, and that it satisfies

$$\sum_{(q,r)\in\mathbb{N}\times\mathbb{N}} \left(f^{\star q} \star g^{\star r}\right) = \left(\sum_{q\in\mathbb{N}} f^{\star q}\right) \star \left(\sum_{r\in\mathbb{N}} g^{\star r}\right).$$

In particular, the family $\left(f^{\star q} \star g^{\star r}\right)_{(q,r)\in\mathbb{N}\times\mathbb{N}} \in \left(\operatorname{Hom}\left(C, A\right)\right)^{\mathbb{N}\times\mathbb{N}}$ is pointwise finitely supported. In other words, for each $x \in C$,

$$\text{the family } \left(\left(f^{\star q} \star g^{\star r}\right)\left(x\right)\right)_{(q,r)\in\mathbb{N}\times\mathbb{N}} \in A^{\mathbb{N}\times\mathbb{N}} \text{ is finitely supported.} \tag{13.40.19}$$

Let $x \in C$. Then, all but finitely many $\left(q, r\right) \in \mathbb{N} \times \mathbb{N}$ satisfy $\left(f^{\star q} \star g^{\star r}\right)\left(x\right) = 0$ (because of (13.40.19)). In other words, there exists a finite subset K of $\mathbb{N} \times \mathbb{N}$ such that

$$\text{each } \left(q, r\right) \in \left(\mathbb{N} \times \mathbb{N}\right) \setminus K \text{ satisfies } \left(f^{\star q} \star g^{\star r}\right)\left(x\right) = 0. \tag{13.40.20}$$

Consider this K.

Let $Q = \left\{u + v \mid \left(u, v\right) \in K\right\}$. Thus, Q is a finite set (since K is a finite set). Thus, all but finitely many $n \in \mathbb{N}$ satisfy $n \in \mathbb{N} \setminus Q$.

But every $n \in \mathbb{N} \setminus Q$ satisfies $(f+g)^{\star n}(x) = 0$ [495]. Hence, all but finitely many $n \in \mathbb{N}$ satisfy $(f+g)^{\star n}(x) = 0$ (since all but finitely many $n \in \mathbb{N}$ satisfy $n \in \mathbb{N} \setminus Q$). In other words, the family $\left((f+g)^{\star n}(x)\right)_{n\in\mathbb{N}} \in A^{\mathbb{N}}$ is finitely supported.

Now, forget that we fixed x. We thus have shown that for each $x \in C$, the family $\left((f+g)^{\star n}(x)\right)_{n\in\mathbb{N}} \in A^{\mathbb{N}}$ is finitely supported. In other words, the family $\left((f+g)^{\star n}\right)_{n\in\mathbb{N}} \in (\operatorname{Hom}(C,A))^{\mathbb{N}}$ is pointwise finitely supported. In other words, the map $f+g$ is pointwise $\star$-nilpotent (by the definition of "pointwise $\star$-nilpotent"). In other words, $f+g \in \mathfrak{n}(C,A)$ (since $\mathfrak{n}(C,A)$ is the set of all pointwise $\star$-nilpotent maps in $\operatorname{Hom}(C,A)$). This proves Proposition 1.7.11(c).

(d) Let $\lambda \in \mathbf{k}$ and $f \in \mathfrak{n}(C,A)$. We must prove that $\lambda f \in \mathfrak{n}(C,A)$.

We have $f \in \mathfrak{n}(C,A)$. In other words, f is a pointwise $\star$-nilpotent map in $\operatorname{Hom}(C,A)$. Thus, the family $(f^{\star n})_{n\in\mathbb{N}}$ is pointwise finitely supported. Hence, Proposition 1.7.7 (applied to $Q = \mathbb{N}$, $(f_q)_{q\in Q} = (f^{\star n})_{n\in\mathbb{N}}$ and $(\lambda_q)_{q\in Q} = (\lambda^n)_{n\in\mathbb{N}}$) shows that the family $(\lambda^n f^{\star n})_{n\in\mathbb{N}} \in (\operatorname{Hom}(C,A))^{\mathbb{N}}$ is pointwise finitely supported. In other words, the family $\left((\lambda f)^{\star n}\right)_{n\in\mathbb{N}} \in (\operatorname{Hom}(C,A))^{\mathbb{N}}$ is pointwise finitely supported (since $(\lambda f)^{\star n} = \lambda^n f^{\star n}$ for each $n \in \mathbb{N}$). In other words, the map λf is pointwise $\star$-nilpotent. In other words, $\lambda f \in \mathfrak{n}(C,A)$. This proves Proposition 1.7.11(d).

(e) Let $f \in \mathfrak{n}(C,A)$ and $g \in \operatorname{Hom}(C,A)$ be such that $f \star g = g \star f$. We must prove that $f \star g \in \mathfrak{n}(C,A)$.

From $f \star g = g \star f$, we conclude that the elements f and g of the $\mathbf{k}$-algebra $(\operatorname{Hom}(C,A), \star)$ commute.

Let $\mathfrak{G}$ be the $\mathbf{k}$-subalgebra of $(\operatorname{Hom}(C,A), \star)$ generated by the two elements f and g. Thus, the $\mathbf{k}$-algebra $\mathfrak{G}$ is generated by commuting elements (because the elements f and g of the $\mathbf{k}$-algebra $(\operatorname{Hom}(C,A), \star)$ commute), and therefore is commutative (since any $\mathbf{k}$-algebra generated by commuting elements must be commutative). Hence, the usual laws for exponentiation hold in this $\mathbf{k}$-algebra $\mathfrak{G}$. In particular, we have

$$(f \star g)^{\star n} = f^{\star n} \star g^{\star n} \qquad \text{for each } n \in \mathbb{N} \tag{13.40.22}$$

(since the multiplication in the $\mathbf{k}$-algebra $\mathfrak{G}$ is the convolution $\star$).

We have $f \in \mathfrak{n}(C,A)$. In other words, f is a pointwise $\star$-nilpotent map in $\operatorname{Hom}(C,A)$. Thus, the family $(f^{\star n})_{n\in\mathbb{N}}$ is pointwise finitely supported.

Thus, Proposition 1.7.8 (applied to $Q = \mathbb{N}$, $(f_q)_{q\in Q} = (f^{\star n})_{n\in\mathbb{N}}$ and $(g_q)_{q\in Q} = (g^{\star n})_{n\in\mathbb{N}}$) shows that the family $(f^{\star n} \star g^{\star n})_{n\in\mathbb{N}} \in (\operatorname{Hom}(C,A))^{\mathbb{N}}$ is pointwise finitely supported. In other words, the family $\left((f \star g)^{\star n}\right)_{n\in\mathbb{N}} \in (\operatorname{Hom}(C,A))^{\mathbb{N}}$ is pointwise finitely supported (because of (13.40.22)). In other words, the map $f \star g$ is pointwise $\star$-nilpotent. In other words, $f \star g \in \mathfrak{n}(C,A)$. This proves Proposition 1.7.11(e).

(f) Let $v \in \mathbf{k}[[T]]$ be a power series whose constant term is 0. Let $f \in \mathfrak{n}(C,A)$. We must show that $v^{\star}(f) \in \mathfrak{n}(C,A)$.

We know that the constant term of v is 0. Thus, the power series v is divisible by T in the ring $\mathbf{k}[[T]]$. In other words, there exists a power series $u \in \mathbf{k}[[T]]$ such that $v = Tu$. Consider this u. Define $g \in \operatorname{Hom}(C,A)$ by $g = u^{\star}(f)$.

[495] *Proof.* Let $n \in \mathbb{N} \setminus Q$. We must show that $(f+g)^{\star n}(x) = 0$.

We have $n \in \mathbb{N} \setminus Q$. In other words, $n \in \mathbb{N}$ and $n \notin Q$.

Let $i \in \{0,1,\ldots,n\}$ be arbitrary. We shall show that $\left(f^{\star i} \star g^{\star(n-i)}\right)(x) = 0$ first.

From $i \in \{0,1,\ldots,n\}$, we obtain $i \in \mathbb{N}$ and $n-i \in \mathbb{N}$. Thus, $(i, n-i) \in \mathbb{N} \times \mathbb{N}$.

If we had $(i, n-i) \in K$, then we would have

$$i + (n-i) \in \{u+v \mid (u,v) \in K\} = Q \qquad (\text{since } Q = \{u+v \mid (u,v) \in K\}),$$

which would contradict $i + (n-i) = n \notin Q$. Thus, we cannot have $(i, n-i) \in K$. In other words, we have $(i, n-i) \notin K$.

Combining $(i, n-i) \in \mathbb{N} \times \mathbb{N}$ with $(i, n-i) \notin K$, we obtain $(i, n-i) \in (\mathbb{N} \times \mathbb{N}) \setminus K$. Hence, (13.40.20) (applied to $(q,r) = (i, n-i)$) yields $\left(f^{\star i} \star g^{\star(n-i)}\right)(x) = 0$.

Now, forget that we fixed i. We thus have shown that

$$\left(f^{\star i} \star g^{\star(n-i)}\right)(x) = 0 \qquad \text{for each } i \in \{0,1,\ldots,n\}. \tag{13.40.21}$$

Now, applying both sides of the equality (13.40.18) to x, we obtain

$$(f+g)^{\star n}(x) = \left(\sum_{i=0}^{n} \binom{n}{i} f^{\star i} \star g^{\star(n-i)}\right)(x) = \sum_{i=0}^{n} \binom{n}{i} \underbrace{\left(f^{\star i} \star g^{\star(n-i)}\right)(x)}_{\substack{=0 \\ (\text{by } (13.40.21))}} = \sum_{i=0}^{n} \binom{n}{i} 0 = 0.$$

Qed.

Applying (1.7.1) to $k = 1$, we find $\left(T^1\right)^\star(f) = f^{\star 1} = f$. Since $T^1 = T$, this rewrites as $T^\star(f) = f$. From $v = Tu$, we obtain

$$\begin{aligned} v^\star(f) &= (Tu)^\star(f) = \underbrace{T^\star(f)}_{=f} \star \underbrace{u^\star(f)}_{=g} \qquad \text{(by (1.7.3) (applied to T and u instead of u and v))} \\ &= f \star g. \end{aligned} \tag{13.40.23}$$

On the other hand, from $v = Tu = u \cdot T$, we obtain

$$\begin{aligned} v^\star(f) &= (u \cdot T)^\star(f) = \underbrace{u^\star(f)}_{=g} \star \underbrace{T^\star(f)}_{=f} \qquad \text{(by (1.7.3) (applied to T instead of v))} \\ &= g \star f. \end{aligned}$$

Comparing this with (13.40.23), we obtain $f \star g = g \star f$. Hence, Proposition 1.7.11(e) yields $f \star g \in \mathfrak{n}(C, A)$. In light of (13.40.23), this rewrites as $v^\star(f) \in \mathfrak{n}(C, A)$. This proves Proposition 1.7.11(f).

(g) Let us first prove the following fact:

Fact G.1: Let $v \in \mathbf{k}[[T]]$ be any power series. Let $f \in \mathfrak{n}(C, A)$. Then,

$$(v^n)^\star(f) = (v^\star(f))^{\star n} \qquad \text{for each } n \in \mathbb{N}. \tag{13.40.24}$$

[*Proof of Fact G.1:* We shall prove (13.40.24) by induction over n:

Induction base: We have $\left(\underbrace{v^0}_{=1}\right)^\star(f) = 1^\star(f) = u_A \epsilon_C$ (by (1.7.6)). Comparing this with

$$(v^\star(f))^{\star 0} = (\text{the unity of the } \mathbf{k}\text{-algebra } (\operatorname{Hom}(C, A), \star)) = u_A \epsilon_C,$$

we obtain $\left(v^0\right)^\star(f) = (v^\star(f))^{\star 0}$. In other words, (13.40.24) holds for $n = 0$. This completes the induction base.

Induction step: Let $N \in \mathbb{N}$. Assume that (13.40.24) holds for $n = N$. We must prove that (13.40.24) holds for $n = N + 1$.

We have assumed that (13.40.24) holds for $n = N$. In other words, we have $\left(v^N\right)^\star(f) = (v^\star(f))^{\star N}$. Now,

$$\begin{aligned} \left(\underbrace{v^{N+1}}_{=v^N v}\right)^\star(f) &= \left(v^N v\right)^\star(f) = \underbrace{\left(v^N\right)^\star(f)}_{=(v^\star(f))^{\star N}} \star v^\star(f) \qquad \left(\text{by (1.7.3) (applied to } u = v^N\text{)}\right) \\ &= (v^\star(f))^{\star N} \star v^\star(f) = (v^\star(f))^{\star(N+1)}. \end{aligned}$$

In other words, (13.40.24) holds for $n = N + 1$. This completes the induction step. Thus, the induction proof of (13.40.24) is complete. In other words, Fact G.1 is proven.]

Now, let $u, v \in \mathbf{k}[[T]]$ be two power series such that the constant term of v is 0. Let $f \in \mathfrak{n}(C, A)$ be arbitrary. We must prove (1.7.7).

Write the power series u in the form $u = \sum_{n \geq 0} u_n T^n$ with $(u_n)_{n \geq 0} \in \mathbf{k}^{\mathbb{N}}$. Thus,

$$u^\star(v^\star(f)) = \sum_{n \geq 0} u_n (v^\star(f))^{\star n} \tag{13.40.25}$$

(by the definition of $u^\star(v^\star(f))$).

But the definition of the composition $u[v]$ yields $u[v] = \sum_{n \in \mathbb{N}} u_n v^n$ (since $u = \sum_{n \geq 0} u_n T^n = \sum_{n \in \mathbb{N}} u_n T^n$). In particular, the sum $\sum_{n \in \mathbb{N}} u_n v^n$ converges in $\mathbf{k}[[T]]$. Hence, Fact B.7 (applied to $Q = \mathbb{N}$ and $(r_q)_{q \in Q} = (u_n v^n)_{n \in \mathbb{N}}$) shows that the family $\left((u_n v^n)^\star(f)\right)_{n \in \mathbb{N}} \in (\operatorname{Hom}(C, A))^{\mathbb{N}}$ is pointwise finitely supported and satisfies

$$\left(\sum_{n \in \mathbb{N}} u_n v^n\right)^\star(f) = \sum_{n \in \mathbb{N}} (u_n v^n)^\star(f). \tag{13.40.26}$$

Now, recall that $u[v] = \sum_{n\in\mathbb{N}} u_n v^n$. Hence,

$$(u[v])^\star(f) = \left(\sum_{n\in\mathbb{N}} u_n v^n\right)^\star (f) = \sum_{n\in\mathbb{N}} \underbrace{(u_n v^n)^\star (f)}_{\substack{=u_n (v^n)^\star (f) \\ \text{(by (1.7.4) (applied} \\ \text{to } u_n \text{ and } v^n \text{ instead of } \lambda \text{ and } u))}} \qquad \text{(by (13.40.26))}$$

$$= \sum_{n\in\mathbb{N}} u_n \underbrace{(v^n)^\star (f)}_{\substack{=(v^\star(f))^{\star n} \\ \text{(by Fact G.1)}}} = \underbrace{\sum_{n\in\mathbb{N}}}_{=\sum_{n\geq 0}} u_n (v^\star (f))^{\star n} = \sum_{n\geq 0} u_n (v^\star (f))^{\star n} = u^\star (v^\star (f))$$

(by (13.40.25)). Thus, (1.7.7) is proven. This proves Proposition 1.7.11(g).

(h) Let us first prove a simple fact:

Fact H.1: Let C be a graded $\mathbf{k}$-coalgebra. Let $f \in \operatorname{Hom}(C, A)$ be such that $f(C_0) = 0$. Then, for each $i \in \mathbb{N}$, we have

$$f^{\star i}(C_n) = 0 \qquad \text{for every } n \in \mathbb{N} \text{ satisfying } i > n. \tag{13.40.27}$$

[*Proof of Fact H.1:* We shall prove (13.40.27) by induction over i:

Induction base: There exists no $n \in \mathbb{N}$ satisfying $0 > n$ (since each $n \in \mathbb{N}$ satisfies $n \geq 0$). Hence, (13.40.27) is vacuously true for $i = 0$. This completes the induction base.

Induction step: Let $p \in \mathbb{N}$. Assume that (13.40.27) holds for $i = p$. We must prove that (13.40.27) holds for $i = p + 1$.

We have assumed that (13.40.27) holds for $i = p$. In other words, we have

$$f^{\star p}(C_n) = 0 \qquad \text{for every } n \in \mathbb{N} \text{ satisfying } p > n. \tag{13.40.28}$$

Now, let $n \in \mathbb{N}$ be such that $p + 1 > n$.

It is easy to see that

$$(f \otimes f^{\star p})(C_i \otimes C_j) = 0 \tag{13.40.29}$$

for every $(i, j) \in \mathbb{N}^2$ satisfying $i + j = n$ [496].

Recall that the $\mathbf{k}$-coalgebra C is graded. Thus, its comultiplication Δ is graded. In other words, $\Delta(C_k) \subset (C \otimes C)_k$ for each $k \in \mathbb{N}$. Applying this to $k = n$, we obtain

$$\Delta(C_n) \subset (C \otimes C)_n = \bigoplus_{\substack{(i,j)\in\mathbb{N}^2; \\ i+j=n}} C_i \otimes C_j \qquad \text{(by the definition of the grading on } C \otimes C)$$

$$= \sum_{\substack{(i,j)\in\mathbb{N}^2; \\ i+j=n}} C_i \otimes C_j \qquad \text{(since direct sums are sums)}.$$

[496] *Proof of (13.40.29):* Let $(i, j) \in \mathbb{N}^2$ be such that $i + j = n$. We must prove (13.40.29).

From $(i, j) \in \mathbb{N}^2$, we obtain $i \in \mathbb{N}$ and $j \in \mathbb{N}$. If $i = 0$, then

$$(f \otimes f^{\star p})(C_i \otimes C_j) = f\left(\underbrace{C_i}_{\substack{=C_0 \\ (\text{since } i=0)}}\right) \otimes f^{\star p}(C_j) = \underbrace{f(C_0)}_{=0} \otimes f^{\star p}(C_j) = 0 \otimes f^{\star p}(C_j) = 0.$$

Hence, if $i = 0$, then (13.40.29) is proven. Thus, for the rest of the proof of (13.40.29), we can WLOG assume that we don't have $i = 0$. Assume this.

We have $i \neq 0$ (since we don't have $i = 0$), and thus $i \geq 1$ (since $i \in \mathbb{N}$). But $i + j = n$, so that $j = n - \underbrace{i}_{\geq 1} \leq n - 1$. But $p + 1 > n$; thus, $p > n - 1 \geq j$ (since $j \leq n - 1$). Hence, (13.40.28) (applied to j instead of n) shows that $f^{\star p}(C_j) = 0$. Now,

$$(f \otimes f^{\star p})(C_i \otimes C_j) = f(C_i) \otimes \underbrace{f^{\star p}(C_j)}_{=0} = f(C_i) \otimes 0 = 0.$$

This proves (13.40.29).

Applying the map $f \otimes f^{\star p}$ to both sides of this relation, we obtain

$$\begin{aligned}
\left(f \otimes f^{\star p}\right)\left(\Delta\left(C_n\right)\right) \subset \left(f \otimes f^{\star p}\right)\left(\sum_{\substack{(i,j)\in\mathbb{N}^2;\\ i+j=n}} C_i \otimes C_j\right) &= \sum_{\substack{(i,j)\in\mathbb{N}^2;\\ i+j=n}} \underbrace{\left(f \otimes f^{\star p}\right)\left(C_i \otimes C_j\right)}_{\substack{=0\\ \text{(by (13.40.29))}}} \\
&= \sum_{\substack{(i,j)\in\mathbb{N}^2;\\ i+j=n}} 0 = 0. \qquad (13.40.30)
\end{aligned}$$

But the definition of convolution yields $f \star f^{\star p} = m \circ \left(f \otimes f^{\star p}\right) \circ \Delta$ (where m denotes the multiplication map $A \otimes A \to A$). We have

$$f^{\star(p+1)} = f \star f^{\star p} = m \circ \left(f \otimes f^{\star p}\right) \circ \Delta$$

and thus

$$\begin{aligned}
\underbrace{f^{\star(p+1)}}_{=m\circ(f\otimes f^{\star p})\circ\Delta}\left(C_n\right) = \left(m \circ \left(f \otimes f^{\star p}\right) \circ \Delta\right)\left(C_n\right) = m\left(\underbrace{\left(f \otimes f^{\star p}\right)\left(\Delta\left(C_n\right)\right)}_{\substack{\subset 0\\ \text{(by (13.40.30))}}}\right) \\
\subset m\left(0\right) = 0.
\end{aligned}$$

In other words, $f^{\star(p+1)}\left(C_n\right) = 0$.

Now, forget that we fixed n. We thus have proven that

$$f^{\star(p+1)}\left(C_n\right) = 0 \qquad \text{for every } n \in \mathbb{N} \text{ satisfying } p+1 > n.$$

In other words, (13.40.27) holds for $i = p+1$. This completes the induction step. Thus, (13.40.27) is proven by induction. Hence, Fact H.1 is proven.]

Now, let us actually prove Proposition 1.7.11(h).

Assume that C is a graded $\mathbf{k}$-coalgebra. Assume that $f \in \operatorname{Hom}(C, A)$ satisfies $f\left(C_0\right) = 0$. We must show that $f \in \mathfrak{n}(C, A)$.

Let $x \in C$. Thus, x is a sum of finitely many homogeneous elements of C (since C is graded). In other words, there exist some $k \in \mathbb{N}$ and some homogeneous elements $x_1, x_2, \ldots, x_k \in C$ satisfying $x = \sum_{g=1}^{k} x_g$. Consider this k and these $x_1, x_2, \ldots, x_k$.

For each $g \in \{1, 2, \ldots, k\}$, there exists some $n_g \in \mathbb{N}$ satisfying $x_g \in C_{n_g}$ (since x_g is a homogeneous element of C). Consider this n_g. The set $\{n_1, n_2, \ldots, n_k\}$ is a finite subset of $\mathbb{N}$, and thus has an upper bound (since any finite subset of $\mathbb{N}$ has an upper bound). In other words, there exists some $N \in \mathbb{N}$ such that

$$\text{each } n \in \{n_1, n_2, \ldots, n_k\} \text{ satisfies } n \leq N. \qquad (13.40.31)$$

Consider this N.

Let $Q = \{0, 1, \ldots, N\}$. Then, Q is a finite subset of $\mathbb{N}$. Thus, all but finitely many $n \in \mathbb{N}$ satisfy $n \in \mathbb{N} \setminus Q$.

Now, let $n \in \mathbb{N} \setminus Q$ be arbitrary. Thus, $n \in \mathbb{N} \setminus Q = \{N+1, N+2, N+3, \ldots\}$ (since $Q = \{0, 1, \ldots, N\}$). Hence, $n > N$.

Thus, each $g \in \{1, 2, \ldots, k\}$ satisfies

$$f^{\star n}\left(x_g\right) = 0 \qquad (13.40.32)$$

[497]. Applying the map $f^{\star n}$ to the equality $x = \sum_{g=1}^{k} x_g$, we obtain

$$f^{\star n}(x) = f^{\star n}\left(\sum_{g=1}^{k} x_g\right) = \sum_{g=1}^{k} \underbrace{f^{\star n}(x_g)}_{\substack{=0 \\ \text{(by (13.40.32))}}} = \sum_{g=1}^{k} 0 = 0.$$

Now, forget that we fixed n. We thus have shown that each $n \in \mathbb{N} \setminus Q$ satisfies $f^{\star n}(x) = 0$. Hence, all but finitely many $n \in \mathbb{N}$ satisfy $f^{\star n}(x) = 0$ (since all but finitely many $n \in \mathbb{N}$ satisfy $n \in \mathbb{N} \setminus Q$). In other words, the family $(f^{\star n}(x))_{n\in\mathbb{N}} \in A^{\mathbb{N}}$ is finitely supported.

Now, forget that we fixed x. We thus have shown that for each $x \in C$, the family $(f^{\star n}(x))_{n\in\mathbb{N}} \in A^{\mathbb{N}}$ is finitely supported. In other words, the family $(f^{\star n})_{n\in\mathbb{N}} \in (\operatorname{Hom}(C, A))^{\mathbb{N}}$ is pointwise finitely supported. In other words, the map f is pointwise $\star$-nilpotent. In other words, $f \in \mathfrak{n}(C, A)$. This proves Proposition 1.7.11(h).

(i) Before we prove this, let us state a general fact:

> *Fact H.7:* Let V and W be two $\mathbf{k}$-modules. Let $\varphi : V \to W$ be a $\mathbf{k}$-linear map. Let $(v_q)_{q\in Q} \in V^Q$ be a finitely supported family of elements of V. Then, the family $(\varphi(v_q))_{q\in Q} \in W^Q$ is also finitely supported, and satisfies $\sum_{q\in Q} \varphi(v_q) = \varphi\left(\sum_{q\in Q} v_q\right)$.

[*Proof of Fact H.7:* Fact H.7 is a basic property of linear maps (essentially, it says that any $\mathbf{k}$-linear map preserves sums in which all but finitely many addends are zero), and we omit its simple proof.]

Now, let us start proving Proposition 1.7.11(i).

Let B be any $\mathbf{k}$-algebra. Let $s : A \to B$ be any $\mathbf{k}$-algebra homomorphism.

Proposition 1.4.3 (applied to $A' = B$, $C' = C$, $\alpha = s$ and $\gamma = \operatorname{id}_C$) shows that the map

$$\operatorname{Hom}(C, A) \to \operatorname{Hom}(C, B), \qquad f \mapsto s \circ f \circ \operatorname{id}_C$$

is a $\mathbf{k}$-algebra homomorphism from the $\mathbf{k}$-algebra $(\operatorname{Hom}(C, A), \star)$ to the $\mathbf{k}$-algebra $(\operatorname{Hom}(C, B), \star)$. Since each $f \in \operatorname{Hom}(C, A)$ satisfies $s \circ f \circ \operatorname{id}_C = s \circ f$, this rewrites as follows: The map

$$\operatorname{Hom}(C, A) \to \operatorname{Hom}(C, B), \qquad f \mapsto s \circ f$$

is a $\mathbf{k}$-algebra homomorphism from the $\mathbf{k}$-algebra $(\operatorname{Hom}(C, A), \star)$ to the $\mathbf{k}$-algebra $(\operatorname{Hom}(C, B), \star)$. Let us denote this $\mathbf{k}$-algebra homomorphism by Φ.

Now, let $u \in \mathbf{k}[[T]]$ and $f \in \mathfrak{n}(C, A)$. We must show that

$$s \circ f \in \mathfrak{n}(C, B) \qquad \text{and} \qquad u^{\star}(s \circ f) = s \circ (u^{\star}(f)). \tag{13.40.33}$$

Write the power series u in the form $u = \sum_{n\geq 0} u_n T^n$ with $(u_n)_{n\geq 0} \in \mathbf{k}^{\mathbb{N}}$. Thus, $u^{\star}(f) = \sum_{n\geq 0} u_n f^{\star n}$ (by the definition of $u^{\star}(f)$).

The definition of Φ yields $\Phi(f) = s \circ f$. Each $n \in \mathbb{N}$ satisfies

$$\begin{aligned} s \circ f^{\star n} &= \Phi(f^{\star n}) \qquad (\text{since } \Phi(f^{\star n}) = s \circ f^{\star n} \text{ (by the definition of } \Phi)) \\ &= \left(\underbrace{\Phi(f)}_{=s\circ f}\right)^{\star n} \qquad (\text{since } \Phi \text{ is a } \mathbf{k}\text{-algebra homomorphism}) \\ &= (s \circ f)^{\star n}. \end{aligned} \tag{13.40.34}$$

Thus, each $n \in \mathbb{N}$ satisfies

$$s(f^{\star n}(x)) = \underbrace{(s \circ f^{\star n})}_{\substack{=(s\circ f)^{\star n} \\ \text{(by (13.40.34))}}}(x) = (s \circ f)^{\star n}(x). \tag{13.40.35}$$

[497] *Proof of (13.40.32):* Let $g \in \{1, 2, \ldots, k\}$. Then, $n_g \in \{n_1, n_2, \ldots, n_k\}$. Hence, (13.40.31) (applied to $n = n_g$) yields $n_g \leq N$. Hence, $N \geq n_g$, so that $n > N \geq n_g$. Hence, (13.40.27) (applied to n and n_g instead of i and n) yields $f^{\star n}(C_{n_g}) = 0$. But $x_g \in C_{n_g}$ (by the definition of n_g) and thus $f^{\star n}\left(\underbrace{x_g}_{\in C_{n_g}}\right) \in f^{\star n}(C_{n_g}) = 0$. In other words, $f^{\star n}(x_g) = 0$. This proves (13.40.32).

We have $f \in \mathfrak{n}(C, A)$. In other words, f is a pointwise $\star$-nilpotent map in $\operatorname{Hom}(C, A)$. Thus, the family $\left(f^{\star n}\right)_{n \in \mathbb{N}}$ is pointwise finitely supported. Hence, Proposition 1.7.7 (applied to $Q = \mathbb{N}$, $\left(f_q\right)_{q \in Q} = \left(f^{\star n}\right)_{n \in \mathbb{N}}$ and $\left(\lambda_q\right)_{q \in Q} = \left(u_n\right)_{n \in \mathbb{N}}$) shows that the family $\left(u_n f^{\star n}\right)_{n \in \mathbb{N}} \in \left(\operatorname{Hom}(C, A)\right)^{\mathbb{N}}$ is pointwise finitely supported. In other words, for each $x \in C$, the family $\left(\left(u_n f^{\star n}\right)(x)\right)_{n \in \mathbb{N}} \in A^{\mathbb{N}}$ is finitely supported. In other words, for each $x \in C$,

$$\text{the family } \left(u_n f^{\star n}(x)\right)_{n \in \mathbb{N}} \in A^{\mathbb{N}} \text{ is finitely supported} \tag{13.40.36}$$

(since each $n \in \mathbb{N}$ satisfies $\left(u_n f^{\star n}\right)(x) = u_n f^{\star n}(x)$).

Also, recall that the family $\left(f^{\star n}\right)_{n \in \mathbb{N}}$ is pointwise finitely supported. In other words, for each $x \in C$,

$$\text{the family } \left(f^{\star n}(x)\right)_{n \in \mathbb{N}} \in A^{\mathbb{N}} \text{ is finitely supported.} \tag{13.40.37}$$

Let $x \in C$. Then, (13.40.37) shows that the family $\left(f^{\star n}(x)\right)_{n \in \mathbb{N}} \in A^{\mathbb{N}}$ is finitely supported. Hence, Fact H.7 (applied to $V = A$, $W = B$, $\varphi = s$, $Q = \mathbb{N}$ and $\left(v_q\right)_{q \in Q} = \left(f^{\star n}(x)\right)_{n \in \mathbb{N}}$) shows that the family $\left(s\left(f^{\star n}(x)\right)\right)_{n \in \mathbb{N}} \in B^{\mathbb{N}}$ is also finitely supported, and that it satisfies $\sum_{n \in \mathbb{N}} s\left(f^{\star n}(x)\right) = s\left(\sum_{n \in \mathbb{N}} f^{\star n}(x)\right)$. Thus, in particular, the family $\left(s\left(f^{\star n}(x)\right)\right)_{n \in \mathbb{N}} \in B^{\mathbb{N}}$ is finitely supported. Since each $n \in \mathbb{N}$ satisfies (13.40.35), this rewrites as follows:

$$\text{The family } \left((s \circ f)^{\star n}(x)\right)_{n \in \mathbb{N}} \in B^{\mathbb{N}} \text{ is finitely supported.} \tag{13.40.38}$$

From (13.40.36), we know that the family $\left(u_n f^{\star n}(x)\right)_{n \in \mathbb{N}} \in A^{\mathbb{N}}$ is finitely supported. Hence, Fact H.7 (applied to $V = A$, $W = B$, $\varphi = s$, $Q = \mathbb{N}$ and $\left(v_q\right)_{q \in Q} = \left(u_n f^{\star n}(x)\right)_{n \in \mathbb{N}}$) shows that the family $\left(s\left(u_n f^{\star n}(x)\right)\right)_{n \in \mathbb{N}} \in B^{\mathbb{N}}$ is also finitely supported, and that it satisfies

$$\sum_{n \in \mathbb{N}} s\left(u_n f^{\star n}(x)\right) = s\left(\sum_{n \in \mathbb{N}} u_n f^{\star n}(x)\right). \tag{13.40.39}$$

From (13.40.39), we obtain

$$\begin{aligned} s\left(\sum_{n \in \mathbb{N}} u_n f^{\star n}(x)\right) &= \sum_{n \in \mathbb{N}} \underbrace{s\left(u_n f^{\star n}(x)\right)}_{\substack{= u_n s\left(f^{\star n}(x)\right) \\ \text{(since the map } s \text{ is } \mathbf{k}\text{-linear)}}} = \sum_{n \in \mathbb{N}} u_n \underbrace{s\left(f^{\star n}(x)\right)}_{\substack{= (s \circ f)^{\star n}(x) \\ \text{(by (13.40.35))}}} \\ &= \sum_{n \in \mathbb{N}} u_n (s \circ f)^{\star n}(x). \end{aligned} \tag{13.40.40}$$

Now, forget that we fixed x. We thus have proven that each $x \in C$ satisfies (13.40.40) and (13.40.38).

In particular, each $x \in C$ satisfies (13.40.38). In other words, for each $x \in C$, the family $\left((s \circ f)^{\star n}(x)\right)_{n \in \mathbb{N}} \in B^{\mathbb{N}}$ is finitely supported. In other words, the family $\left((s \circ f)^{\star n}\right)_{n \in \mathbb{N}} \in \left(\operatorname{Hom}(C, B)\right)^{\mathbb{N}}$ is pointwise finitely supported. In other words, the map $s \circ f : C \to B$ is pointwise $\star$-nilpotent. In other words, $s \circ f \in \mathfrak{n}(C, B)$. Hence, $u^{\star}(s \circ f)$ is well-defined.

Recall that $u = \sum_{n \geq 0} u_n T^n$. Thus,

$$u^{\star}(s \circ f) = \underbrace{\sum_{n \geq 0}}_{= \sum_{n \in \mathbb{N}}} u_n (s \circ f)^{\star n} = \sum_{n \in \mathbb{N}} u_n (s \circ f)^{\star n}.$$

Hence, each $x \in C$ satisfies

$$\begin{aligned} \left(u^{\star}(s \circ f)\right)(x) &= \left(\sum_{n \in \mathbb{N}} u_n (s \circ f)^{\star n}\right)(x) = \sum_{n \in \mathbb{N}} u_n (s \circ f)^{\star n}(x) = s\left(\sum_{n \in \mathbb{N}} u_n f^{\star n}(x)\right) \qquad \text{(by (13.40.40))} \\ &= \left(s \circ \left(u^{\star}(f)\right)\right)(x) \end{aligned}$$

(since

$$(s \circ (u^\star (f)))(x) = s\left(\underbrace{(u^\star (f))}_{=\sum_{n\geq 0} u_n f^{\star n}}(x)\right) = s\left(\underbrace{\left(\sum_{n\geq 0} u_n f^{\star n}\right)(x)}_{=\sum_{n\geq 0} u_n f^{\star n}(x)}\right)$$

$$= s\left(\underbrace{\sum_{n\geq 0}}_{=\sum_{n\in\mathbb{N}}} u_n f^{\star n}(x)\right) = s\left(\sum_{n\in\mathbb{N}} u_n f^{\star n}(x)\right)$$

). In other words, we have $u^\star (s \circ f) = s \circ (u^\star (f))$.

We have now proven that $s \circ f \in \mathfrak{n}(C, B)$ and $u^\star (s \circ f) = s \circ (u^\star (f))$. This proves (13.40.33). Hence, Proposition 1.7.11(i) is proven.

(j) Let C be a connected graded $\mathbf{k}$-bialgebra. Let $F : C \to A$ be a $\mathbf{k}$-algebra homomorphism. We must prove that $F - u_A\epsilon_C \in \mathfrak{n}(C, A)$.

We have $F(1_C) = 1_A$ (since F is a $\mathbf{k}$-algebra homomorphism). But the axioms of a $\mathbf{k}$-bialgebra yield $\epsilon_C(1_C) = 1$.

The definition of the map u_A yields $u_A(1) = 1 \cdot 1_A = 1_A$. Now,

$$(F - u_A\epsilon_C)(1_C) = \underbrace{F(1_C)}_{=1_A} - \underbrace{(u_A\epsilon_C)(1_C)}_{=u_A(\epsilon_C(1_C))} = 1_A - u_A\left(\underbrace{\epsilon_C(1_C)}_{=1}\right) = 1_A - \underbrace{u_A(1)}_{=1_A} = 1_A - 1_A = 0.$$

But Exercise 1.3.20(c) (applied to C instead of A) shows that $C_0 = \mathbf{k} \cdot 1_C$. Applying the map $F - u_A\epsilon_C$ to both sides of this relation, we obtain

$$(F - u_A\epsilon_C)(C_0) = (F - u_A\epsilon_C)(\mathbf{k} \cdot 1_C) = \mathbf{k} \cdot \underbrace{(F - u_A\epsilon_C)(1_C)}_{=0} \qquad \text{(since the map } F - u_A\epsilon_C \text{ is } \mathbf{k}\text{-linear)}$$

$$= \mathbf{k} \cdot 0 = 0.$$

Hence, Proposition 1.7.11(h) (applied to $f = F - u_A\epsilon_C$) shows that $F - u_A\epsilon_C \in \mathfrak{n}(C, A)$. Thus, Proposition 1.7.11(j) is proven. □

Thus, Proposition 1.7.11 is proven, so that Exercise 1.7.13 is solved.

13.41. **Solution to Exercise 1.7.20.** *Solution to Exercise 1.7.20.*

Proof of Proposition 1.7.15. Proposition 1.7.15 is a well-known fact that is often used in enumerative combinatorics (for computing generating functions). We shall give a purely algebraic proof (somewhat similar to the one given in [138, Example 7.67]). Other proofs (some combinatorial, some analytic) can be found in the literature.

The proof will rely on several simple facts about power series. Keep in mind that all of the following facts assume that $\mathbf{k}$ is a commutative $\mathbb{Q}$-algebra.

First, we notice that

$$\overline{\exp} = \underbrace{\exp}_{=\sum_{n\geq 0}\frac{1}{n!}T^n} - 1 = \sum_{n\geq 0}\frac{1}{n!}T^n - 1 = \underbrace{\frac{1}{0!}}_{=\frac{1}{1}=1}\underbrace{T^0}_{=1} + \sum_{n\geq 1}\frac{1}{n!}T^n - 1$$

$$\text{(here, we have split off the addend for } n=0 \text{ from the sum)}$$

$$= 1 + \sum_{n\geq 1}\frac{1}{n!}T^n - 1 = \sum_{n\geq 1}\frac{1}{n!}T^n. \tag{13.41.1}$$

Hence, the power series $\overline{\exp}$ has constant term 0. Hence, the power series $\overline{\log}\left[\overline{\exp}\right]$ is well-defined.

Also,

$$\overline{\log} = \log\left(1+T\right) = \sum_{n\geq 1}\frac{\left(-1\right)^{n-1}}{n}T^n. \tag{13.41.2}$$

Hence, the power series $\overline{\log}$ has constant term 0. Hence, the power series $\overline{\exp}\left[\overline{\log}\right]$ is well-defined.

For each $n \geq 1$, we have

$$\left(\text{the constant term of } \overline{\log}^n\right) = 0 \tag{13.41.3}$$

[498].

Substituting $\overline{\log}$ for T on both sides of the equality (13.41.1), we obtain

$$\overline{\exp}\left[\overline{\log}\right] = \sum_{n\geq 1}\frac{1}{n!}\overline{\log}^n.$$

Hence,

$$\left(\text{the constant term of } \overline{\exp}\left[\overline{\log}\right]\right) = \left(\text{the constant term of } \sum_{n\geq 1}\frac{1}{n!}\overline{\log}^n\right)$$
$$= \sum_{n\geq 1}\frac{1}{n!}\underbrace{\left(\text{the constant term of } \overline{\log}^n\right)}_{\substack{=0 \\ \text{(by (13.41.3))}}} = \sum_{n\geq 1}\frac{1}{n!}0 = 0.$$

In other words, the power series $\overline{\exp}\left[\overline{\log}\right]$ has constant term 0. A similar argument (with the roles of $\overline{\exp}$ and $\overline{\log}$ switched) shows that the power series $\overline{\log}\left[\overline{\exp}\right]$ has constant term 0.

> *Fact I.1:* Let $u \in \mathbf{k}\left[\left[T\right]\right]$ and $v \in \mathbf{k}\left[\left[T\right]\right]$ be two power series having the same constant term. Assume that $\frac{d}{dT}u = \frac{d}{dT}v$. Then, $u = v$.

[*Proof of Fact I.1:* Write the power series u in the form $u = \sum_{n\geq 0} u_n T^n$ with $\left(u_n\right)_{n\geq 0} \in \mathbf{k}^{\mathbb{N}}$. Thus, $\frac{d}{dT}u = \sum_{n\geq 1} n u_n T^{n-1}$ (by the definition of the derivative).

Write the power series v in the form $v = \sum_{n\geq 0} v_n T^n$ with $\left(v_n\right)_{n\geq 0} \in \mathbf{k}^{\mathbb{N}}$. Thus, $\frac{d}{dT}v = \sum_{n\geq 1} n v_n T^{n-1}$ (by the definition of the derivative).

Now,

$$\sum_{n\geq 1} n u_n T^{n-1} = \frac{d}{dT}u = \frac{d}{dT}v = \sum_{n\geq 1} n v_n T^{n-1}.$$

Comparing coefficients in front of T^{n-1} on both sides of this equality, we obtain

$$n u_n = n v_n \qquad \text{for each integer } n \geq 1. \tag{13.41.4}$$

[498] *Proof of (13.41.3):* Let $n \geq 1$. The power series $\overline{\log}$ is divisible by T (since it has constant term 0). Hence, the power series $\overline{\log}^n$ is divisible by T^n. Thus, the power series $\overline{\log}^n$ is also divisible by T (since T^n is divisible by T (since $n \geq 1$)), and therefore has constant term 0. In other words, we have $\left(\text{the constant term of } \overline{\log}^n\right) = 0$. This proves (13.41.3).

On the other hand, the power series u has constant term u_0 (since $u = \sum_{n\geq 0} u_n T^n$), and the power series v has constant term v_0 (similarly). Thus, the constant terms of u and v are u_0 and v_0, respectively. Therefore, $u_0 = v_0$ (since the power series u and v have the same constant term).

Now, each $n \in \mathbb{N}$ satisfies $u_n = v_n$ [499]. Hence, $\sum_{n\geq 0} \underbrace{u_n}_{=v_n} T^n = \sum_{n\geq 0} v_n T^n$. Thus, $u = \sum_{n\geq 0} u_n T^n = \sum_{n\geq 0} v_n T^n = v$. This proves Fact I.1.]

Fact I.2: Let $w \in \mathbf{k}[[T]]$ be a power series having constant term 0. Then,

$$\frac{d}{dT}\left(\overline{\exp}[w]\right) = \left(\frac{d}{dT} w\right) \cdot \exp[w] \tag{13.41.5}$$

and

$$\frac{d}{dT}\left(\overline{\log}[w]\right) = \left(\frac{d}{dT} w\right) \cdot \frac{1}{1+w}. \tag{13.41.6}$$

[*Proof of Fact I.2:* It is easy to see that each positive integer n satisfies

$$\frac{d}{dT}\left(w^n\right) = n\left(\frac{d}{dT} w\right) w^{n-1}. \tag{13.41.7}$$

(Indeed, (13.41.7) can be proven by a straightforward induction on n, using the Leibniz identity $\frac{d}{dT}(uv) = \left(\frac{d}{dT}u\right) v + u\frac{d}{dT}v$.)

Substituting w for T on both sides of the equality (13.41.1), we obtain

$$\overline{\exp}[w] = \sum_{n\geq 1} \frac{1}{n!} w^n.$$

Applying the operator $\frac{d}{dT}$ to this equality, we find

$$\frac{d}{dT}\overline{\exp}[w] = \frac{d}{dT}\sum_{n\geq 1}\frac{1}{n!}w^n = \sum_{n\geq 1}\frac{1}{n!}\cdot \underbrace{\frac{d}{dT}\left(w^n\right)}_{\substack{=n\left(\frac{d}{dT}w\right)w^{n-1} \\ \text{(by (13.41.7))}}} = \sum_{n\geq 1}\underbrace{\frac{1}{n!}\cdot n}_{=\frac{1}{(n-1)!}}\left(\frac{d}{dT}w\right)w^{n-1}$$

$$= \sum_{n\geq 1}\frac{1}{(n-1)!}\left(\frac{d}{dT}w\right)w^{n-1} = \sum_{n\geq 0}\frac{1}{n!}\left(\frac{d}{dT}w\right)w^n$$

$$\left(\text{here, we have substituted } n \text{ for } n-1 \text{ in the sum}\right).$$

Comparing this with

$$\left(\frac{d}{dT}w\right)\cdot \underbrace{\exp[w]}_{\substack{=\sum_{n\geq 0}\frac{1}{n!}w^n \\ \left(\text{since } \exp=\sum_{n\geq 0}\frac{1}{n!}T^n\right)}} = \left(\frac{d}{dT}w\right)\cdot\sum_{n\geq 0}\frac{1}{n!}w^n = \sum_{n\geq 0}\frac{1}{n!}\left(\frac{d}{dT}w\right)w^n,$$

we obtain $\frac{d}{dT}\left(\overline{\exp}[w]\right) = \left(\frac{d}{dT}w\right)\cdot\exp[w]$. This proves (13.41.5).

[499] *Proof.* Let $n \in \mathbb{N}$. We must prove that $u_n = v_n$.

If $n = 0$, then this follows immediately from $u_0 = v_0$. Hence, we WLOG assume that we don't have $n = 0$. Thus, $n \geq 1$ (since $n \in \mathbb{N}$). Therefore, (13.41.4) yields $nu_n = nv_n$. We can multiply both sides of this equality by $\frac{1}{n}$ (since $\mathbf{k}$ is a $\mathbb{Q}$-algebra), and thus obtain $u_n = v_n$, qed.

Substituting w for T on both sides of the equality (13.41.2), we obtain

$$\overline{\log}[w] = \sum_{n\geq 1} \frac{(-1)^{n-1}}{n} w^n.$$

Applying the operator $\dfrac{d}{dT}$ to this equality, we find

$$\frac{d}{dT}\overline{\log}[w] = \frac{d}{dT}\sum_{n\geq 1} \frac{(-1)^{n-1}}{n} w^n = \sum_{n\geq 1} \frac{(-1)^{n-1}}{n} \cdot \underbrace{\frac{d}{dT}(w^n)}_{\substack{=n\left(\frac{d}{dT}w\right)w^{n-1} \\ \text{(by (13.41.7))}}} = \sum_{n\geq 1} \frac{(-1)^{n-1}}{n} \cdot n\left(\frac{d}{dT}w\right)w^{n-1}$$
$$= \sum_{n\geq 1} (-1)^{n-1}\left(\frac{d}{dT}w\right)w^{n-1} = \sum_{n\geq 0} (-1)^{n}\left(\frac{d}{dT}w\right)w^{n}$$
$$\text{(here, we have substituted } n \text{ for } n-1 \text{ in the sum)}.$$

Comparing this with

$$\left(\frac{d}{dT}w\right) \cdot \underbrace{\frac{1}{1+w}}_{=\sum_{n\geq 0}(-1)^n w^n} = \left(\frac{d}{dT}w\right) \cdot \sum_{n\geq 0}(-1)^n w^n = \sum_{n\geq 0}(-1)^n \left(\frac{d}{dT}w\right) w^n,$$

we obtain $\dfrac{d}{dT}\left(\overline{\log}[w]\right) = \left(\dfrac{d}{dT}w\right) \cdot \dfrac{1}{1+w}$. This proves (13.41.6). Thus, Fact I.2 is proven.]

Fact I.3: Let $u \in \mathbf{k}[[T]]$ and $v \in \mathbf{k}[[T]]$ be two power series having constant term 1. Assume that $\left(\dfrac{d}{dT}u\right) \cdot v = \left(\dfrac{d}{dT}v\right) \cdot u$. Then, $u = v$.

[*Proof of Fact I.3:* The power series v has constant term 1, and thus has a multiplicative inverse v^{-1}. The Leibniz rule (applied to v and v^{-1}) yields

$$\frac{d}{dT}\left(v \cdot v^{-1}\right) = \left(\frac{d}{dT}v\right)v^{-1} + v\frac{d}{dT}\left(v^{-1}\right).$$

Comparing this with $\dfrac{d}{dT}\underbrace{\left(v \cdot v^{-1}\right)}_{=1} = \dfrac{d}{dT}1 = 0$, we obtain $\left(\dfrac{d}{dT}v\right)v^{-1} + v\dfrac{d}{dT}\left(v^{-1}\right) = 0$. Solving this equality for $\dfrac{d}{dT}\left(v^{-1}\right)$, we find

$$\frac{d}{dT}\left(v^{-1}\right) = -\frac{1}{v}\left(\frac{d}{dT}v\right)v^{-1} = -v^{-2}\left(\frac{d}{dT}v\right).$$

Now, the Leibniz rule (applied to u and v^{-1}) yields

$$\frac{d}{dT}\left(uv^{-1}\right) = \left(\frac{d}{dT}u\right)v^{-1} + u\underbrace{\frac{d}{dT}\left(v^{-1}\right)}_{=-v^{-2}\left(\frac{d}{dT}v\right)} = \left(\frac{d}{dT}u\right)v^{-1} + u\left(-v^{-2}\left(\frac{d}{dT}v\right)\right)$$
$$= v^{-2}\underbrace{\left(\left(\frac{d}{dT}u\right)\cdot v - \left(\frac{d}{dT}v\right)\cdot u\right)}_{\substack{=0 \\ \left(\text{since } \left(\frac{d}{dT}u\right)\cdot v = \left(\frac{d}{dT}v\right)\cdot u\right)}} = v^{-2}0 = 0 = \frac{d}{dT}1.$$

Moreover, the power series uv^{-1} and 1 have the same constant term[500]. Hence, Fact I.1 (applied to uv^{-1} and 1 instead of u and v) shows that $uv^{-1} = 1$. Thus, $u = v$. This proves Fact I.3.]

The equality (13.41.6) (applied to $w = T$) yields $\dfrac{d}{dT}\left(\overline{\log}\left[T\right]\right) = \underbrace{\left(\dfrac{d}{dT}T\right)}_{=1} \cdot \dfrac{1}{1+T} = \dfrac{1}{1+T}$. In other words, $\dfrac{d}{dT}\overline{\log} = \dfrac{1}{1+T}$ (since $\overline{\log} = \overline{\log}\left[T\right]$).

Now, (13.41.5) (applied to $w = \overline{\log}$) shows that

$$\frac{d}{dT}\left(\overline{\exp}\left[\overline{\log}\right]\right) = \underbrace{\left(\frac{d}{dT}\overline{\log}\right)}_{=\frac{1}{1+T}} \cdot \exp\left[\overline{\log}\right] = \frac{1}{1+T} \cdot \exp\left[\overline{\log}\right].$$

But $\overline{\exp} = \exp - 1$ and thus $\exp = \overline{\exp} + 1$. Substituting $\overline{\log}$ for T in this equality, we find $\exp\left[\overline{\log}\right] = \overline{\exp}\left[\overline{\log}\right] + 1$. Hence,

$$\frac{d}{dT}\left(\exp\left[\overline{\log}\right]\right) = \frac{d}{dT}\left(\overline{\exp}\left[\overline{\log}\right] + 1\right) = \frac{d}{dT}\overline{\exp}\left[\overline{\log}\right] + \underbrace{\frac{d}{dT}1}_{=0} = \frac{d}{dT}\overline{\exp}\left[\overline{\log}\right] = \frac{1}{1+T} \cdot \exp\left[\overline{\log}\right].$$

Multiplying this equality by $1 + T$, we find

$$\left(\frac{d}{dT}\left(\exp\left[\overline{\log}\right]\right)\right) \cdot (1+T) = \exp\left[\overline{\log}\right].$$

Comparing this with $\underbrace{\left(\dfrac{d}{dT}(1+T)\right)}_{=1} \cdot \exp\left[\overline{\log}\right] = \exp\left[\overline{\log}\right]$, we find

$$\left(\frac{d}{dT}\left(\exp\left[\overline{\log}\right]\right)\right) \cdot (1+T) = \left(\frac{d}{dT}(1+T)\right) \cdot \exp\left[\overline{\log}\right].$$

Since both power series $\exp\left[\overline{\log}\right]$ and $1 + T$ have constant term 1 [501], we can thus apply Fact I.3 to $u = \exp\left[\overline{\log}\right]$ and $v = 1 + T$. We thus conclude that $\exp\left[\overline{\log}\right] = 1 + T$. Comparing this with $\exp\left[\overline{\log}\right] = \overline{\exp}\left[\overline{\log}\right] + 1$, we obtain $\overline{\exp}\left[\overline{\log}\right] + 1 = 1 + T$. Subtracting 1 from this equality, we find $\overline{\exp}\left[\overline{\log}\right] = T$.

The equality (13.41.5) (applied to $w = T$) yields $\dfrac{d}{dT}\left(\overline{\exp}\left[T\right]\right) = \underbrace{\left(\dfrac{d}{dT}T\right)}_{=1} \cdot \underbrace{\exp\left[T\right]}_{=\exp} = \exp$. In other words, $\dfrac{d}{dT}\overline{\exp} = \exp$ (since $\overline{\exp} = \overline{\exp}\left[T\right]$).

On the other hand, (13.41.6) (applied to $w = \overline{\exp}$) shows that

$$\frac{d}{dT}\left(\overline{\log}\left[\overline{\exp}\right]\right) = \underbrace{\left(\frac{d}{dT}\overline{\exp}\right)}_{=\exp=\overline{\exp}+1=1+\overline{\exp}} \cdot \frac{1}{1+\overline{\exp}} = \left(1+\overline{\exp}\right) \cdot \frac{1}{1+\overline{\exp}} = 1 = \frac{d}{dT}T.$$

Since the two power series $\overline{\log}\left[\overline{\exp}\right]$ and T have the same constant term[502], we can thus apply Fact I.1 to $u = \overline{\log}\left[\overline{\exp}\right]$ and $v = T$. We thus conclude that $\overline{\log}\left[\overline{\exp}\right] = T$. The proof of Proposition 1.7.15 is thus complete. □

[500] *Proof.* The power series v has constant term 1. Hence, its inverse v^{-1} has constant term $1^{-1} = 1$. Now, both power series u and v^{-1} have constant term 1. Hence, their product uv^{-1} has constant term $1 \cdot 1 = 1$. Since the power series 1 also has constant term 1, this shows that the power series uv^{-1} and 1 have the same constant term (namely, 1).

[501] *Proof.* It is clear that the power series $1+T$ has constant term 1. Thus, it remain to prove that the power series $\exp\left[\overline{\log}\right]$ has constant term 1.

Recall that the power series $\overline{\exp}\left[\overline{\log}\right]$ has constant term 0. Hence, the power series $\overline{\exp}\left[\overline{\log}\right] + 1$ has constant term $0+1 = 1$. In other words, the power series $\exp\left[\overline{\log}\right]$ has constant term 1 (since $\exp\left[\overline{\log}\right] = \overline{\exp}\left[\overline{\log}\right] + 1$). Qed.

[502] This is because the power series $\overline{\log}\left[\overline{\exp}\right]$ has constant term 0, and the power series T also has constant term 0.

Proof of Lemma 1.7.16. The power series $\overline{\log}$ has constant term 0 (by Proposition 1.7.15). Hence, Proposition 1.7.11(f) (applied to $v = \overline{\log}$ and $f = g - u_A\epsilon_C$) yields that $\overline{\log}^{\star}(g - u_A\epsilon_C) \in \mathfrak{n}(C, A)$. This proves Lemma 1.7.16. □

Proof of Proposition 1.7.18. (a) Let $f \in \mathfrak{n}(C, A)$. We must prove that $\exp^{\star} f - u_A\epsilon_C \in \mathfrak{n}(C, A)$ and $\log^{\star}(\exp^{\star} f) = f$.

The power series $\overline{\exp}$ has constant term 0 (by Proposition 1.7.15). Hence, Proposition 1.7.11(f) (applied to $v = \overline{\exp}$) yields that $\overline{\exp}^{\star} f \in \mathfrak{n}(C, A)$. But $\overline{\exp} = \exp - 1$, so that $\exp = \overline{\exp} + 1$. Hence,

$$\begin{aligned}\exp^{\star}(f) = (\overline{\exp}+1)^{\star}(f) &= \underbrace{\overline{\exp}^{\star}(f)}_{=\overline{\exp}^{\star} f} + \underbrace{1^{\star}(f)}_{\substack{=u_A\epsilon_C \\ \text{(by (1.7.6))}}} \qquad \text{(by (1.7.2))} \\ &= \overline{\exp}^{\star} f + u_A\epsilon_C.\end{aligned}$$

Solving this equation for $\overline{\exp}^{\star} f$, we obtain $\overline{\exp}^{\star} f = \exp^{\star}(f) - u_A\epsilon_C = \exp^{\star} f - u_A\epsilon_C$. Hence, $\exp^{\star} f - u_A\epsilon_C = \overline{\exp}^{\star} f \in \mathfrak{n}(C, A)$. It thus remains to prove that $\log^{\star}(\exp^{\star} f) = f$.

The map $\exp^{\star} f$ satisfies $\exp^{\star} f - u_A\epsilon_C \in \mathfrak{n}(C, A)$. Hence, the map $\log^{\star}(\exp^{\star} f) \in \operatorname{Hom}(C, A)$ is well-defined, and satisfies

$$\begin{aligned}\log^{\star}(\exp^{\star} f) &= \overline{\log}^{\star}\left(\underbrace{\exp^{\star} f - u_A\epsilon_C}_{=\overline{\exp}^{\star} f}\right) \qquad \text{(by the definition of } \log^{\star}(\exp^{\star} f)\text{)} \\ &= \overline{\log}^{\star}(\overline{\exp}^{\star} f). \qquad (13.41.8)\end{aligned}$$

But the power series $\overline{\exp}$ has constant term 0. Thus, (1.7.7) (applied to $u = \overline{\log}$ and $v = \overline{\exp}$) yields $\left(\overline{\log}\left[\overline{\exp}\right]\right)^{\star}(f) = \overline{\log}^{\star}(\overline{\exp}^{\star}(f)) = \overline{\log}^{\star}(\overline{\exp}^{\star} f)$. Since $\overline{\log}\left[\overline{\exp}\right] = T$ (by Proposition 1.7.15), this rewrites as $T^{\star}(f) = \overline{\log}^{\star}(\overline{\exp}^{\star} f)$.

Applying (1.7.1) to $k = 1$, we obtain $\left(T^1\right)^{\star}(f) = f^{\star 1} = f$. Since $T^1 = T$, this rewrites as $T^{\star}(f) = f$. Compared with $T^{\star}(f) = \overline{\log}^{\star}(\overline{\exp}^{\star} f)$, this yields $\overline{\log}^{\star}(\overline{\exp}^{\star} f) = f$. Now, (13.41.8) becomes $\log^{\star}(\exp^{\star} f) = \overline{\log}^{\star}(\overline{\exp}^{\star} f) = f$. This proves Proposition 1.7.18(a).

(b) Let $g \in \operatorname{Hom}(C, A)$ be such that $g - u_A\epsilon_C \in \mathfrak{n}(C, A)$. We must prove that $\exp^{\star}(\log^{\star} g) = g$.

Set $f = g - u_A\epsilon_C$. Thus, $f = g - u_A\epsilon_C \in \mathfrak{n}(C, A)$.

The power series $\overline{\log}$ has constant term 0 (by Proposition 1.7.15). Hence, Proposition 1.7.11(f) (applied to $v = \overline{\log}$) yields that $\overline{\log}^{\star} f \in \mathfrak{n}(C, A)$.

The definition of $\log^{\star} g$ yields $\log^{\star} g = \overline{\log}^{\star}\underbrace{(g - u_A\epsilon_C)}_{=f} = \overline{\log}^{\star} f \in \mathfrak{n}(C, A)$.

But the power series $\overline{\log}$ has constant term 0. Thus, (1.7.7) (applied to $u = \overline{\exp}$ and $v = \overline{\log}$) yields $\left(\overline{\exp}\left[\overline{\log}\right]\right)^{\star}(f) = \overline{\exp}^{\star}\left(\overline{\log}^{\star}(f)\right) = \overline{\exp}^{\star}\left(\overline{\log}^{\star} f\right)$. Since $\overline{\exp}\left[\overline{\log}\right] = T$ (by Proposition 1.7.15), this rewrites as $T^{\star}(f) = \overline{\exp}^{\star}\left(\overline{\log}^{\star} f\right)$.

Applying (1.7.1) to $k = 1$, we obtain $\left(T^1\right)^{\star}(f) = f^{\star 1} = f$. Since $T^1 = T$, this rewrites as $T^{\star}(f) = f$. Compared with $T^{\star}(f) = \overline{\exp}^{\star}\left(\overline{\log}^{\star} f\right)$, this yields $\overline{\exp}^{\star}\left(\overline{\log}^{\star} f\right) = f$.

But $\overline{\exp} = \exp - 1$, so that $\exp = \overline{\exp} + 1$. Hence,

$$\begin{aligned}\exp^{\star}(\log^{\star} g) = (\overline{\exp}+1)^{\star}\left(\underbrace{\log^{\star} g}_{=\overline{\log}^{\star} f}\right) &= (\overline{\exp}+1)^{\star}\left(\overline{\log}^{\star} f\right) = \underbrace{\overline{\exp}^{\star}\left(\overline{\log}^{\star} f\right)}_{=f=g-u_A\epsilon_C} + \underbrace{1^{\star}\left(\overline{\log}^{\star} f\right)}_{\substack{=u_A\epsilon_C \\ \text{(by (1.7.6))}}} \\ &\qquad \text{(by an application of (1.7.2))} \\ &= (g - u_A\epsilon_C) + u_A\epsilon_C = g.\end{aligned}$$

This proves Proposition 1.7.18(b).

(c) Let $f, g \in \mathfrak{n}(C, A)$ be such that $f \star g = g \star f$. Then, Proposition 1.7.11(c) shows that $f + g \in \mathfrak{n}(C, A)$. Hence, $\exp^{\star}(f + g)$ is well-defined.

Let us recall the following facts, which have been proven in the proof of Proposition 1.7.11(c) (during the solution to Exercise 1.7.13):

- We have

$$(f+g)^{\star n} = \sum_{i=0}^{n} \binom{n}{i} f^{\star i} \star g^{\star (n-i)} \qquad \text{for each } n \in \mathbb{N}. \tag{13.41.9}$$

- The family $\left(f^{\star n}\right)_{n\in\mathbb{N}}$ is pointwise finitely supported.
- For each $x \in C$,

$$\text{the family } \left(\left(f^{\star q} \star g^{\star r}\right)(x)\right)_{(q,r)\in\mathbb{N}\times\mathbb{N}} \in A^{\mathbb{N}\times\mathbb{N}} \text{ is finitely supported.} \tag{13.41.10}$$

Furthermore, the family $\left(\dfrac{f^{\star q}}{q!}\right)_{q\in\mathbb{N}}$ is pointwise finitely supported[503]. Similarly, the family $\left(\dfrac{g^{\star r}}{r!}\right)_{r\in\mathbb{N}}$ is pointwise finitely supported. Hence, Proposition 1.7.6 (applied to $Q=\mathbb{N}$, $R=\mathbb{N}$, $\left(f_q\right)_{q\in Q} = \left(\dfrac{f^{\star q}}{q!}\right)_{q\in\mathbb{N}}$ and $\left(g_r\right)_{r\in R} = \left(\dfrac{g^{\star r}}{r!}\right)_{r\in\mathbb{N}}$) shows that the family $\left(\dfrac{f^{\star q}}{q!} \star \dfrac{g^{\star r}}{r!}\right)_{(q,r)\in\mathbb{N}\times\mathbb{N}} \in \left(\operatorname{Hom}(C,A)\right)^{\mathbb{N}\times\mathbb{N}}$ is pointwise finitely supported, and that it satisfies

$$\sum_{(q,r)\in\mathbb{N}\times\mathbb{N}} \left(\frac{f^{\star q}}{q!} \star \frac{g^{\star r}}{r!}\right) = \left(\sum_{q\in\mathbb{N}} \frac{f^{\star q}}{q!}\right) \star \left(\sum_{r\in\mathbb{N}} \frac{g^{\star r}}{r!}\right). \tag{13.41.11}$$

Let $x \in C$. Then, the family $\left(\left(f^{\star q} \star g^{\star r}\right)(x)\right)_{(q,r)\in\mathbb{N}\times\mathbb{N}} \in A^{\mathbb{N}\times\mathbb{N}}$ is finitely supported (by (13.41.10)). In other words, all but finitely many $(q,r) \in \mathbb{N}\times\mathbb{N}$ satisfy $\left(f^{\star q} \star g^{\star r}\right)(x) = 0$. In other words, there exists a finite subset K of $\mathbb{N}\times\mathbb{N}$ such that

$$\text{each } (q,r) \in (\mathbb{N}\times\mathbb{N}) \setminus K \text{ satisfies } \left(f^{\star q} \star g^{\star r}\right)(x) = 0. \tag{13.41.12}$$

Consider this K. Then,

$$\text{each } (q,r) \in \mathbb{N}\times\mathbb{N} \text{ satisfying } (q,r) \notin K \text{ satisfies } \left(\frac{f^{\star q}}{q!} \star \frac{g^{\star r}}{r!}\right)(x) = 0 \tag{13.41.13}$$

[504].

Let $Q = \{u+v \mid (u,v) \in K\}$. Thus, Q is a finite set (since K is a finite set). Also, $Q \subset \mathbb{N}$ (since each $(u,v) \in K$ satisfies $(u,v) \in K \subset \mathbb{N}\times\mathbb{N}$ and thus $u+v \in \mathbb{N}$). Thus, the set $\mathbb{N}$ is the union of its two disjoint subsets Q and $\mathbb{N} \setminus Q$.

But

$$\text{every } n \in \mathbb{N} \setminus Q \text{ satisfies } (f+g)^{\star n}(x) = 0 \tag{13.41.14}$$

[505].

[503] *Proof.* We know that the family $\left(f^{\star n}\right)_{n\in\mathbb{N}}$ is pointwise finitely supported. Hence, Proposition 1.7.7 (applied to $Q=\mathbb{N}$, $\left(f_q\right)_{q\in Q} = \left(f^{\star n}\right)_{n\in\mathbb{N}}$ and $\left(\lambda_q\right)_{q\in Q} = \left(\dfrac{1}{n!}\right)_{n\in\mathbb{N}}$) shows that the family $\left(\dfrac{1}{n!} f^{\star n}\right)_{n\in\mathbb{N}}$ is pointwise finitely supported. Since $\dfrac{1}{n!} f^{\star n} = \dfrac{f^{\star n}}{n!}$ for each $n \in \mathbb{N}$, this result rewrites as follows: The family $\left(\dfrac{f^{\star n}}{n!}\right)_{n\in\mathbb{N}}$ is pointwise finitely supported. Renaming the index n as q in this statement, we obtain the following: The family $\left(\dfrac{f^{\star q}}{q!}\right)_{q\in\mathbb{N}}$ is pointwise finitely supported. Qed.

[504] *Proof of (13.41.13):* Let $(q,r) \in \mathbb{N}\times\mathbb{N}$ be such that $(q,r) \notin K$. Then, $(q,r) \in (\mathbb{N}\times\mathbb{N}) \setminus K$ (since $(q,r) \in \mathbb{N}\times\mathbb{N}$ but $(q,r) \notin K$). Thus,

$$\left(\frac{f^{\star q}}{q!} \star \frac{g^{\star r}}{r!}\right)(x) = \frac{1}{q!} \cdot \frac{1}{r!} \cdot \underbrace{\left(f^{\star q} \star g^{\star r}\right)(x)}_{\substack{=0 \\ \text{(by (13.41.12))}}} = \frac{1}{q!} \cdot \frac{1}{r!} \cdot 0 = 0.$$

This proves (13.41.13).

[505] This has already been proven during our proof of Proposition 1.7.11(c).

Let $n \in \mathbb{N}$. Let us first observe that the map

$$\{0,1,\ldots,n\} \to \{(q,r) \in \mathbb{N} \times \mathbb{N} \mid q+r=n\}, \\ i \mapsto (i, n-i)$$

is a bijection. Hence, we can substitute $(i, n-i)$ for (q,r) in the sum $\sum_{\substack{(q,r)\in\mathbb{N}\times\mathbb{N};\\ q+r=n}} \frac{f^{\star q}}{q!} \star \frac{g^{\star r}}{r!}$. We thus obtain

$$\sum_{\substack{(q,r)\in\mathbb{N}\times\mathbb{N};\\ q+r=n}} \frac{f^{\star q}}{q!} \star \frac{g^{\star r}}{r!} = \sum_{i\in\{0,1,\ldots,n\}} \frac{f^{\star i}}{i!} \star \frac{g^{\star(n-i)}}{(n-i)!}. \tag{13.41.15}$$

But (13.41.9) yields

$$\begin{aligned}(f+g)^{\star n} &= \underbrace{\sum_{i=0}^{n}}_{=\sum_{i\in\{0,1,\ldots,n\}}} \underbrace{\binom{n}{i}}_{=\dfrac{n!}{i!\,(n-i)!}} f^{\star i} \star g^{\star(n-i)} \\ &= \sum_{i\in\{0,1,\ldots,n\}} \frac{n!}{i!\,(n-i)!} f^{\star i} \star g^{\star(n-i)} = n! \cdot \sum_{i\in\{0,1,\ldots,n\}} \frac{f^{\star i}}{i!} \star \frac{g^{\star(n-i)}}{(n-i)!}.\end{aligned}$$

Multiplying this equality by $\frac{1}{n!}$, we obtain

$$\frac{1}{n!}(f+g)^{\star n} = \sum_{i\in\{0,1,\ldots,n\}} \frac{f^{\star i}}{i!} \star \frac{g^{\star(n-i)}}{(n-i)!} = \sum_{\substack{(q,r)\in\mathbb{N}\times\mathbb{N};\\ q+r=n}} \frac{f^{\star q}}{q!} \star \frac{g^{\star r}}{r!} \tag{13.41.16}$$

(by (13.41.15)). Applying both sides of this equality to x, we find

$$\begin{aligned}\frac{1}{n!}(f+g)^{\star n}(x) &= \sum_{\substack{(q,r)\in\mathbb{N}\times\mathbb{N};\\ q+r=n}} \left(\frac{f^{\star q}}{q!} \star \frac{g^{\star r}}{r!}\right)(x) \\ &= \underbrace{\sum_{\substack{(q,r)\in\mathbb{N}\times\mathbb{N};\\ q+r=n;\\ (q,r)\in K}}}_{\substack{=\sum_{\substack{(q,r)\in K;\\ q+r=n}}\\ (\text{since } K\subset\mathbb{N}\times\mathbb{N})}} \left(\frac{f^{\star q}}{q!} \star \frac{g^{\star r}}{r!}\right)(x) + \sum_{\substack{(q,r)\in\mathbb{N}\times\mathbb{N};\\ q+r=n;\\ (q,r)\notin K}} \underbrace{\left(\frac{f^{\star q}}{q!} \star \frac{g^{\star r}}{r!}\right)(x)}_{\substack{=0\\ (\text{by } (13.41.13))}} \\ &= \sum_{\substack{(q,r)\in K;\\ q+r=n}} \left(\frac{f^{\star q}}{q!} \star \frac{g^{\star r}}{r!}\right)(x) + \underbrace{\sum_{\substack{(q,r)\in\mathbb{N}\times\mathbb{N};\\ q+r=n;\\ (q,r)\notin K}} 0}_{=0} = \sum_{\substack{(q,r)\in K;\\ q+r=n}} \left(\frac{f^{\star q}}{q!} \star \frac{g^{\star r}}{r!}\right)(x). \end{aligned} \tag{13.41.17}$$

Now, forget that we fixed n. We thus have proven (13.41.17) for each $n \in \mathbb{N}$.

From $\exp = \sum_{n\geq 0} \frac{1}{n!} T^n$, we obtain

$$\begin{aligned}\exp^{\star}(f+g) &= \underbrace{\sum_{n\geq 0}}_{=\sum_{n\in\mathbb{N}}} \frac{1}{n!}(f+g)^{\star n} \qquad (\text{by the definition of } \exp^{\star}(f+g)) \\ &= \sum_{n\in\mathbb{N}} \frac{1}{n!}(f+g)^{\star n}.\end{aligned}$$

Applying both sides of this equality to x, we find

$$\begin{aligned}
\left(\exp^{\star}(f+g)\right)(x) &= \left(\sum_{n\in\mathbb{N}} \frac{1}{n!}(f+g)^{\star n}\right)(x) = \sum_{n\in\mathbb{N}} \frac{1}{n!}(f+g)^{\star n}(x) \\
&= \sum_{n\in Q} \frac{1}{n!}(f+g)^{\star n}(x) + \sum_{n\in\mathbb{N}\setminus Q} \frac{1}{n!} \underbrace{(f+g)^{\star n}(x)}_{\substack{=0 \\ \text{(by (13.41.14))}}} \\
&\qquad \text{(since the set } \mathbb{N} \text{ is the union of its two disjoint subsets } Q \text{ and } \mathbb{N}\setminus Q\text{)} \\
&= \sum_{n\in Q} \frac{1}{n!}(f+g)^{\star n}(x) + \underbrace{\sum_{n\in\mathbb{N}\setminus Q} \frac{1}{n!}0}_{=0} = \sum_{n\in Q} \underbrace{\frac{1}{n!}(f+g)^{\star n}(x)}_{\substack{=\sum_{\substack{(q,r)\in K;\\ q+r=n}} \left(\frac{f^{\star q}}{q!}\star\frac{g^{\star r}}{r!}\right)(x) \\ \text{(by (13.41.17))}}} \\
&= \underbrace{\sum_{n\in Q} \sum_{\substack{(q,r)\in K;\\ q+r=n}}}_{\substack{=\sum_{(q,r)\in K} \sum_{\substack{n\in Q;\\ q+r=n}} \\ \text{(since both sums } \sum_{n\in Q} \text{ and } \sum_{\substack{(q,r)\in K;\\ q+r=n}} \\ \text{are finite)}}} \left(\frac{f^{\star q}}{q!}\star\frac{g^{\star r}}{r!}\right)(x) \\
&= \sum_{(q,r)\in K} \sum_{\substack{n\in Q;\\ q+r=n}} \left(\frac{f^{\star q}}{q!}\star\frac{g^{\star r}}{r!}\right)(x) . \qquad (13.41.18)
\end{aligned}$$

But each $(q,r)\in K$ satisfies

$$\sum_{\substack{n\in Q;\\ q+r=n}} \left(\frac{f^{\star q}}{q!}\star\frac{g^{\star r}}{r!}\right)(x) = \left(\frac{f^{\star q}}{q!}\star\frac{g^{\star r}}{r!}\right)(x) \qquad (13.41.19)$$

[506].

Hence, (13.41.18) becomes

$$\left(\exp^{\star}(f+g)\right)(x) = \sum_{(q,r)\in K} \underbrace{\sum_{\substack{n\in Q;\\ q+r=n}} \left(\frac{f^{\star q}}{q!}\star\frac{g^{\star r}}{r!}\right)(x)}_{\substack{=\left(\frac{f^{\star q}}{q!}\star\frac{g^{\star r}}{r!}\right)(x) \\ \text{(by (13.41.19))}}} = \sum_{(q,r)\in K} \left(\frac{f^{\star q}}{q!}\star\frac{g^{\star r}}{r!}\right)(x) . \qquad (13.41.20)$$

[506] *Proof of (13.41.19):* Let $(q,r)\in K$. Thus, $q+r\in\{u+v \mid (u,v)\in K\} = Q$ (since $Q=\{u+v \mid (u,v)\in K\}$).

Hence, there exists some $n\in Q$ satisfying $q+r=n$ (namely, $n=q+r$). Furthermore, this n is unique (because the condition $q+r=n$ clearly determines n uniquely). Thus, there exists a **unique** $n\in Q$ satisfying $q+r=n$. Therefore, the sum $\sum_{\substack{n\in Q;\\ q+r=n}} \left(\frac{f^{\star q}}{q!}\star\frac{g^{\star r}}{r!}\right)(x)$ has exactly one addend. Consequently, this sum simplifies as follows:

$$\sum_{\substack{n\in Q;\\ q+r=n}} \left(\frac{f^{\star q}}{q!}\star\frac{g^{\star r}}{r!}\right)(x) = \left(\frac{f^{\star q}}{q!}\star\frac{g^{\star r}}{r!}\right)(x) .$$

This proves (13.41.19).

From $\exp = \sum_{n\geq 0} \frac{1}{n!} T^n$, we obtain

$$\begin{aligned}
\exp^\star f &= \underbrace{\sum_{n\geq 0}}_{=\sum_{n\in\mathbb{N}}} \frac{1}{n!} f^{\star n} \qquad \left(\text{by the definition of } \exp^\star f\right) \\
&= \sum_{n\in\mathbb{N}} \frac{1}{n!} f^{\star n}.
\end{aligned}$$

The same argument (applied to g instead of f) shows that $\exp^\star g = \sum_{n\in\mathbb{N}} \frac{1}{n!} g^{\star n}$. Now,

$$\begin{aligned}
\underbrace{\left(\exp^\star f\right)}_{\substack{=\sum_{n\in\mathbb{N}} \frac{1}{n!} f^{\star n} \\ =\sum_{q\in\mathbb{N}} \frac{1}{q!} f^{\star q}}} \star \underbrace{\left(\exp^\star g\right)}_{\substack{=\sum_{n\in\mathbb{N}} \frac{1}{n!} g^{\star n} \\ =\sum_{r\in\mathbb{N}} \frac{1}{r!} g^{\star r}}} &= \left(\sum_{q\in\mathbb{N}} \underbrace{\frac{1}{q!} f^{\star q}}_{=\frac{f^{\star q}}{q!}}\right) \star \left(\sum_{r\in\mathbb{N}} \underbrace{\frac{1}{r!} g^{\star r}}_{=\frac{g^{\star r}}{r!}}\right) = \left(\sum_{q\in\mathbb{N}} \frac{f^{\star q}}{q!}\right) \star \left(\sum_{r\in\mathbb{N}} \frac{g^{\star r}}{r!}\right) \\
&= \sum_{(q,r)\in\mathbb{N}\times\mathbb{N}} \left(\frac{f^{\star q}}{q!} \star \frac{g^{\star r}}{r!}\right) \qquad \left(\text{by (13.41.11)}\right).
\end{aligned}$$

Applying both sides of this equality to x, we obtain

$$\begin{aligned}
\left(\left(\exp^\star f\right) \star \left(\exp^\star g\right)\right)(x) &= \left(\sum_{(q,r)\in\mathbb{N}\times\mathbb{N}} \left(\frac{f^{\star q}}{q!} \star \frac{g^{\star r}}{r!}\right)\right)(x) = \sum_{(q,r)\in\mathbb{N}\times\mathbb{N}} \left(\frac{f^{\star q}}{q!} \star \frac{g^{\star r}}{r!}\right)(x) \\
&= \underbrace{\sum_{\substack{(q,r)\in\mathbb{N}\times\mathbb{N}; \\ (q,r)\in K}}}_{\substack{=\sum_{(q,r)\in K} \\ (\text{since } K\subset\mathbb{N}\times\mathbb{N})}} \left(\frac{f^{\star q}}{q!} \star \frac{g^{\star r}}{r!}\right)(x) + \sum_{\substack{(q,r)\in\mathbb{N}\times\mathbb{N}; \\ (q,r)\notin K}} \underbrace{\left(\frac{f^{\star q}}{q!} \star \frac{g^{\star r}}{r!}\right)(x)}_{\substack{=0 \\ (\text{by (13.41.13)})}} \\
&= \sum_{(q,r)\in K} \left(\frac{f^{\star q}}{q!} \star \frac{g^{\star r}}{r!}\right)(x) + \underbrace{\sum_{\substack{(q,r)\in\mathbb{N}\times\mathbb{N}; \\ (q,r)\notin K}} 0}_{=0} = \sum_{(q,r)\in K} \left(\frac{f^{\star q}}{q!} \star \frac{g^{\star r}}{r!}\right)(x) \\
&= \left(\exp^\star (f+g)\right)(x) \qquad \left(\text{by (13.41.20)}\right).
\end{aligned}$$

Now, forget that we fixed x. We thus have shown that $\left(\left(\exp^\star f\right) \star \left(\exp^\star g\right)\right)(x) = \left(\exp^\star (f+g)\right)(x)$ for each $x \in C$. In other words, $\left(\exp^\star f\right) \star \left(\exp^\star g\right) = \exp^\star (f+g)$. In other words, $\exp^\star (f+g) = \left(\exp^\star f\right) \star \left(\exp^\star g\right)$. This completes the proof of Proposition 1.7.18(c).

(d) Consider the $\mathbf{k}$-linear map $0 : C \to A$. It satisfies $0 \in \mathfrak{n}(C, A)$ [507]. It remains to show that $\exp^\star 0 = u_A \epsilon_C$.

We have $\exp = \sum_{n\geq 0} \frac{1}{n!} T^n$. Thus, the definition of $\exp^\star 0$ yields

$$\begin{aligned}
\exp^\star 0 &= \sum_{n\geq 0} \frac{1}{n!} 0^{\star n} = \underbrace{\frac{1}{0!}}_{=\frac{1}{1}=1} 0^{\star 0} + \sum_{n\geq 1} \frac{1}{n!} \underbrace{0^{\star n}}_{\substack{=0 \\ (\text{since } n\geq 1)}} = 0^{\star 0} + \underbrace{\sum_{n\geq 1} \frac{1}{n!} 0}_{=0} \\
&= 0^{\star 0} = \left(\text{the unity of the } \mathbf{k}\text{-algebra } \left(\operatorname{Hom}(C, A), \star\right)\right) = u_A \epsilon_C.
\end{aligned}$$

Thus, Proposition 1.7.18(d) is proven.

[507] This can easily be checked by the reader (using the fact that $0^{\star n} = 0$ for every positive integer n).

(e) Let $f \in \mathfrak{n}(C, A)$. We must prove that each $n \in \mathbb{N}$ satisfies

$$nf \in \mathfrak{n}(C, A) \qquad \text{and} \qquad \exp^{\star}(nf) = (\exp^{\star} f)^{\star n}. \tag{13.41.21}$$

[*Proof of (13.41.21):* We shall prove (13.41.21) by induction over n:

Induction base: We have $0f \in \mathfrak{n}(C, A)$ [508]. Furthermore,

$$\exp^{\star}\left(\underbrace{0f}_{=0}\right) = \exp^{\star} 0 = u_A \epsilon_C \qquad \text{(by Proposition 1.7.18(d))}$$
$$= (\text{the unity of the } \mathbf{k}\text{-algebra } (\operatorname{Hom}(C, A), \star)) = (\exp^{\star} f)^{\star 0}.$$

Thus, we have shown that $0f \in \mathfrak{n}(C, A)$ and $\exp^{\star}(0f) = (\exp^{\star} f)^{\star 0}$. In other words, (13.41.21) holds for $n = 0$. This completes the induction base.

Induction step: Let $k \in \mathbb{N}$. Assume that (13.41.21) holds for $n = k$. We must prove that (13.41.21) holds for $n = k + 1$.

We have assumed that (13.41.21) holds for $n = k$. In other words,

$$kf \in \mathfrak{n}(C, A) \qquad \text{and} \qquad \exp^{\star}(kf) = (\exp^{\star} f)^{\star k}.$$

Now, $f \star (kf) = kf \star f = (kf) \star f$. Hence, Proposition 1.7.18(c) (applied to $g = kf$) yields that $f + kf \in \mathfrak{n}(C, A)$ and $\exp^{\star}(f + kf) = (\exp^{\star} f) \star (\exp^{\star}(kf))$. Now, $(k+1) f = f + kf \in \mathfrak{n}(C, A)$ and

$$\exp^{\star}\left(\underbrace{(k+1) f}_{=f+kf}\right) = \exp^{\star}(f + kf) = (\exp^{\star} f) \star \underbrace{(\exp^{\star}(kf))}_{=(\exp^{\star} f)^{\star k}} = (\exp^{\star} f) \star (\exp^{\star} f)^{\star k} = (\exp^{\star} f)^{\star(k+1)}.$$

Thus, we have shown that

$$(k+1) f \in \mathfrak{n}(C, A) \qquad \text{and} \qquad \exp^{\star}((k+1) f) = (\exp^{\star} f)^{\star(k+1)}.$$

In other words, (13.41.21) holds for $n = k+1$. This completes the induction step. Thus, the induction proof of (13.41.21) is finished.]

Now, (13.41.21) is proven. In other words, Proposition 1.7.18(e) is proven.

(f) Let $f \in \mathfrak{n}(C, A)$. Define a map $g \in \operatorname{Hom}(C, A)$ by $g = f + u_A \epsilon_C$. Thus, $g - u_A \epsilon_C = f \in \mathfrak{n}(C, A)$. Hence, $\log^{\star} g$ is well-defined.

For each $n \in \mathbb{N}$, define an element $\lambda_n \in \mathbf{k}$ by $\lambda_n = \begin{cases} \dfrac{(-1)^{n-1}}{n} 1_{\mathbf{k}}, & \text{if } n \geq 1; \\ 0, & \text{if } n = 0 \end{cases}$. Then, $\lambda_0 = 0$, whereas

$$\text{every integer } n \geq 1 \text{ satisfies } \lambda_n = \frac{(-1)^{n-1}}{n} 1_{\mathbf{k}}. \tag{13.41.22}$$

Hence,

$$\sum_{n \geq 0} \lambda_n T^n = \underbrace{\lambda_0}_{=0} T^0 + \sum_{n \geq 1} \underbrace{\lambda_n}_{\substack{= \frac{(-1)^{n-1}}{n} 1_{\mathbf{k}} \\ \text{(by (13.41.22))}}} T^n = \underbrace{0 T^0}_{=0} + \sum_{n \geq 1} \frac{(-1)^{n-1}}{n} \underbrace{1_{\mathbf{k}} T^n}_{=T^n}$$
$$= \sum_{n \geq 1} \frac{(-1)^{n-1}}{n} T^n = \log(1 + T) = \overline{\log}.$$

In other words, $\overline{\log} = \sum_{n \geq 0} \lambda_n T^n$. Hence, the definition of $\overline{\log}^{\star} f$ yields

$$\overline{\log}^{\star} f = \sum_{n \geq 0} \lambda_n f^{\star n} = \underbrace{\lambda_0}_{=0} f^{\star 0} + \sum_{n \geq 1} \underbrace{\lambda_n}_{\substack{= \frac{(-1)^{n-1}}{n} 1_{\mathbf{k}} \\ \text{(by (13.41.22))}}} f^{\star n} = \underbrace{0 f^{\star 0}}_{=0} + \sum_{n \geq 1} \frac{(-1)^{n-1}}{n} \underbrace{1_{\mathbf{k}} f^{\star n}}_{=f^{\star n}} = \sum_{n \geq 1} \frac{(-1)^{n-1}}{n} f^{\star n}.$$

[508] *Proof.* Proposition 1.7.11(d) (applied to $\lambda = 0 \cdot 1_{\mathbf{k}}$) yields $0 \cdot 1_{\mathbf{k}} f \in \mathfrak{n}(C, A)$. Thus, $0f = 0 = 0 \cdot 1_{\mathbf{k}} f \in \mathfrak{n}(C, A)$.

Now, the definition of $\log^\star g$ yields

$$\log^\star g = \overline{\log}^\star f = \sum_{n\geq 1} \frac{(-1)^{n-1}}{n} f^{\star n}.$$

Since $g = f + u_A\epsilon_C$, this rewrites as $\log^\star (f + u_A\epsilon_C) = \sum_{n\geq 1} \dfrac{(-1)^{n-1}}{n} f^{\star n}$. This proves Proposition 1.7.18(f). □

Thus, Proposition 1.7.15, Lemma 1.7.16 and Proposition 1.7.18 are proven. Hence, Exercise 1.7.20 is solved.

13.42. **Solution to Exercise 1.7.28.** *Solution to Exercise 1.7.28.*

Proof of Proposition 1.7.21. We know that $\gamma : C \to C'$ is a $\mathbf{k}$-coalgebra morphism. In other words, $\gamma : C \to C'$ is a $\mathbf{k}$-linear map satisfying $\Delta_{C'} \circ \gamma = (\gamma \otimes \gamma) \circ \Delta_C$ and $\epsilon_{C'} \circ \gamma = \epsilon_C$.

We know that $\alpha : A \to A'$ is a $\mathbf{k}$-algebra morphism. In other words, $\alpha : A \to A'$ is a $\mathbf{k}$-linear map satisfying $\alpha \circ m_A = m_{A'} \circ (\alpha \otimes \alpha)$ and $\alpha \circ u_A = u_{A'}$.

(a) Let $f \in \operatorname{Hom}(C, A)$, $g \in \operatorname{Hom}(C, A)$, $f' \in \operatorname{Hom}(C', A')$ and $g' \in \operatorname{Hom}(C', A')$ be such that $f' \circ \gamma = \alpha \circ f$ and $g' \circ \gamma = \alpha \circ g$. We must prove that $(f' \star g') \circ \gamma = \alpha \circ (f \star g)$.

The definition of convolution yields $f \star g = m_A \circ (f \otimes g) \circ \Delta_C$ and $f' \star g' = m_{A'} \circ (f' \otimes g') \circ \Delta_{C'}$. Now,

$$\begin{aligned}
\underbrace{(f' \star g')}_{=m_{A'}\circ(f'\otimes g')\circ\Delta_{C'}} \circ\gamma &= m_{A'} \circ (f' \otimes g') \circ \underbrace{\Delta_{C'} \circ \gamma}_{=(\gamma\otimes\gamma)\circ\Delta_C} = m_{A'} \circ \underbrace{(f' \otimes g') \circ (\gamma \otimes \gamma)}_{=(f'\circ\gamma)\otimes(g'\circ\gamma)} \circ\Delta_C \\
&= m_{A'} \circ \left(\underbrace{(f' \circ \gamma)}_{=\alpha\circ f} \otimes \underbrace{(g' \circ \gamma)}_{=\alpha\circ g} \right) \circ \Delta_C = m_{A'} \circ \underbrace{((\alpha \circ f) \otimes (\alpha \circ g))}_{=(\alpha\otimes\alpha)\circ(f\otimes g)} \circ\Delta_C \\
&= \underbrace{m_{A'} \circ (\alpha \otimes \alpha)}_{=\alpha\circ m_A} \circ (f \otimes g) \circ \Delta_C = \alpha \circ \underbrace{m_A \circ (f \otimes g) \circ \Delta_C}_{=f\star g} = \alpha \circ (f \star g).
\end{aligned}$$

This proves Proposition 1.7.21(a).

(b) Let $f \in \operatorname{Hom}(C, A)$ and $f' \in \operatorname{Hom}(C', A')$ be such that $f' \circ \gamma = \alpha \circ f$. We must prove that every $n \in \mathbb{N}$ satisfies

$$(f')^{\star n} \circ \gamma = \alpha \circ f^{\star n}. \qquad (13.42.1)$$

[*Proof of (13.42.1):* We shall prove (13.42.1) by induction over n:

Induction base: We have $f^{\star 0} =$ (the unity of the $\mathbf{k}$-algebra $(\operatorname{Hom}(C, A), \star)$) $= u_A \circ \epsilon_C$ and similarly $(f')^{\star 0} = u_{A'} \circ \epsilon_{C'}$. Hence,

$$\underbrace{(f')^{\star 0}}_{=u_{A'}\circ\epsilon_{C'}} \circ\gamma = \underbrace{u_{A'}}_{=\alpha\circ u_A} \circ \underbrace{\epsilon_{C'} \circ \gamma}_{=\epsilon_C} = \alpha \circ \underbrace{u_A \circ \epsilon_C}_{=f^{\star 0}} = \alpha \circ f^{\star 0}.$$

In other words, (13.42.1) holds for $n = 0$. This completes the induction base.

Induction step: Fix $k \in \mathbb{N}$. Assume that (13.42.1) holds for $n = k$. We must now prove that (13.42.1) holds for $n = k + 1$.

We have assumed that (13.42.1) holds for $n = k$. In other words, we have $(f')^{\star k} \circ \gamma = \alpha \circ f^{\star k}$.

Now,

$$\begin{aligned}
\underbrace{(f')^{\star(k+1)}}_{=f'\star(f')^{\star k}} \circ\gamma &= \left(f' \star (f')^{\star k} \right) \circ \gamma = \alpha \circ \underbrace{\left(f \star f^{\star k} \right)}_{=f^{\star(k+1)}} \\
&\qquad \left(\text{by Proposition 1.7.21(a) (applied to } g = f^{\star k} \text{ and } g' = (f')^{\star k} \text{)} \right) \\
&= \alpha \circ f^{\star(k+1)}.
\end{aligned}$$

In other words, (13.42.1) holds for $n = k+1$. This completes the induction step. Thus, the induction proof of (13.42.1) is finished.]

Hence, (13.42.1) is proven. In other words, we have proven Proposition 1.7.21(b). □

Before we move on to the proof of Proposition 1.7.22, let us show a simple lemma (which can also easily be obtained as a consequence of Exercise 1.4.4(b)):

Lemma 13.42.1. *Let C be a $\mathbf{k}$-coalgebra. Let A be a $\mathbf{k}$-algebra. Let $F \in \operatorname{Hom}(C, A)$ and $G \in \operatorname{Hom}(C, A)$. Then, each $k \in \mathbb{N}$ satisfies*

$$(F \otimes G)^{\star k} = F^{\star k} \otimes G^{\star k} \tag{13.42.2}$$

(as maps $C \otimes C \to A \otimes A$).

Proof of Lemma 13.42.1. We must prove that (13.42.2) holds for each $k \in \mathbb{N}$. We shall prove this by induction over k:

Induction base: Let s denote the canonical $\mathbf{k}$-algebra isomorphism $\mathbf{k} \to \mathbf{k} \otimes \mathbf{k}$. We have

$$\begin{aligned}(F \otimes G)^{\star 0} &= (\text{the unity of the } \mathbf{k}\text{-algebra } (\operatorname{Hom}(C \otimes C, A \otimes A), \star)) \\ &= \underbrace{u_{A \otimes A}}_{\substack{=(u_A \otimes u_A) \circ s \\ \text{(by the definition of} \\ \text{the } \mathbf{k}\text{-algebra } A \otimes A)}} \circ \underbrace{\epsilon_{C \otimes C}}_{\substack{=s^{-1} \circ (\epsilon_C \otimes \epsilon_C) \\ \text{(by the definition of} \\ \text{the } \mathbf{k}\text{-coalgebra } C \otimes C)}} = (u_A \otimes u_A) \circ \underbrace{s \circ s^{-1}}_{=\operatorname{id}_{\mathbf{k} \otimes \mathbf{k}}} \circ (\epsilon_C \otimes \epsilon_C) = (u_A \otimes u_A) \circ (\epsilon_C \otimes \epsilon_C) \\ &= \underbrace{(u_A \circ \epsilon_C)}_{\substack{=(\text{the unity of the } \mathbf{k}\text{-algebra } (\operatorname{Hom}(C,A),\star)) \\ =F^{\star 0}}} \otimes \underbrace{(u_A \circ \epsilon_C)}_{\substack{=(\text{the unity of the } \mathbf{k}\text{-algebra } (\operatorname{Hom}(C,A),\star)) \\ =G^{\star 0}}} = F^{\star 0} \otimes G^{\star 0}.\end{aligned}$$

In other words, (13.42.2) holds for $k = 0$. This completes the induction base.

Induction step: Let $\ell \in \mathbb{N}$. Assume that (13.42.2) holds for $k = \ell$. We must prove that (13.42.2) holds for $k = \ell + 1$.

We have assumed that (13.42.2) holds for $k = \ell$. In other words, we have

$$(F \otimes G)^{\star \ell} = F^{\star \ell} \otimes G^{\star \ell}. \tag{13.42.3}$$

Now, Exercise 1.4.4(a) (applied to C, A, $F^{\star \ell}$, F, $G^{\star \ell}$ and G instead of D, B, f, f', g and g') shows that

$$\left(F^{\star \ell} \otimes G^{\star \ell}\right) \star (F \otimes G) = \left(F^{\star \ell} \star F\right) \otimes \left(G^{\star \ell} \star G\right)$$

in the convolution algebra $\operatorname{Hom}(C \otimes C, A \otimes A)$. Thus,

$$\left(F^{\star \ell} \otimes G^{\star \ell}\right) \star (F \otimes G) = \underbrace{\left(F^{\star \ell} \star F\right)}_{=F^{\star(\ell+1)}} \otimes \underbrace{\left(G^{\star \ell} \star G\right)}_{=G^{\star(\ell+1)}} = F^{\star(\ell+1)} \otimes G^{\star(\ell+1)}.$$

Hence,

$$F^{\star(\ell+1)} \otimes G^{\star(\ell+1)} = \underbrace{\left(F^{\star \ell} \otimes G^{\star \ell}\right)}_{\substack{=(F \otimes G)^{\star \ell} \\ \text{(by (13.42.3))}}} \star (F \otimes G) = (F \otimes G)^{\star \ell} \star (F \otimes G) = (F \otimes G)^{\star(\ell+1)}.$$

In other words, (13.42.2) holds for $k = \ell + 1$. This completes the induction step. Hence, (13.42.2) is proven by induction. Thus, Lemma 13.42.1 is proven. □

Proof of Proposition 1.7.22. Let $\mathfrak{i}$ be the unity of the $\mathbf{k}$-algebra $(\operatorname{Hom}(C, A), \star)$. Thus,

$$\mathfrak{i} = (\text{the unity of the } \mathbf{k}\text{-algebra } (\operatorname{Hom}(C, A), \star)) = u_A \circ \epsilon_C.$$

The axioms of a $\mathbf{k}$-bialgebra show that $\epsilon(1_C) = 1_{\mathbf{k}}$ (since C is a $\mathbf{k}$-bialgebra). In other words, $\epsilon(1) = 1$.

Let us first show that every $x \in C$ and $y \in C$ satisfy

$$f(xy) = \epsilon(y) f(x) + \epsilon(x) f(y). \tag{13.42.4}$$

[*Proof of (13.42.4):* Let $x \in C$ and $y \in C$. The map ϵ is $\mathbf{k}$-linear; thus,

$$\epsilon(x - \epsilon(x) 1) = \epsilon(x) - \epsilon(x) \underbrace{\epsilon(1)}_{=1} = \epsilon(x) - \epsilon(x) = 0.$$

In other words, $x - \epsilon(x) 1 \in \ker \epsilon$. The same argument (applied to y instead of x) shows that $y - \epsilon(y) 1 \in \ker \epsilon$. But

$$(x - \epsilon(x) 1)(y - \epsilon(y) 1) = xy - \epsilon(x) y - \epsilon(y) x + \epsilon(x) \epsilon(y) 1,$$

so that

$$xy - \epsilon(x) y - \epsilon(y) x + \epsilon(x) \epsilon(y) 1 = \underbrace{(x - \epsilon(x) 1)}_{\in \ker \epsilon} \underbrace{(y - \epsilon(y) 1)}_{\in \ker \epsilon} \in (\ker \epsilon)(\ker \epsilon) = (\ker \epsilon)^2 .$$

Applying the map f to both sides of this relation, we find

$$f(xy - \epsilon(x) y - \epsilon(y) x + \epsilon(x) \epsilon(y) 1) \in f\left((\ker \epsilon)^2\right) = 0,$$

so that $f(xy - \epsilon(x) y - \epsilon(y) x + \epsilon(x) \epsilon(y) 1) = 0$. Hence,

$$\begin{aligned} 0 &= f(xy - \epsilon(x) y - \epsilon(y) x + \epsilon(x) \epsilon(y) 1) \\ &= f(xy) - \epsilon(x) f(y) - \epsilon(y) f(x) + \epsilon(x) \epsilon(y) \underbrace{f(1)}_{=0} \qquad \text{(since the map } f \text{ is } \mathbf{k}\text{-linear)} \\ &= f(xy) - \epsilon(x) f(y) - \epsilon(y) f(x) . \end{aligned}$$

Solving this equality for $f(xy)$, we obtain $f(xy) = \epsilon(y) f(x) + \epsilon(x) f(y)$. This proves (13.42.4).]

Now, we shall show that

$$(f \circ m_C)(z) = (m_A \circ (f \otimes \mathfrak{i} + \mathfrak{i} \otimes f))(z) \tag{13.42.5}$$

for each $z \in C \otimes C$.

[*Proof of (13.42.5):* Let $z \in C \otimes C$. We are going to prove the equality (13.42.5).

Both sides of the equality (13.42.5) are $\mathbf{k}$-linear in z. Hence, for the proof of (13.42.5), we can WLOG assume that z is a pure tensor (since any $z \in C \otimes C$ is a $\mathbf{k}$-linear combination of pure tensors). Assume this.

There exist $x \in C$ and $y \in C$ satisfying $z = x \otimes y$ (since z is a pure tensor). Consider these x and y. We have

$$\underbrace{\mathfrak{i}}_{= u_A \circ \epsilon_C}(x) = \left(u_A \circ \underbrace{\epsilon_C}_{= \epsilon}\right)(x) = (u_A \circ \epsilon)(x) = u_A(\epsilon(x)) = \epsilon(x) 1_A$$

(by the definition of the map u_A). The same argument (applied to y instead of x) shows that $\mathfrak{i}(y) = \epsilon(y) 1_A$.

From $z = x \otimes y$, we obtain $m_C(z) = m_C(x \otimes y) = xy$ (by the definition of m_C). Now,

$$(f \circ m_C)(z) = f\left(\underbrace{m_C(z)}_{= xy}\right) = f(xy) = \epsilon(y) f(x) + \epsilon(x) f(y) \qquad \text{(by (13.42.4))} .$$

Comparing this with

$$\begin{aligned} (m_A \circ (f \otimes \mathfrak{i} + \mathfrak{i} \otimes f))(z) &= m_A\left((f \otimes \mathfrak{i} + \mathfrak{i} \otimes f)\left(\underbrace{z}_{= x \otimes y}\right)\right) = m_A\left(\underbrace{(f \otimes \mathfrak{i} + \mathfrak{i} \otimes f)(x \otimes y)}_{= (f \otimes \mathfrak{i})(x \otimes y) + (\mathfrak{i} \otimes f)(x \otimes y)}\right) \\ &= m_A\left(\underbrace{(f \otimes \mathfrak{i})(x \otimes y)}_{= f(x) \otimes \mathfrak{i}(y)} + \underbrace{(\mathfrak{i} \otimes f)(x \otimes y)}_{= \mathfrak{i}(x) \otimes f(y)}\right) = m_A(f(x) \otimes \mathfrak{i}(y) + \mathfrak{i}(x) \otimes f(y)) \\ &= f(x) \underbrace{\mathfrak{i}(y)}_{= \epsilon(y) 1_A} + \underbrace{\mathfrak{i}(x)}_{= \epsilon(x) 1_A} f(y) \qquad \text{(by the definition of } m_A\text{)} \\ &= \epsilon(y) f(x) + \epsilon(x) f(y) , \end{aligned}$$

we obtain $(f \circ m_C)(z) = (m_A \circ (f \otimes \mathfrak{i} + \mathfrak{i} \otimes f))(z)$. Thus, (13.42.5) is proven.]

Now, we have proven that (13.42.5) holds for each $z \in C$. In other words, we have

$$f \circ m_C = m_A \circ (f \otimes \mathfrak{i} + \mathfrak{i} \otimes f) . \tag{13.42.6}$$

The axioms of a $\mathbf{k}$-bialgebra show that the comultiplication $\Delta_C : C \to C \otimes C$ of the $\mathbf{k}$-coalgebra C is a $\mathbf{k}$-algebra homomorphism (since C is a $\mathbf{k}$-bialgebra). They also show that the multiplication $m_C : C \otimes C \to C$ of the $\mathbf{k}$-algebra C is a $\mathbf{k}$-coalgebra homomorphism (since C is a $\mathbf{k}$-bialgebra).

Exercise 1.5.6(b) shows that the $\mathbf{k}$-algebra A is commutative if and only if its multiplication $m_A : A \otimes A \to A$ is a $\mathbf{k}$-algebra homomorphism. Thus, its multiplication $m_A : A \otimes A \to A$ is a $\mathbf{k}$-algebra homomorphism (since A is commutative).

Thus, Proposition 1.7.21(b) (applied to $C \otimes C$, C, $A \otimes A$, A, m_C, m_A, $f \otimes \mathfrak{i} + \mathfrak{i} \otimes f$ and f instead of C, C', A, A', γ, α, f and f') shows that each $n \in \mathbb{N}$ satisfies

$$f^{\star n} \circ m_C = m_A \circ (f \otimes \mathfrak{i} + \mathfrak{i} \otimes f)^{\star n} \tag{13.42.7}$$

(because of (13.42.6)).

The two $\mathbf{k}$-linear maps $f \otimes \mathfrak{i}$ and $\mathfrak{i} \otimes f$ in $\operatorname{Hom}(C \otimes C, A \otimes A)$ satisfy

$$(f \otimes \mathfrak{i}) \star (\mathfrak{i} \otimes f) = (\mathfrak{i} \otimes f) \star (f \otimes \mathfrak{i}) \tag{13.42.8}$$

[509]. In other words, the two elements $f \otimes \mathfrak{i}$ and $\mathfrak{i} \otimes f$ of the $\mathbf{k}$-algebra $(\operatorname{Hom}(C \otimes C, A \otimes A), \star)$ commute.

Hence, it is easy to see

$$(f \otimes \mathfrak{i} + \mathfrak{i} \otimes f)^{\star n} = \sum_{i=0}^{n} \binom{n}{i} f^{\star i} \otimes f^{\star (n-i)} \qquad \text{for each } n \in \mathbb{N}. \tag{13.42.10}$$

[*Proof of (13.42.10):* Let $n \in \mathbb{N}$. Let $\mathfrak{G}$ be the $\mathbf{k}$-subalgebra of $(\operatorname{Hom}(C \otimes C, A \otimes A), \star)$ generated by the two elements $f \otimes \mathfrak{i}$ and $\mathfrak{i} \otimes f$. Thus, the $\mathbf{k}$-algebra $\mathfrak{G}$ is generated by commuting elements (because the elements $f \otimes \mathfrak{i}$ and $\mathfrak{i} \otimes f$ of the $\mathbf{k}$-algebra $(\operatorname{Hom}(C \otimes C, A \otimes A), \star)$ commute), and therefore is commutative (since any $\mathbf{k}$-algebra generated by commuting elements must be commutative). Hence, the binomial formula holds in this $\mathbf{k}$-algebra $\mathfrak{G}$. In other words, any $\alpha \in \mathfrak{G}$ and $\beta \in \mathfrak{G}$ and $m \in \mathbb{N}$ satisfy

$$(\alpha + \beta)^{\star m} = \sum_{i=0}^{m} \binom{m}{i} \alpha^{\star i} \star \beta^{\star (m-i)}$$

[509] *Proof of (13.42.8):* Exercise 1.4.4(a) (applied to C, A, f, $\mathfrak{i}$, $\mathfrak{i}$ and f instead of D, B, f, f', g and g') shows that

$$(f \otimes \mathfrak{i}) \star (\mathfrak{i} \otimes f) = \underbrace{(f \star \mathfrak{i})}_{=f} \otimes \underbrace{(\mathfrak{i} \star f)}_{=f} = f \otimes f. \tag{13.42.9}$$

But Exercise 1.4.4(a) (applied to C, A, $\mathfrak{i}$, f, f and $\mathfrak{i}$ instead of D, B, f, f', g and g') shows that

$$(\mathfrak{i} \otimes f) \star (f \otimes \mathfrak{i}) = \underbrace{(\mathfrak{i} \star f)}_{=f} \otimes \underbrace{(f \star \mathfrak{i})}_{=f} = f \otimes f.$$

Comparing this with (13.42.9), we obtain $(f \otimes \mathfrak{i}) \star (\mathfrak{i} \otimes f) = (\mathfrak{i} \otimes f) \star (f \otimes \mathfrak{i})$. This proves (13.42.8).

(since the multiplication in the $\mathbf{k}$-algebra $\mathfrak{G}$ is the convolution $\star$). Applying this to $\alpha = f \otimes \mathfrak{i}$, $\beta = \mathfrak{i} \otimes f$ and $m = n$, we obtain

$$\begin{aligned}
&\left(f \otimes \mathfrak{i} + \mathfrak{i} \otimes f\right)^{\star n} \\
&= \sum_{i=0}^{n} \binom{n}{i} \underbrace{\left(f \otimes \mathfrak{i}\right)^{\star i}}_{\substack{= f^{\star i} \otimes \mathfrak{i}^{\star i} \\ \text{(by (13.42.2) (applied} \\ \text{to } f,\ \mathfrak{i} \text{ and } i \text{ instead of } F,\ G \text{ and } k))}} \star \underbrace{\left(\mathfrak{i} \otimes f\right)^{\star (n-i)}}_{\substack{= \mathfrak{i}^{\star (n-i)} \otimes f^{\star (n-i)} \\ \text{(by (13.42.2) (applied} \\ \text{to } \mathfrak{i},\ f \text{ and } n-i \text{ instead of } F,\ G \text{ and } k))}} \\
&= \sum_{i=0}^{n} \binom{n}{i} \left(f^{\star i} \otimes \underbrace{\mathfrak{i}^{\star i}}_{\substack{= \mathfrak{i} \\ \text{(since } \mathfrak{i} \text{ is the unity} \\ \text{of the } \mathbf{k}\text{-algebra } (\operatorname{Hom}(C,A),\star))}} \right) \star \left(\underbrace{\mathfrak{i}^{\star (n-i)}}_{\substack{= \mathfrak{i} \\ \text{(since } \mathfrak{i} \text{ is the unity} \\ \text{of the } \mathbf{k}\text{-algebra } (\operatorname{Hom}(C,A),\star))}} \otimes f^{\star (n-i)} \right) \\
&= \sum_{i=0}^{n} \binom{n}{i} \underbrace{\left(f^{\star i} \otimes \mathfrak{i}\right) \star \left(\mathfrak{i} \otimes f^{\star (n-i)}\right)}_{\substack{= \left(f^{\star i} \star \mathfrak{i}\right) \otimes \left(\mathfrak{i} \star f^{\star (n-i)}\right) \\ \text{(by Exercise 1.4.4(a) (applied} \\ \text{to } C,\ A,\ f^{\star i},\ \mathfrak{i},\ \mathfrak{i},\ f^{\star (n-i)} \\ \text{instead of } D,\ B,\ f,\ f',\ g \text{ and } g'))}} = \sum_{i=0}^{n} \binom{n}{i} \underbrace{\left(f^{\star i} \star \mathfrak{i}\right)}_{\substack{= f^{\star i} \\ \text{(since } \mathfrak{i} \text{ is the unity} \\ \text{of the } \mathbf{k}\text{-algebra } (\operatorname{Hom}(C,A),\star))}} \otimes \underbrace{\left(\mathfrak{i} \star f^{\star (n-i)}\right)}_{\substack{= f^{\star (n-i)} \\ \text{(since } \mathfrak{i} \text{ is the unity} \\ \text{of the } \mathbf{k}\text{-algebra } (\operatorname{Hom}(C,A),\star))}} \\
&= \sum_{i=0}^{n} \binom{n}{i} f^{\star i} \otimes f^{\star (n-i)}.
\end{aligned}$$

This proves (13.42.10).]

Now, let $x, y \in C$ and $n \in \mathbb{N}$ be arbitrary. Then, the definition of m_C yields $m_C(x \otimes y) = xy$. Hence,

$$\left(f^{\star n} \circ m_C\right)(x \otimes y) = f^{\star n} \left(\underbrace{m_C(x \otimes y)}_{= xy} \right) = f^{\star n}(xy).$$

Hence,

$$\begin{aligned}
f^{\star n}(xy) &= \underbrace{\left(f^{\star n} \circ m_C\right)}_{\substack{= m_A \circ (f \otimes \mathfrak{i} + \mathfrak{i} \otimes f)^{\star n} \\ \text{(by (13.42.7))}}} (x \otimes y) = \left(m_A \circ \left(f \otimes \mathfrak{i} + \mathfrak{i} \otimes f\right)^{\star n}\right)(x \otimes y) \\
&= m_A \left(\underbrace{\left(f \otimes \mathfrak{i} + \mathfrak{i} \otimes f\right)^{\star n}}_{\substack{= \sum_{i=0}^{n} \binom{n}{i} f^{\star i} \otimes f^{\star (n-i)} \\ \text{(by (13.42.10))}}} (x \otimes y) \right) = m_A \left(\left(\sum_{i=0}^{n} \binom{n}{i} f^{\star i} \otimes f^{\star (n-i)} \right) (x \otimes y) \right) \\
&= m_A \left(\underbrace{\left(\sum_{i=0}^{n} \binom{n}{i} f^{\star i} \otimes f^{\star (n-i)} \right) (x \otimes y)}_{= \sum_{i=0}^{n} \binom{n}{i} \left(f^{\star i} \otimes f^{\star (n-i)}\right)(x \otimes y)} \right) = m_A \left(\sum_{i=0}^{n} \binom{n}{i} \underbrace{\left(f^{\star i} \otimes f^{\star (n-i)}\right)(x \otimes y)}_{= f^{\star i}(x) \otimes f^{\star (n-i)}(y)} \right) \\
&= m_A \left(\sum_{i=0}^{n} \binom{n}{i} f^{\star i}(x) \otimes f^{\star (n-i)}(y) \right) = \sum_{i=0}^{n} \binom{n}{i} f^{\star i}(x) f^{\star (n-i)}(y)
\end{aligned}$$

(by the definition of m_A). This proves Proposition 1.7.22. □

Before we prove Proposition 1.7.23, let us show two lemmas:[510]

Lemma 13.42.2. *Let C be a $\mathbf{k}$-coalgebra. Let A be a $\mathbf{k}$-algebra. Let $f \in \mathfrak{n}(C, A)$.*

(a) *If $z \in C$ is arbitrary, then the family $\left(\frac{1}{n!} f^{\star n}(z)\right)_{n \in \mathbb{N}} \in A^{\mathbb{N}}$ is finitely supported and satisfies*

$$(\exp^\star f)(z) = \sum_{n \in \mathbb{N}} \frac{1}{n!} f^{\star n}(z).$$

(b) *If $x \in C$ and $y \in C$ are arbitrary, then the family $\left(\sum_{i=0}^{n} \frac{1}{i! \cdot (n-i)!} f^{\star i}(x) f^{\star(n-i)}(y)\right)_{n \in \mathbb{N}} \in A^{\mathbb{N}}$ is finitely supported and satisfies*

$$(\exp^\star f)(x) \cdot (\exp^\star f)(y) = \sum_{n \in \mathbb{N}} \sum_{i=0}^{n} \frac{1}{i! \cdot (n-i)!} f^{\star i}(x) f^{\star(n-i)}(y).$$

Proof of Lemma 13.42.2. We have $f \in \mathfrak{n}(C, A)$. In other words, f is a pointwise $\star$-nilpotent map in $\operatorname{Hom}(C, A)$. Thus, the family $(f^{\star n})_{n \in \mathbb{N}} \in (\operatorname{Hom}(C, A))^{\mathbb{N}}$ is pointwise finitely supported. In other words, for each $x \in C$,

$$\text{the family } (f^{\star n}(x))_{n \in \mathbb{N}} \in A^{\mathbb{N}} \text{ is finitely supported.} \tag{13.42.11}$$

(a) Let $z \in C$ be arbitrary. Recall that $\exp = \sum_{n \geq 0} \frac{1}{n!} T^n$. Hence, $\exp^\star f = \sum_{n \geq 0} \frac{1}{n!} f^{\star n}$ (by the definition of $\exp^\star f$). Hence,

$$(\exp^\star f)(z) = \left(\sum_{n \geq 0} \frac{1}{n!} f^{\star n}\right)(z) = \underbrace{\sum_{n \geq 0}}_{= \sum_{n \in \mathbb{N}}} \frac{1}{n!} f^{\star n}(z) = \sum_{n \in \mathbb{N}} \frac{1}{n!} f^{\star n}(z).$$

In particular, the sum $\sum_{n \in \mathbb{N}} \frac{1}{n!} f^{\star n}(z)$ is well-defined. Hence, the family $\left(\frac{1}{n!} f^{\star n}(z)\right)_{n \in \mathbb{N}} \in A^{\mathbb{N}}$ is finitely supported. Thus, the proof of Lemma 13.42.2(a) is complete.

(b) Let $x \in C$ and $y \in C$ be arbitrary.

Lemma 13.42.2(a) (applied to $z = x$) shows that the family $\left(\frac{1}{n!} f^{\star n}(x)\right)_{n \in \mathbb{N}} \in A^{\mathbb{N}}$ is finitely supported and satisfies $(\exp^\star f)(x) = \sum_{n \in \mathbb{N}} \frac{1}{n!} f^{\star n}(x)$. Renaming the index n as q in this sentence, we obtain the following: The family $\left(\frac{1}{q!} f^{\star q}(x)\right)_{q \in \mathbb{N}} \in A^{\mathbb{N}}$ is finitely supported.and satisfies

$$(\exp^\star f)(x) = \sum_{q \in \mathbb{N}} \frac{1}{q!} f^{\star q}(x). \tag{13.42.12}$$

Similarly, the family $\left(\frac{1}{r!} f^{\star r}(y)\right)_{r \in \mathbb{N}} \in A^{\mathbb{N}}$ is finitely supported and satisfies

$$(\exp^\star f)(y) = \sum_{r \in \mathbb{N}} \frac{1}{r!} f^{\star r}(y). \tag{13.42.13}$$

Recall that sums of the form $\sum_{q \in Q} f_q$ (where $(f_q)_{q \in Q}$ is a finitely supported family) satisfy the usual rules for finite sums, even though their indexing set Q may be infinite. This pertains, in particular, to the sums $\sum_{q \in \mathbb{N}} \frac{1}{q!} f^{\star q}(x)$ and $\sum_{r \in \mathbb{N}} \frac{1}{r!} f^{\star r}(y)$ (because the families $\left(\frac{1}{q!} f^{\star q}(x)\right)_{q \in \mathbb{N}}$ and $\left(\frac{1}{r!} f^{\star r}(y)\right)_{r \in \mathbb{N}}$ are finitely

[510] Recall that we are still using the conventions that are in place throughout Section 1.7.

supported). Thus, the following manipulations make sense:

$$\begin{aligned}
&\left(\sum_{q\in\mathbb{N}} \frac{1}{q!} f^{\star q}(x)\right)\left(\sum_{r\in\mathbb{N}} \frac{1}{r!} f^{\star r}(y)\right) \\
&= \sum_{q\in\mathbb{N}} \sum_{r\in\mathbb{N}} \frac{1}{q!} f^{\star q}(x) \frac{1}{r!} f^{\star r}(y) = \sum_{q\in\mathbb{N}} \underbrace{\sum_{r\in\mathbb{N}} \frac{1}{q!\cdot r!} f^{\star q}(x) f^{\star r}(y)}_{\substack{=\sum_{\substack{n\in\mathbb{N};\\ n\geq q}} \frac{1}{q!\cdot(n-q)!} f^{\star q}(x) f^{\star(n-q)}(y) \\ \text{(here, we have substituted } n-q \text{ for } r \text{ in the sum)}}} \\
&= \underbrace{\sum_{q\in\mathbb{N}} \sum_{\substack{n\in\mathbb{N};\\ n\geq q}}}_{=\sum_{n\in\mathbb{N}} \sum_{\substack{q\in\mathbb{N};\\ n\geq q}}} \frac{1}{q!\cdot(n-q)!} f^{\star q}(x) f^{\star(n-q)}(y) = \sum_{n\in\mathbb{N}} \underbrace{\sum_{\substack{q\in\mathbb{N};\\ n\geq q}}}_{=\sum_{\substack{q\in\mathbb{N};\\ q\leq n}} = \sum_{q=0}^{n}} \frac{1}{q!\cdot(n-q)!} f^{\star q}(x) f^{\star(n-q)}(y) \\
&= \sum_{n\in\mathbb{N}} \sum_{q=0}^{n} \frac{1}{q!\cdot(n-q)!} f^{\star q}(x) f^{\star(n-q)}(y) = \sum_{n\in\mathbb{N}} \sum_{i=0}^{n} \frac{1}{i!\cdot(n-i)!} f^{\star i}(x) f^{\star(n-i)}(y)
\end{aligned}$$

(here, we have renamed the summation index q as i). Comparing this with

$$\underbrace{\left(\sum_{q\in\mathbb{N}} \frac{1}{q!} f^{\star q}(x)\right)}_{\substack{=(\exp^\star f)(x) \\ \text{(by (13.42.12))}}} \underbrace{\left(\sum_{r\in\mathbb{N}} \frac{1}{r!} f^{\star r}(y)\right)}_{\substack{=(\exp^\star f)(y) \\ \text{(by (13.42.13))}}} = (\exp^\star f)(x) \cdot (\exp^\star f)(y),$$

we obtain

$$(\exp^\star f)(x) \cdot (\exp^\star f)(y) = \sum_{n\in\mathbb{N}} \sum_{i=0}^{n} \frac{1}{i!\cdot(n-i)!} f^{\star i}(x) f^{\star(n-i)}(y).$$

In particular, the sum on the right hand side of this equality is well-defined. Thus, the family $\left(\sum_{i=0}^{n} \dfrac{1}{i!\cdot(n-i)!} f^{\star i}(x) f^{\star(n-i)}(y)\right)_{n\in\mathbb{N}} \in A^{\mathbb{N}}$ is finitely supported. This completes the proof of Lemma 13.42.2(b). □

Lemma 13.42.3. *Let C be a* $\mathbf{k}$*-bialgebra. Let A be a* $\mathbf{k}$*-algebra. Let $f \in \operatorname{Hom}(C, A)$. Every $n \in \mathbb{N}$ satisfies*

$$f^{\star n}(1) = (f(1))^n. \tag{13.42.14}$$

Proof of Lemma 13.42.3. We must prove that (13.42.14) holds for each $n \in \mathbb{N}$. We shall prove this by induction over n:

Induction base: The axioms of a $\mathbf{k}$-bialgebra yield $\epsilon_C(1) = 1$ (since C is a $\mathbf{k}$-bialgebra). Now,

$$f^{\star 0} = (\text{the unity of the } \mathbf{k}\text{-algebra } (\operatorname{Hom}(C, A), \star)) = u_A \circ \epsilon_C.$$

Thus,

$$\underbrace{f^{\star 0}}_{=u_A\circ\epsilon_C}(1) = (u_A \circ \epsilon_C)(1) = u_A\left(\underbrace{\epsilon_C(1)}_{=1}\right) = u_A(1) = 1 \cdot 1_A = 1_A = (f(1))^0.$$

In other words, (13.42.14) holds for $n = 0$. This completes the induction base.

Induction step: Let $k \in \mathbb{N}$. Assume that (13.42.14) holds for $n = k$. We must show that (13.42.14) holds for $n = k + 1$.

The axioms of a $\mathbf{k}$-bialgebra yield $\Delta_C(1) = 1 \otimes 1$ (since C is a $\mathbf{k}$-bialgebra).

We have assumed that (13.42.14) holds for $n = k$. In other words, we have $f^{\star k}(1) = (f(1))^k$. Now, $f^{\star(k+1)} = f \star f^{\star k} = m_A \circ \left(f \otimes f^{\star k}\right) \circ \Delta_C$ (by the definition of convolution). Applying both sides of this equality to $1 \in C$, we obtain

$$\begin{aligned} f^{\star(k+1)}(1) &= \left(m_A \circ \left(f \otimes f^{\star k}\right) \circ \Delta_C\right)(1) = m_A \left(\left(f \otimes f^{\star k}\right) \left(\underbrace{\Delta_C(1)}_{=1 \otimes 1} \right) \right) = m_A \left(\underbrace{\left(f \otimes f^{\star k}\right)(1 \otimes 1)}_{=f(1) \otimes f^{\star k}(1)} \right) \\ &= m_A \left(f(1) \otimes f^{\star k}(1) \right) = f(1) \underbrace{f^{\star k}(1)}_{=(f(1))^k} \qquad \text{(by the definition of } m_A\text{)} \\ &= f(1) \cdot (f(1))^k = (f(1))^{k+1}. \end{aligned}$$

In other words, (13.42.14) holds for $n = k+1$. This completes the induction step. Thus, (13.42.14) is proven by induction. Hence, Lemma 13.42.3 is proven. □

Proof of Proposition 1.7.23. Let $x \in C$ and $y \in C$.

Lemma 13.42.2(a) (applied to $z = xy$) shows that the family $\left(\frac{1}{n!} f^{\star n}(xy)\right)_{n \in \mathbb{N}} \in A^{\mathbb{N}}$ is finitely supported and satisfies

$$(\exp^\star f)(xy) = \sum_{n \in \mathbb{N}} \frac{1}{n!} f^{\star n}(xy). \tag{13.42.15}$$

Lemma 13.42.2(b) shows that the family $\left(\sum_{i=0}^n \frac{1}{i! \cdot (n-i)!} f^{\star i}(x) f^{\star(n-i)}(y)\right)_{n \in \mathbb{N}} \in A^{\mathbb{N}}$ is finitely supported and satisfies

$$(\exp^\star f)(x) \cdot (\exp^\star f)(y) = \sum_{n \in \mathbb{N}} \sum_{i=0}^n \frac{1}{i! \cdot (n-i)!} f^{\star i}(x) f^{\star(n-i)}(y). \tag{13.42.16}$$

Comparing (13.42.16) with

$$\begin{aligned} (\exp^\star f)(xy) &= \sum_{n \in \mathbb{N}} \frac{1}{n!} \underbrace{f^{\star n}(xy)}_{\substack{=\sum_{i=0}^n \binom{n}{i} f^{\star i}(x) f^{\star(n-i)}(y) \\ \text{(by Proposition 1.7.22)}}} \qquad \text{(by (13.42.15))} \\ &= \sum_{n \in \mathbb{N}} \frac{1}{n!} \sum_{i=0}^n \binom{n}{i} f^{\star i}(x) f^{\star(n-i)}(y) = \sum_{n \in \mathbb{N}} \sum_{i=0}^n \underbrace{\frac{1}{n!} \binom{n}{i}}_{\substack{= \frac{1}{i! \cdot (n-i)!} \\ \left(\text{since } \binom{n}{i} = \frac{n!}{i! \cdot (n-i)!}\right)}} f^{\star i}(x) f^{\star(n-i)}(y) \\ &= \sum_{n \in \mathbb{N}} \sum_{i=0}^n \frac{1}{i! \cdot (n-i)!} f^{\star i}(x) f^{\star(n-i)}(y), \end{aligned}$$

we obtain $(\exp^\star f)(xy) = (\exp^\star f)(x) \cdot (\exp^\star f)(y)$.

Now, forget that we fixed x and y. We thus have shown that

$$(\exp^\star f)(xy) = (\exp^\star f)(x) \cdot (\exp^\star f)(y) \qquad \text{for all } x \in C \text{ and } y \in C. \tag{13.42.17}$$

Now, $\exp = \sum_{n\geq 0} \frac{1}{n!} T^n$. Hence, $\exp^\star f = \sum_{n\geq 0} \frac{1}{n!} f^{\star n}$ (by the definition of $\exp^\star f$). Hence,

$$\begin{aligned}
\left(\exp^\star f\right)(1) &= \left(\sum_{n\geq 0} \frac{1}{n!} f^{\star n}\right)(1) = \sum_{n\geq 0} \frac{1}{n!} \underbrace{f^{\star n}(1)}_{\substack{=(f(1))^n \\ \text{(by Lemma 13.42.3)}}} = \sum_{n\geq 0} \frac{1}{n!} \Big(\underbrace{f(1)}_{=0}\Big)^n \\
&= \sum_{n\geq 0} \frac{1}{n!} 0^n = \underbrace{\frac{1}{0!}}_{=\frac{1}{1}=1} \underbrace{0^0}_{=1} + \sum_{n\geq 1} \frac{1}{n!} \underbrace{0^n}_{\substack{=0 \\ \text{(since } n\geq 1\text{)}}} = 1 + \underbrace{\sum_{n\geq 1} \frac{1}{n!} 0}_{=0} = 1.
\end{aligned}$$

Thus, we know that $\exp^\star f$ is a $\mathbf{k}$-linear map $C \to A$ (since $\exp^\star f \in \operatorname{Hom}(C, A)$) satisfying (13.42.17) and $\left(\exp^\star f\right)(1) = 1$. In other words, $\exp^\star f : C \to A$ is a $\mathbf{k}$-algebra homomorphism. This proves Proposition 1.7.23. □

Proof of Lemma 1.7.24. The matrix $\left(i^{N+1-j}\right)_{i,j=1,2,\ldots,N+1} \in \mathbb{Q}^{(N+1)\times(N+1)}$ is invertible (since its determinant is the Vandermonde determinant $\prod_{1\leq i<j\leq N+1} \underbrace{(i-j)}_{\neq 0} \neq 0$). Let $\left(s_{i,j}\right)_{i,j=1,2,\ldots,N+1} \in \mathbb{Q}^{(N+1)\times(N+1)}$ be its inverse matrix. Then, $\left(s_{i,j}\right)_{i,j=1,2,\ldots,N+1} \cdot \left(i^{N+1-j}\right)_{i,j=1,2,\ldots,N+1} = I_{N+1}$ (the $(N+1)\times(N+1)$ identity matrix). Comparing the entries on both sides of this equality, we see: For every $(u,v) \in \{1,2,\ldots,N+1\}^2$, we have

$$\sum_{j=1}^{N+1} s_{u,j} j^{N+1-v} = \delta_{u,v}$$

(because $\sum_{j=1}^{N+1} s_{u,j} j^{N+1-v}$ is the (u,v)-th entry of the matrix $\left(s_{i,j}\right)_{i,j=1,2,\ldots,N+1} \cdot \left(i^{N+1-j}\right)_{i,j=1,2,\ldots,N+1}$, while $\delta_{u,v}$ is the (u,v)-th entry of the matrix I_{N+1}).

The entries $s_{i,j}$ of the matrix $\left(s_{i,j}\right)_{i,j=1,2,\ldots,N+1}$ are rational numbers, and therefore there exists a positive integer M such that every $(i,j) \in \{1,2,\ldots,N+1\}^2$ satisfies $M s_{i,j} \in \mathbb{Z}$ (because finitely many rational numbers always have a common denominator). Consider this M. For every $(i,j) \in \{1,2,\ldots,N+1\}^2$, define an element $t_{i,j} \in \mathbb{Z}$ by $t_{i,j} = M s_{i,j}$. Then,

$$\sum_{j=1}^{N+1} \underbrace{t_{u,j}}_{=M s_{u,j}} j^{N+1-v} = M \underbrace{\sum_{j=1}^{N+1} s_{u,j} j^{N+1-v}}_{=\delta_{u,v}} = M \delta_{u,v} \tag{13.42.18}$$

for every $(u,v) \in \{1,2,\ldots,N+1\}^2$.

Now, let $i \in \{0,1,\ldots,N\}$. Then,

$$\begin{aligned}
\sum_{j=1}^{N+1} t_{N+1-i,j} \sum_{k=0}^{N} w_k j^k &= \sum_{k=0}^{N} \underbrace{\left(\sum_{j=1}^{N+1} t_{N+1-i,j} j^k\right)}_{\substack{=M\delta_{N+1-i,N+1-k} \\ \text{(by (13.42.18),} \\ \text{applied to } u=N+1-i \text{ and } v=N+1-k\text{)}}} w_k = M \sum_{k=0}^{N} \underbrace{\delta_{N+1-i,N+1-k}}_{=\delta_{i,k}} w_k \\
&= M \underbrace{\sum_{k=0}^{N} \delta_{i,k} w_k}_{=w_i} = M w_i.
\end{aligned}$$

Hence,

$$M w_i = \sum_{j=1}^{N+1} t_{N+1-i,j} \underbrace{\sum_{k=0}^{N} w_k j^k}_{\substack{=0 \\ \text{(by (1.7.9), applied} \\ \text{to } n=j\text{)}}} = \sum_{j=1}^{N+1} t_{N+1-i,j} 0 = 0.$$

Since M is a positive integer, we can cancel M from this equality (because V is torsionfree), and thus obtain $w_i = 0$.

Now forget that we fixed i. We thus have proven that $w_i = 0$ for every $i \in \{0, 1, \ldots, N\}$. In other words, $w_k = 0$ for every $k \in \{0, 1, \ldots, N\}$. Lemma 1.7.24 is thus proven. □

Proof of Lemma 1.7.25. The family $(w_k)_{k\in\mathbb{N}}$ is finitely supported. In other words, all but finitely many $k \in \mathbb{N}$ satisfy $w_k = 0$. In other words, there exists a finite subset K of $\mathbb{N}$ such that

$$\text{each } k \in \mathbb{N} \setminus K \text{ satisfies } w_k = 0. \tag{13.42.19}$$

Consider this K.

The set K is a finite subset of $\mathbb{N}$, and thus has an upper bound (since any finite subset of $\mathbb{N}$ has an upper bound). In other words, there exists some $N \in \mathbb{N}$ such that

$$\text{each } n \in K \text{ satisfies } n \leq N. \tag{13.42.20}$$

Consider this N.

Now,

$$\text{each } k \in \mathbb{N} \text{ satisfying } k > N \text{ satisfies } w_k = 0 \tag{13.42.21}$$

[511].

We have assumed that $\sum_{k\in\mathbb{N}} w_k n^k = 0$ for all $n \in \mathbb{N}$. Thus, for all $n \in \mathbb{N}$, we have

$$0 = \sum_{k\in\mathbb{N}} w_k n^k = \underbrace{\sum_{\substack{k\in\mathbb{N};\\ k\leq N}}}_{=\sum_{k=0}^{N}} w_k n^k + \sum_{\substack{k\in\mathbb{N};\\ k>N}} \underbrace{w_k}_{\substack{=0\\ \text{(by (13.42.21))}}} n^k = \sum_{k=0}^{N} w_k n^k + \underbrace{\sum_{\substack{k\in\mathbb{N};\\ k>N}} 0 n^k}_{=0} = \sum_{k=0}^{N} w_k n^k.$$

Hence, for all $n \in \mathbb{N}$, we have $\sum_{k=0}^{N} w_k n^k = 0$. Thus, Lemma 1.7.24 shows that

$$w_k = 0 \qquad \text{for every } k \in \{0, 1, \ldots, N\}. \tag{13.42.22}$$

Now, it is easy to see that $w_k = 0$ for every $k \in \mathbb{N}$ [512]. This proves Lemma 1.7.25. □

Before we come to the proof of Proposition 1.7.26, we state another lemma, which is an easy consequence of Lemma 1.7.25:

Lemma 13.42.4. *Let V be a torsionfree abelian group (written additively). Let $(a_k)_{k\in\mathbb{N}} \in V^{\mathbb{N}}$ and $(b_k)_{k\in\mathbb{N}} \in V^{\mathbb{N}}$ be two finitely supported families of elements of V. Assume that*

$$\sum_{k\in\mathbb{N}} a_k t^k = \sum_{k\in\mathbb{N}} b_k t^k \qquad \text{for all } t \in \mathbb{N}. \tag{13.42.23}$$

Then, $a_k = b_k$ for every $k \in \mathbb{N}$.

Proof of Lemma 13.42.4. The families $(a_k)_{k\in\mathbb{N}}$ and $(b_k)_{k\in\mathbb{N}}$ are finitely supported. Hence, (by a straightforward and well-known argument) it follows that the family $(a_k - b_k)_{k\in\mathbb{N}} \in V^{\mathbb{N}}$ is also finitely supported.

Now, let $t \in \mathbb{N}$. Then, the family $\left(a_k t^k\right)_{k\in\mathbb{N}} \in V^{\mathbb{N}}$ is finitely supported.[513] Similarly, the family $\left(b_k t^k\right)_{k\in\mathbb{N}} \in V^{\mathbb{N}}$ is finitely supported.

Recall that sums of the form $\sum_{q\in Q} f_q$ (where $(f_q)_{q\in Q}$ is a finitely supported family) satisfy the usual rules for finite sums, even though their indexing set Q may be infinite. In particular, this pertains to the

[511] *Proof of (13.42.21):* Let $k \in \mathbb{N}$ be such that $k > N$. We must prove that $w_k = 0$.

If we had $k \in K$, then we would have $k \leq N$ (by (13.42.20) (applied to $n = k$)), which would contradict $k > N$. Hence, we cannot have $k \in K$. Thus, we have $k \notin K$. Combining $k \in \mathbb{N}$ with $k \notin K$, we obtain $k \in \mathbb{N} \setminus K$. Hence, (13.42.19) yields $w_k = 0$. This proves (13.42.21).

[512] *Proof.* Let $k \in \mathbb{N}$. We must prove that $w_k = 0$.

If $k \in \{0, 1, \ldots, N\}$, then this follows immediately from (13.42.22). Hence, for the rest of this proof, we WLOG assume that we don't have $k \in \{0, 1, \ldots, N\}$. Thus, $k \notin \{0, 1, \ldots, N\}$. Hence, $k \in \mathbb{N} \setminus \{0, 1, \ldots, N\} = \{N+1, N+2, N+3, \ldots\}$. Therefore, $k \geq N+1 > N$. Thus, (13.42.21) shows that $w_k = 0$. Qed.

[513] This is an easy consequence of the fact that the family $(a_k)_{k\in\mathbb{N}}$ is finitely supported.

sums $\sum_{k\in\mathbb{N}} a_k t^k$ and $\sum_{k\in\mathbb{N}} b_k t^k$ (since the families $\left(a_k t^k\right)_{k\in\mathbb{N}}$ and $\left(b_k t^k\right)_{k\in\mathbb{N}}$ are finitely supported). Hence, the following manipulations are valid:

$$\sum_{k\in\mathbb{N}} a_k t^k - \sum_{k\in\mathbb{N}} b_k t^k = \sum_{k\in\mathbb{N}} \underbrace{\left(a_k t^k - b_k t^k\right)}_{=(a_k-b_k)t^k} = \sum_{k\in\mathbb{N}} \left(a_k - b_k\right) t^k .$$

Hence,

$$\sum_{k\in\mathbb{N}} \left(a_k - b_k\right) t^k = \sum_{k\in\mathbb{N}} a_k t^k - \sum_{k\in\mathbb{N}} b_k t^k = 0 \qquad (13.42.24)$$

(by (13.42.23)).

Now, forget that we fixed t. We thus have proven that $\sum_{k\in\mathbb{N}} \left(a_k - b_k\right) t^k = 0$ for all $t \in \mathbb{N}$. Renaming the index t as n in this statement, we conclude that $\sum_{k\in\mathbb{N}} \left(a_k - b_k\right) n^k = 0$ for all $n \in \mathbb{N}$. Hence, Lemma 1.7.25 (applied to $(w_k)_{k\in\mathbb{N}} = (a_k - b_k)_{k\in\mathbb{N}}$) shows that $a_k - b_k = 0$ for every $k \in \mathbb{N}$. In other words, $a_k = b_k$ for every $k \in \mathbb{N}$. This proves Lemma 13.42.4. □

Proof of Proposition 1.7.26. Any **k**-module naturally becomes a $\mathbb{Q}$-module (since **k** is a $\mathbb{Q}$-algebra). Thus, in particular, the **k**-module A becomes a $\mathbb{Q}$-module. Hence, the **k**-module A is a torsionfree abelian group (written additively).

Now, fix $x \in \ker\epsilon$ and $y \in \ker\epsilon$. We shall show that $f(xy) = 0$.

We have $f(C_0) = 0$. Thus, Proposition 1.7.11(h) yields $f \in \mathfrak{n}(C, A)$.

Lemma 13.42.2(a) (applied to $z = xy$) shows that the family $\left(\dfrac{1}{n!} f^{\star n}(xy)\right)_{n\in\mathbb{N}} \in A^{\mathbb{N}}$ is finitely supported and satisfies

$$\left(\exp^{\star} f\right)(xy) = \sum_{n\in\mathbb{N}} \frac{1}{n!} f^{\star n}(xy) .$$

In particular, the family $\left(\dfrac{1}{n!} f^{\star n}(xy)\right)_{n\in\mathbb{N}} \in A^{\mathbb{N}}$ is finitely supported. Renaming the index n as k in this statement, we conclude that the family $\left(\dfrac{1}{k!} f^{\star k}(xy)\right)_{k\in\mathbb{N}} \in A^{\mathbb{N}}$ is finitely supported.

Lemma 13.42.2(b) yields that the family $\left(\sum_{i=0}^{n} \dfrac{1}{i!\cdot(n-i)!} f^{\star i}(x) f^{\star(n-i)}(y)\right)_{n\in\mathbb{N}} \in A^{\mathbb{N}}$ is finitely supported. In other words, the family $\left(\sum_{i=0}^{n} \dfrac{f^{\star i}(x)}{i!} \cdot \dfrac{f^{\star(n-i)}(y)}{(n-i)!}\right)_{n\in\mathbb{N}} \in A^{\mathbb{N}}$ is finitely supported (since $\dfrac{1}{i!\cdot(n-i)!} f^{\star i}(x) f^{\star(n-i)}(y) = \dfrac{f^{\star i}(x)}{i!} \cdot \dfrac{f^{\star(n-i)}(y)}{(n-i)!}$ for every $n \in \mathbb{N}$ and $i \in \{0, 1, \ldots, n\}$). Renaming the index n as k in this statement, we obtain the following: The family $\left(\sum_{i=0}^{k} \dfrac{f^{\star i}(x)}{i!} \cdot \dfrac{f^{\star(k-i)}(y)}{(k-i)!}\right)_{k\in\mathbb{N}} \in A^{\mathbb{N}}$ is finitely supported.

On the other hand, let $t \in \mathbb{N}$ be arbitrary. Then, Proposition 1.7.18(e) (applied to $n = t$) shows that $tf \in \mathfrak{n}(C, A)$ and $\exp^{\star}(tf) = \left(\exp^{\star} f\right)^{\star t}$.

Exercise 1.5.11(b) (applied to C, t and $\exp^{\star} f$ instead of H, k and f_i) shows that the map $\underbrace{\left(\exp^{\star} f\right) \star \left(\exp^{\star} f\right) \star \cdots \star \left(\exp^{\star} f\right)}_{t\text{ times}}$ is a **k**-algebra homomorphism $C \to A$. In light of

$$\underbrace{\left(\exp^{\star} f\right) \star \left(\exp^{\star} f\right) \star \cdots \star \left(\exp^{\star} f\right)}_{t\text{ times}} = \left(\exp^{\star} f\right)^{\star t} = \exp^{\star}(tf) \qquad \left(\text{since } \exp^{\star}(tf) = \left(\exp^{\star} f\right)^{\star t}\right),$$

this rewrites as follows: The map $\exp^{\star}(tf)$ is a **k**-algebra homomorphism $C \to A$.

Lemma 13.42.2(a) (applied to tf and xy instead of f and z) shows that the family $\left(\dfrac{1}{n!} (tf)^{\star n}(xy)\right)_{n\in\mathbb{N}} \in A^{\mathbb{N}}$ is finitely supported and satisfies

$$\left(\exp^{\star}(tf)\right)(xy) = \sum_{n\in\mathbb{N}} \frac{1}{n!} (tf)^{\star n}(xy) . \qquad (13.42.25)$$

Lemma 13.42.2(b) (applied to tf instead of f) shows that the family $\left(\sum_{i=0}^{n}\dfrac{1}{i!\cdot(n-i)!}(tf)^{\star i}(x)(tf)^{\star(n-i)}(y)\right)_{n\in\mathbb{N}}\in A^{\mathbb{N}}$ is finitely supported and satisfies

$$(\exp^{\star}(tf))(x)\cdot(\exp^{\star}(tf))(y)=\sum_{n\in\mathbb{N}}\sum_{i=0}^{n}\frac{1}{i!\cdot(n-i)!}(tf)^{\star i}(x)(tf)^{\star(n-i)}(y).$$

Comparing

$$\begin{aligned}
&(\exp^{\star}(tf))(x)\cdot(\exp^{\star}(tf))(y)\\
&=\sum_{n\in\mathbb{N}}\sum_{i=0}^{n}\frac{1}{i!\cdot(n-i)!}\underbrace{(tf)^{\star i}(x)}_{=t^i f^{\star i}(x)}\underbrace{(tf)^{\star(n-i)}(y)}_{=t^{n-i}f^{\star(n-i)}(y)}=\sum_{n\in\mathbb{N}}\sum_{i=0}^{n}\underbrace{\frac{1}{i!\cdot(n-i)!}t^i f^{\star i}(x)t^{n-i}f^{\star(n-i)}(y)}_{=\dfrac{f^{\star i}(x)}{i!}\cdot\dfrac{f^{\star(n-i)}(y)}{(n-i)!}t^i t^{n-i}}\\
&=\sum_{n\in\mathbb{N}}\sum_{i=0}^{n}\frac{f^{\star i}(x)}{i!}\cdot\frac{f^{\star(n-i)}(y)}{(n-i)!}\underbrace{t^i t^{n-i}}_{=t^{i+(n-i)}=t^n}=\sum_{n\in\mathbb{N}}\sum_{i=0}^{n}\frac{f^{\star i}(x)}{i!}\cdot\frac{f^{\star(n-i)}(y)}{(n-i)!}t^n\\
&=\sum_{n\in\mathbb{N}}\left(\sum_{i=0}^{n}\frac{f^{\star i}(x)}{i!}\cdot\frac{f^{\star(n-i)}(y)}{(n-i)!}\right)t^n=\sum_{k\in\mathbb{N}}\left(\sum_{i=0}^{k}\frac{f^{\star i}(x)}{i!}\cdot\frac{f^{\star(k-i)}(y)}{(k-i)!}\right)t^k\\
&\qquad\text{(here, we have renamed the summation index } n \text{ as } k\text{)}
\end{aligned}$$

with

$$\begin{aligned}
&(\exp^{\star}(tf))(x)\cdot(\exp^{\star}(tf))(y)\\
&=(\exp^{\star}(tf))(xy)\qquad\text{(since }\exp^{\star}(tf)\text{ is a }\mathbf{k}\text{-algebra homomorphism)}\\
&=\sum_{n\in\mathbb{N}}\frac{1}{n!}\underbrace{(tf)^{\star n}(xy)}_{\substack{=t^n f^{\star n}(xy)\\=f^{\star n}(xy)t^n}}\qquad\text{(by (13.42.25))}\\
&=\sum_{n\in\mathbb{N}}\frac{1}{n!}f^{\star n}(xy)t^n=\sum_{k\in\mathbb{N}}\frac{1}{k!}f^{\star k}(xy)t^k\qquad\text{(here, we have renamed the summation index } n \text{ as } k\text{)},
\end{aligned}$$

we obtain

$$\sum_{k\in\mathbb{N}}\frac{1}{k!}f^{\star k}(xy)t^k=\sum_{k\in\mathbb{N}}\left(\sum_{i=0}^{k}\frac{f^{\star i}(x)}{i!}\cdot\frac{f^{\star(k-i)}(y)}{(k-i)!}\right)t^k.$$

Now, forget that we fixed t. We thus have shown that

$$\sum_{k\in\mathbb{N}}\frac{1}{k!}f^{\star k}(xy)t^k=\sum_{k\in\mathbb{N}}\left(\sum_{i=0}^{k}\frac{f^{\star i}(x)}{i!}\cdot\frac{f^{\star(k-i)}(y)}{(k-i)!}\right)t^k$$

for all $t\in\mathbb{N}$.

Hence, we can apply Lemma 13.42.4 to $V=A$, $(a_k)_{k\in\mathbb{N}}=\left(\dfrac{1}{k!}f^{\star k}(xy)\right)_{k\in\mathbb{N}}$ and $(b_k)_{k\in\mathbb{N}}=\left(\sum_{i=0}^{k}\dfrac{f^{\star i}(x)}{i!}\cdot\dfrac{f^{\star(k-i)}(y)}{(k-i)!}\right)_{k\in\mathbb{N}}$ (since A is a torsionfree abelian group (written additively), and since the families $\left(\dfrac{1}{k!}f^{\star k}(xy)\right)_{k\in\mathbb{N}}$ and $\left(\sum_{i=0}^{k}\dfrac{f^{\star i}(x)}{i!}\cdot\dfrac{f^{\star(k-i)}(y)}{(k-i)!}\right)_{k\in\mathbb{N}}$ are finitely supported). As a result, we obtain

$$\frac{1}{k!}f^{\star k}(xy)=\sum_{i=0}^{k}\frac{f^{\star i}(x)}{i!}\cdot\frac{f^{\star(k-i)}(y)}{(k-i)!}\qquad\text{for every }k\in\mathbb{N}.$$

Applying this to $k = 1$, we find

$$\begin{aligned}\frac{1}{1!} f^{\star 1}(xy) &= \sum_{i=0}^{1} \frac{f^{\star i}(x)}{i!} \cdot \frac{f^{\star(1-i)}(y)}{(1-i)!} = \frac{f^{\star 0}(x)}{0!} \cdot \frac{f^{\star(1-0)}(y)}{(1-0)!} + \frac{f^{\star 1}(x)}{1!} \cdot \frac{f^{\star(1-1)}(y)}{(1-1)!} \\ &= \frac{f^{\star 0}(x)}{0!} \cdot \frac{f^{\star 1}(y)}{1!} + \frac{f^{\star 1}(x)}{1!} \cdot \frac{f^{\star 0}(y)}{0!} \qquad (13.42.26)\end{aligned}$$

(since $1 - 0 = 1$ and $1 - 1 = 0$).

But recall that $x \in \ker \underbrace{\epsilon}_{=\epsilon_C} = \ker(\epsilon_C)$ and thus $\epsilon_C(x) = 0$. But

$$f^{\star 0} = (\text{the unity of the } \mathbf{k}\text{-algebra } (\operatorname{Hom}(C, A), \star)) = u_A \circ \epsilon_C,$$

and thus $f^{\star 0}(x) = (u_A \circ \epsilon_C)(x) = u_A\left(\underbrace{\epsilon_C(x)}_{=0}\right) = u_A(0) = 0$. Hence, $\dfrac{f^{\star 0}(x)}{0!} = \dfrac{0}{0!} = 0$. Similarly, $\dfrac{f^{\star 0}(y)}{0!} = 0$. Hence, (13.42.26) becomes

$$\frac{1}{1!} f^{\star 1}(xy) = \underbrace{\frac{f^{\star 0}(x)}{0!}}_{=0} \cdot \frac{f^{\star 1}(y)}{1!} + \frac{f^{\star 1}(x)}{1!} \cdot \underbrace{\frac{f^{\star 0}(y)}{0!}}_{=0} = 0 \cdot \frac{f^{\star 1}(y)}{1!} + \frac{f^{\star 1}(x)}{1!} \cdot 0 = 0.$$

Comparing this with $\underbrace{\frac{1}{1!}}_{=\frac{1}{1}=1} \underbrace{f^{\star 1}}_{=f}(xy) = f(xy)$, we obtain $f(xy) = 0$.

Now, forget that we fixed x and y. We thus have shown that

$$f(xy) = 0 \qquad \text{for every } x \in \ker \epsilon \text{ and } y \in \ker \epsilon.$$

Since the map f is $\mathbf{k}$-linear, this entails that $f\left((\ker \epsilon)^2\right) = 0$ (because the $\mathbf{k}$-module $(\ker \epsilon)^2$ is spanned by elements of the form xy with $x \in \ker \epsilon$ and $y \in \ker \epsilon$). This proves Proposition 1.7.26. □

Proof of Proposition 1.7.27. We know that $f(C)$ generates the $\mathbf{k}$-algebra A. Thus, any $\mathbf{k}$-subalgebra of A that contains $f(C)$ as a subset must be the whole A. In other words,

$$\left(\begin{array}{c}\text{if } D \text{ is a } \mathbf{k}\text{-subalgebra of } A \text{ satisfying } f(C) \subset D, \\ \text{then } D = A\end{array}\right). \qquad (13.42.27)$$

Define a $\mathbf{k}$-linear map $F \in \operatorname{Hom}(C, A)$ by $F = \exp^\star f$. (This is well-defined, since $f \in \mathfrak{n}(C, A)$.)

Proposition 1.7.23 shows that $\exp^\star f : C \to A$ is a $\mathbf{k}$-algebra homomorphism. In other words, $F : C \to A$ is a $\mathbf{k}$-algebra homomorphism (since $F = \exp^\star f$).

Proposition 1.7.18(a) yields that $\exp^\star f - u_A \epsilon_C \in \mathfrak{n}(C, A)$ and $\log^\star(\exp^\star f) = f$.

Define an element $\widetilde{F}$ of $\mathfrak{n}(C, A)$ by $\widetilde{F} = F - u_A \epsilon_C$. (This is well-defined, since $\underbrace{F}_{=\exp^\star f} - u_A \epsilon_C = \exp^\star f - u_A \epsilon_C \in \mathfrak{n}(C, A)$.) From $\log^\star(\exp^\star f) = f$, we obtain

$$\begin{aligned} f &= \log^\star \underbrace{(\exp^\star f)}_{=F} = \log^\star F = \overline{\log}^\star \underbrace{(F - u_A \epsilon_C)}_{=\widetilde{F}} \qquad (\text{by the definition of } \log^\star F) \\ &= \overline{\log}^\star \widetilde{F}. \qquad (13.42.28)\end{aligned}$$

On the other hand, Proposition 1.7.11(j) (applied to C and id_C instead of A and F) shows that $\operatorname{id}_C - u_C \epsilon_C \in \mathfrak{n}(C, C)$ (since $\operatorname{id}_C : C \to C$ is a $\mathbf{k}$-algebra homomorphism).

Define an element $\widetilde{\operatorname{id}}$ of $\mathfrak{n}(C,C)$ by $\widetilde{\operatorname{id}} = \operatorname{id}_C - u_C\epsilon_C$. (This is well-defined, since $\operatorname{id}_C - u_C\epsilon_C \in \mathfrak{n}(C,C)$.) Then, $F \circ \widetilde{\operatorname{id}} = \widetilde{F}$ [514]. Hence, (13.42.28) becomes

$$f = \overline{\log}^\star \underbrace{\widetilde{F}}_{= F\circ\widetilde{\operatorname{id}}} = \overline{\log}^\star \left(F \circ \widetilde{\operatorname{id}}\right). \tag{13.42.29}$$

Proposition 1.7.11(i) (applied to C, A, F, $\overline{\log}$ and $\widetilde{\operatorname{id}}$ instead of A, B, s, u and f) shows that

$$F \circ \widetilde{\operatorname{id}} \in \mathfrak{n}(C,A) \qquad \text{and} \qquad \overline{\log}^\star \left(F \circ \widetilde{\operatorname{id}}\right) = F \circ \left(\overline{\log}^\star \left(\widetilde{\operatorname{id}}\right)\right).$$

Now, (13.42.29) becomes

$$f = \overline{\log}^\star \left(F \circ \widetilde{\operatorname{id}}\right) = F \circ \left(\overline{\log}^\star \left(\widetilde{\operatorname{id}}\right)\right).$$

Hence,

$$f(C) = \left(F \circ \left(\overline{\log}^\star \left(\widetilde{\operatorname{id}}\right)\right)\right)(C) = F\left(\underbrace{\left(\overline{\log}^\star \left(\widetilde{\operatorname{id}}\right)\right)(C)}_{\subset C}\right) \subset F(C).$$

But $F(C)$ is a $\mathbf{k}$-subalgebra of A (since F is a $\mathbf{k}$-algebra homomorphism). Hence, (13.42.27) (applied to $D = F(C)$) shows that $F(C) = A$ (since $f(C) \subset F(C)$). In other words, the map F is surjective.

Hence, F is a surjective $\mathbf{k}$-algebra homomorphism. In other words, $\exp^\star f$ is a surjective $\mathbf{k}$-algebra homomorphism (since $F = \exp^\star f$). This proves Proposition 1.7.27. □

We have now proven Lemmas 1.7.24 and 1.7.25 and Propositions 1.7.21, 1.7.22, 1.7.23, 1.7.26 and 1.7.27. Thus, Exercise 1.7.28 is solved.

13.43. **Solution to Exercise 1.7.33.**

Solution to Exercise 1.7.33. We begin by proving some simple lemmas:

Lemma 13.43.1. *Let C be a $\mathbf{k}$-coalgebra. Let A be a $\mathbf{k}$-algebra. Let $f \in \operatorname{Hom}(C,A)$. Then, every $n \in \mathbb{N}$ satisfies*

$$f^{\star n}(C) \subset (f(C))^n. \tag{13.43.1}$$

(Here, we set $V^0 = \mathbf{k} \cdot 1_A$ for any $\mathbf{k}$-submodule V of A.)

Proof of Lemma 13.43.1. We must prove the relation (13.43.1). We shall prove it by induction over n:

Induction base: We have $f^{\star 0} = (\text{the unity of the } \mathbf{k}\text{-algebra } (\operatorname{Hom}(C,A), \star)) = u_A\epsilon_C$. Thus, every $x \in C$ satisfies

$$\underbrace{f^{\star 0}}_{=u_A\epsilon_C}(x) = (u_A\epsilon_C)(x) = u_A(\epsilon_C(x)) = \underbrace{\epsilon_C(x)}_{\in\mathbf{k}} \cdot 1_A \qquad (\text{by the definition of } u_A)$$
$$\in \mathbf{k}\cdot 1_A = (f(C))^0 \qquad \left(\text{since } (f(C))^0 = \mathbf{k}\cdot 1_A\right).$$

In other words, we have $f^{\star 0}(C) \subset (f(C))^0$. In other words, (13.43.1) holds for $n = 0$.

Induction step: Let $N \in \mathbb{N}$. Assume that (13.43.1) holds for $n = N$. We must prove that (13.43.1) holds for $n = N + 1$.

[514] *Proof.* Recall that F is a $\mathbf{k}$-algebra homomorphism. In other words, F is a $\mathbf{k}$-linear map satisfying $F \circ m_C = m_A \circ (F \otimes F)$ and $F \circ u_C = u_A$. Now,

$$F \circ \underbrace{\widetilde{\operatorname{id}}}_{=\operatorname{id}_C - u_C\epsilon_C} = F \circ (\operatorname{id}_C - u_C\epsilon_C) = \underbrace{F \circ \operatorname{id}_C}_{=F} - \underbrace{F \circ (u_C\epsilon_C)}_{=F\circ u_C \circ \epsilon_C} \qquad (\text{since composition of } \mathbf{k}\text{-linear maps is } \mathbf{k}\text{-bilinear})$$
$$= F - \underbrace{F \circ u_C}_{=u_A} \circ \epsilon_C = F - u_A \circ \epsilon_C = F - u_A\epsilon_C = \widetilde{F}.$$

If X and Y are two $\mathbf{k}$-submodules of A, then the multiplication $m_A : A \otimes A \to A$ of the $\mathbf{k}$-algebra A satisfies

$$m_A (X \otimes Y) = XY. \tag{13.43.2}$$

(This follows easily from the definitions of m_A and of XY.)

We have assumed that (13.43.1) holds for $n = N$. In other words, we have $f^{\star N}(C) \subset (f(C))^N$. But

$$f^{\star(N+1)} = f \star f^{\star N} = m_A \circ \left(f \otimes f^{\star N}\right) \circ \Delta_C \qquad \text{(by the definition of convolution)}.$$

Hence,

$$\begin{aligned}\underbrace{f^{\star(N+1)}}_{=m_A\circ(f\otimes f^{\star N})\circ\Delta_C}(C) &= \left(m_A \circ \left(f \otimes f^{\star N}\right) \circ \Delta_C\right)(C) = m_A\left(\left(f \otimes f^{\star N}\right)\left(\underbrace{\Delta_C(C)}_{\subset C\otimes C}\right)\right) \\ &\subset m_A\left(\underbrace{\left(f \otimes f^{\star N}\right)(C \otimes C)}_{=f(C)\otimes f^{\star N}(C)}\right) = m_A\left(f(C) \otimes f^{\star N}(C)\right) \\ &= f(C) \cdot \underbrace{f^{\star N}(C)}_{\subset (f(C))^N} \qquad \left(\text{by (13.43.2) (applied to } X = f(C) \text{ and } Y = f^{\star N}(C)\text{)}\right) \\ &\subset f(C) \cdot (f(C))^N = (f(C))^{N+1}.\end{aligned}$$

In other words, (13.43.1) holds for $n = N+1$. This completes the induction step. Thus, the proof of (13.43.1) by induction is complete. In other words, Lemma 13.43.1 is proven. □

Lemma 13.43.2. *Let A be a $\mathbf{k}$-bialgebra. Let $\widetilde{\mathrm{id}}$ be the $\mathbf{k}$-linear map $\mathrm{id}_A - u_A\epsilon_A : A \to A$. Then:*

(a) *We have $\ker \widetilde{\mathrm{id}} = \mathbf{k} \cdot 1_A$.*
(b) *We have $\widetilde{\mathrm{id}}(A) \subset \ker \epsilon$.*

Proof of Lemma 13.43.2. The axioms of a $\mathbf{k}$-bialgebra yield $\epsilon_A(1_A) = 1$ (since A is a $\mathbf{k}$-bialgebra). Also, the definition of u_A yields $u_A(1) = 1 \cdot 1_A = 1_A$.

We have

$$\begin{aligned}\underbrace{\widetilde{\mathrm{id}}}_{=\mathrm{id}_A - u_A\epsilon_A}(1_A) &= (\mathrm{id}_A - u_A\epsilon_A)(1_A) = \underbrace{\mathrm{id}_A(1_A)}_{=1_A} - \underbrace{(u_A\epsilon_A)(1_A)}_{=u_A(\epsilon_A(1_A))} = 1_A - u_A\left(\underbrace{\epsilon_A(1_A)}_{=1}\right) \\ &= 1_A - \underbrace{u_A(1)}_{=1_A} = 1_A - 1_A = 0.\end{aligned}$$

Now, the map $\widetilde{\mathrm{id}} = \mathrm{id}_A - u_A\epsilon_A$ is $\mathbf{k}$-linear (since all three maps id_A, u_A and ϵ_A are $\mathbf{k}$-linear). Therefore, $\widetilde{\mathrm{id}}(\mathbf{k} \cdot 1_A) = \mathbf{k} \cdot \underbrace{\widetilde{\mathrm{id}}(1_A)}_{=0} = \mathbf{k} \cdot 0 = 0$. Hence, $\mathbf{k} \cdot 1_A \subset \ker \widetilde{\mathrm{id}}$.

On the other hand, let $x \in \ker \widetilde{\mathrm{id}}$. Thus, $\widetilde{\mathrm{id}}(x) = 0$. Comparing this to

$$\underbrace{\widetilde{\mathrm{id}}}_{=\mathrm{id}_A - u_A\epsilon_A}(x) = (\mathrm{id}_A - u_A\epsilon_A)(x) = \underbrace{\mathrm{id}_A(x)}_{=x} - \underbrace{(u_A\epsilon_A)(x)}_{\substack{=u_A(\epsilon_A(x)) \\ =\epsilon_A(x)\cdot 1_A \\ \text{(by the definition of } u_A\text{)}}} = x - \epsilon_A(x) \cdot 1_A,$$

we obtain $x - \epsilon_A(x) \cdot 1_A = 0$. Thus, $x = \underbrace{\epsilon_A(x)}_{\in\mathbf{k}} \cdot 1_A \in \mathbf{k} \cdot 1_A$.

Now, forget that we fixed x. We thus have proven that $x \in \mathbf{k} \cdot 1_A$ for each $x \in \ker \widetilde{\mathrm{id}}$. In other words, $\ker \widetilde{\mathrm{id}} \subset \mathbf{k} \cdot 1_A$. Combining this with $\mathbf{k} \cdot 1_A \subset \ker \widetilde{\mathrm{id}}$, we obtain $\ker \widetilde{\mathrm{id}} = \mathbf{k} \cdot 1_A$. This proves Lemma 13.43.2(a).

(b) The axioms of a **k**-bialgebra yield $\epsilon_A \circ u_A = \mathrm{id}_{\mathbf{k}}$ (since A is a **k**-bialgebra). Now,

$$\begin{aligned}\epsilon_A \circ \underbrace{\widetilde{\mathrm{id}}}_{=\mathrm{id}_A - u_A \epsilon_A} &= \epsilon_A \circ \left(\mathrm{id}_A - u_A \epsilon_A\right) = \underbrace{\epsilon_A \circ \mathrm{id}_A}_{=\epsilon_A} - \underbrace{\epsilon_A \circ \left(u_A \epsilon_A\right)}_{=\epsilon_A \circ u_A \circ \epsilon_A} \\ &\qquad \text{(since composition of } \mathbf{k}\text{-linear maps is } \mathbf{k}\text{-bilinear)} \\ &= \epsilon_A - \underbrace{\epsilon_A \circ u_A}_{=\mathrm{id}_{\mathbf{k}}} \circ \epsilon_A = \epsilon_A - \epsilon_A = 0.\end{aligned}$$

Hence, $\epsilon_A \left(\widetilde{\mathrm{id}}\,(A)\right) = \underbrace{\left(\epsilon_A \circ \widetilde{\mathrm{id}}\right)}_{=0} (A) = 0\,(A) = 0$. Therefore, $\widetilde{\mathrm{id}}\,(A) \subset \ker\left(\underbrace{\epsilon_A}_{=\epsilon}\right) = \ker \epsilon$. This proves Lemma 13.43.2(b). □

Lemma 13.43.3. *Let A be a* **k***-bialgebra. Then, $A / \left(\mathbf{k} \cdot 1_A + (\ker \epsilon)^2\right) \cong (\ker \epsilon) / (\ker \epsilon)^2$ as* **k***-modules.*

Proof of Lemma 13.43.3. The axioms of a **k**-bialgebra show that the counit ϵ of A is a **k**-algebra homomorphism (since A is a **k**-bialgebra). Thus, $\ker \epsilon$ is an ideal of A. Now, $(\ker \epsilon)^2 = \underbrace{(\ker \epsilon)}_{\subset A} (\ker \epsilon) \subset A (\ker \epsilon) \subset \ker \epsilon$ (since $\ker \epsilon$ is an ideal of A). Hence, the quotient $(\ker \epsilon) / (\ker \epsilon)^2$ makes sense.

Let $\widetilde{\mathrm{id}}$ be the map $\mathrm{id}_A - u_A \epsilon_A : A \to A$. Then, Lemma 13.43.2(a) shows that $\ker \widetilde{\mathrm{id}} = \mathbf{k} \cdot 1_A$.

Lemma 13.43.2(b) shows that $\widetilde{\mathrm{id}}\,(A) \subset \ker \epsilon$. Thus, each $a \in A$ satisfies $\widetilde{\mathrm{id}}\left(\underbrace{a}_{\in A}\right) \in \widetilde{\mathrm{id}}\,(A) \subset \ker \epsilon$. Hence, we can define a map $\pi : A \to \ker \epsilon$ by

$$\left(\pi(a) = \widetilde{\mathrm{id}}\,(a) \qquad \text{for each } a \in A\right).$$

Consider this map π. This map π differs from $\widetilde{\mathrm{id}}$ only in its target (namely, its target is $\ker \epsilon$, whereas the target of $\widetilde{\mathrm{id}}$ is A); therefore, this map π is **k**-linear (since the map $\widetilde{\mathrm{id}}$ is **k**-linear).

Each $a \in \ker \epsilon$ satisfies

$$\begin{aligned}\pi(a) &= \underbrace{\widetilde{\mathrm{id}}}_{=\mathrm{id}_A - u_A \epsilon_A} (a) = \left(\mathrm{id}_A - u_A \epsilon_A\right)(a) = \underbrace{\mathrm{id}_A(a)}_{=a} - \underbrace{\left(u_A \epsilon_A\right)(a)}_{=u_A(\epsilon_A(a))} = a - u_A\left(\underbrace{\epsilon_A}_{=\epsilon}(a)\right) \\ &= a - u_A\left(\underbrace{\epsilon(a)}_{\substack{=0 \\ (\text{since } a \in \ker \epsilon)}}\right) = a - \underbrace{u_A(0)}_{\substack{=0 \\ (\text{since the map } u_A \text{ is } \mathbf{k}\text{-linear})}} = a.\end{aligned}$$

Hence, each $a \in \ker \epsilon$ satisfies $a = \pi\left(\underbrace{a}_{\in A}\right) \in \pi(A)$. In other words, we have $\ker \epsilon \subset \pi(A)$. Hence, the map π is surjective.

Let γ be the canonical projection $\ker \epsilon \to (\ker \epsilon) / (\ker \epsilon)^2$. Thus, the map γ is surjective and satisfies $\ker \gamma = (\ker \epsilon)^2$.

Both maps γ and π are **k**-linear. Thus, their composition $\gamma \circ \pi$ is **k**-linear. Hence, its kernel $\ker(\gamma \circ \pi)$ is a **k**-submodule of A.

The map $\gamma \circ \pi : A \to (\ker \epsilon) / (\ker \epsilon)^2$ is the composition of two surjective maps (since the two maps γ and π are surjective), and thus is itself surjective. In other words, $(\ker \epsilon) / (\ker \epsilon)^2 = (\gamma \circ \pi)(A)$.

It is known that if V and W are two **k**-modules, and if $\delta : V \to W$ is a **k**-linear map, then $\delta(V) \cong V / \ker \delta$ as **k**-modules. Applying this to $V = A$, $W = (\ker \epsilon) / (\ker \epsilon)^2$ and $\delta = \gamma \circ \pi$, we conclude that $(\gamma \circ \pi)(A) \cong A / \ker(\gamma \circ \pi)$ as **k**-modules. Thus,

$$(\ker \epsilon) / (\ker \epsilon)^2 = (\gamma \circ \pi)(A) \cong A / \ker(\gamma \circ \pi) \tag{13.43.3}$$

as **k**-modules.

Now, let us show that $\ker(\gamma\circ\pi)=\mathbf{k}\cdot 1_A+(\ker\epsilon)^2$.

We first observe that $\mathbf{k}\cdot 1_A\subset\ker(\gamma\circ\pi)$ [515] and $(\ker\epsilon)^2\subset\ker(\gamma\circ\pi)$ [516]. Hence,

$$\underbrace{\mathbf{k}\cdot 1_A}_{\subset\ker(\gamma\circ\pi)}+\underbrace{(\ker\epsilon)^2}_{\subset\ker(\gamma\circ\pi)}\subset\ker(\gamma\circ\pi)+\ker(\gamma\circ\pi)\subset\ker(\gamma\circ\pi)\tag{13.43.4}$$

(since $\ker(\gamma\circ\pi)$ is a **k**-submodule of A).

On the other hand, let $z\in\ker(\gamma\circ\pi)$ be arbitrary. We shall show that $z\in\mathbf{k}\cdot 1_A+(\ker\epsilon)^2$.

We have $\gamma(\pi(z))=(\gamma\circ\pi)(z)=0$ (since $z\in\ker(\gamma\circ\pi)$), so that $\pi(z)\in\ker\gamma=(\ker\epsilon)^2$. But the definition of π yields

$$\pi(z)=\underbrace{\widetilde{\mathrm{id}}}_{=\mathrm{id}_A-u_A\epsilon_A}(z)=(\mathrm{id}_A-u_A\epsilon_A)(z)=\underbrace{\mathrm{id}_A(z)}_{=z}-\underbrace{(u_A\epsilon_A)(z)}_{\substack{=u_A(\epsilon_A(z))\\ =\epsilon_A(z)\cdot 1_A\\ \text{(by the definition of }u_A)}}=z-\epsilon_A(z)\cdot 1_A.$$

Hence,

$$z-\epsilon_A(z)\cdot 1_A=\pi(z)\in(\ker\epsilon)^2.$$

Therefore,

$$z\in\underbrace{\epsilon_A(z)}_{\in\mathbf{k}}\cdot 1_A+(\ker\epsilon)^2\subset\mathbf{k}\cdot 1_A+(\ker\epsilon)^2.$$

Now, forget that we fixed z. We thus have shown that $z\in\mathbf{k}\cdot 1_A+(\ker\epsilon)^2$ for each $z\in\ker(\gamma\circ\pi)$. In other words, $\ker(\gamma\circ\pi)\subset\mathbf{k}\cdot 1_A+(\ker\epsilon)^2$. Combining this with (13.43.4), we obtain $\ker(\gamma\circ\pi)=\mathbf{k}\cdot 1_A+(\ker\epsilon)^2$. Thus, (13.43.3) becomes

$$(\ker\epsilon)/(\ker\epsilon)^2\cong A/\underbrace{\ker(\gamma\circ\pi)}_{=\mathbf{k}\cdot 1_A+(\ker\epsilon)^2}=A/\left(\mathbf{k}\cdot 1_A+(\ker\epsilon)^2\right)$$

as **k**-modules. This proves Lemma 13.43.3. □

Proof of Theorem 1.7.29. Proposition 1.7.11(j) (applied to $C=A$ and $F=\mathrm{id}_A$) yields that $\mathrm{id}_A-u_A\epsilon_A\in\mathfrak{n}(A,A)$ (since $\mathrm{id}_A:A\to A$ is a **k**-algebra homomorphism). Thus, the map $\log^\star(\mathrm{id}_A)\in\mathfrak{n}(A,A)$ is well-defined. This proves Theorem 1.7.29(a).

Define an element $\widetilde{\mathrm{id}}$ of $\mathfrak{n}(A,A)$ by $\widetilde{\mathrm{id}}=\mathrm{id}_A-u_A\epsilon_A$. (This is well-defined, since $\mathrm{id}_A-u_A\epsilon_A\in\mathfrak{n}(A,A)$.)

[515] *Proof.* We have $1_A=\underbrace{1}_{\in\mathbf{k}}\cdot 1_A\in\mathbf{k}\cdot 1_A=\ker\widetilde{\mathrm{id}}$ (since $\ker\widetilde{\mathrm{id}}=\mathbf{k}\cdot 1_A$) and thus $\widetilde{\mathrm{id}}(1_A)=0$. The definition of π yields $\pi(1_A)=\widetilde{\mathrm{id}}(1_A)=0$. Now, $(\gamma\circ\pi)(1_A)=\gamma\left(\underbrace{\pi(1_A)}_{=0}\right)=\gamma(0)=0$ (since the map γ is **k**-linear). Since the map $\gamma\circ\pi$ is **k**-linear, we have $(\gamma\circ\pi)(\mathbf{k}\cdot 1_A)=\mathbf{k}\cdot\underbrace{(\gamma\circ\pi)(1_A)}_{=0}=\mathbf{k}\cdot 0=0$. In other words, $\mathbf{k}\cdot 1_A\subset\ker(\gamma\circ\pi)$.

[516] *Proof.* Let $x\in(\ker\epsilon)^2$. Thus, $x\in(\ker\epsilon)^2\subset\ker\epsilon$, so that $\epsilon(x)=0$.

The definition of π yields

$$\pi(x)=\underbrace{\widetilde{\mathrm{id}}}_{=\mathrm{id}_A-u_A\epsilon_A}(x)=(\mathrm{id}_A-u_A\epsilon_A)(x)=\underbrace{\mathrm{id}_A(x)}_{=x}-\underbrace{(u_A\epsilon_A)(x)}_{\substack{=u_A(\epsilon_A(x))\\ =\epsilon_A(x)\cdot 1_A\\ \text{(by the definition of }u_A)}}=x-\underbrace{\epsilon_A}_{=\epsilon}(x)\cdot 1_A$$
$$=x-\underbrace{\epsilon(x)}_{=0}\cdot 1_A=x-\underbrace{0\cdot 1_A}_{=0}=x\in(\ker\epsilon)^2=\ker\gamma$$

(since $\ker\gamma=(\ker\epsilon)^2$). Therefore, $\gamma(\pi(x))=0$. Now, $(\gamma\circ\pi)(x)=\gamma(\pi(x))=0$, so that $x\in\ker(\gamma\circ\pi)$.

Now, forget that we fixed x. We thus have shown that $x\in\ker(\gamma\circ\pi)$ for each $x\in(\ker\epsilon)^2$. In other words, $(\ker\epsilon)^2\subset\ker(\gamma\circ\pi)$.

(b) The axioms of a $\mathbf{k}$-bialgebra show that $\epsilon : A \to \mathbf{k}$ is a $\mathbf{k}$-algebra homomorphism. Thus, $\ker \epsilon$ is an ideal of A. Now, it is easy to see that if $n \in \mathbb{N}$ and $m \in \mathbb{N}$ satisfy $n \geq m > 0$, then[517]

$$\widetilde{\operatorname{id}}^{\star n}(A) \subset (\ker \epsilon)^m . \tag{13.43.5}$$

[*Proof of (13.43.5):* Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$ be such that $n \geq m > 0$. Then, $(\ker \epsilon)^m$ is an ideal of A (since $\ker \epsilon$ is an ideal of A, and since $m > 0$). But Lemma 13.43.1 (applied to $C = A$ and $f = \widetilde{\operatorname{id}}$) yields

$$\begin{aligned} \widetilde{\operatorname{id}}^{\star n}(A) &\subset \Bigg(\underbrace{\widetilde{\operatorname{id}}(A)}_{\substack{\subset \ker \epsilon \\ \text{(by Lemma 13.43.2(b))}}} \Bigg)^n \subset (\ker \epsilon)^n = \underbrace{(\ker \epsilon)^{n-m}}_{\subset A} (\ker \epsilon)^m \qquad (\text{since } n \geq m) \\ &\subset A (\ker \epsilon)^m \subset (\ker \epsilon)^m \qquad (\text{since } (\ker \epsilon)^m \text{ is an ideal of } A) . \end{aligned}$$

This proves (13.43.5).]

Proposition 1.7.18(f) (applied to $C = A$ and $f = \widetilde{\operatorname{id}}$) yields

$$\log^\star \left(\widetilde{\operatorname{id}} + u_A \epsilon_A \right) = \sum_{n \geq 1} \frac{(-1)^{n-1}}{n} \widetilde{\operatorname{id}}^{\star n}.$$

Since $\widetilde{\operatorname{id}} + u_A \epsilon_A = \operatorname{id}_A$ (because $\widetilde{\operatorname{id}} = \operatorname{id}_A - u_A \epsilon_A$), this rewrites as

$$\log^\star (\operatorname{id}_A) = \sum_{n \geq 1} \frac{(-1)^{n-1}}{n} \widetilde{\operatorname{id}}^{\star n}.$$

Thus,

$$\mathfrak{e} = \log^\star (\operatorname{id}_A) = \sum_{n \geq 1} \frac{(-1)^{n-1}}{n} \widetilde{\operatorname{id}}^{\star n} \tag{13.43.6}$$

$$= \underbrace{\frac{(-1)^{1-1}}{1}}_{=1} \underbrace{\widetilde{\operatorname{id}}^{\star 1}}_{= \widetilde{\operatorname{id}}} + \sum_{n \geq 2} \frac{(-1)^{n-1}}{n} \widetilde{\operatorname{id}}^{\star n} = \widetilde{\operatorname{id}} + \sum_{n \geq 2} \frac{(-1)^{n-1}}{n} \widetilde{\operatorname{id}}^{\star n}. \tag{13.43.7}$$

Furthermore,

$$\mathfrak{e}(1_A) = 0. \tag{13.43.8}$$

[*Proof of (13.43.8):* Lemma 13.42.3 (applied to $C = A$ and $f = \widetilde{\operatorname{id}}$) shows that every $n \in \mathbb{N}$ satisfies

$$\widetilde{\operatorname{id}}^{\star n}(1_A) = \left(\widetilde{\operatorname{id}}(1_A) \right)^n . \tag{13.43.9}$$

But Lemma 13.43.2(a) yields $\ker \widetilde{\operatorname{id}} = \mathbf{k} \cdot 1_A$. Hence, $1_A = \underbrace{1}_{\in \mathbf{k}} \cdot 1_A \in \mathbf{k} \cdot 1_A = \ker \widetilde{\operatorname{id}}$, so that $\widetilde{\operatorname{id}}(1_A) = 0$. Now, if n is any positive integer, then

$$\begin{aligned} \widetilde{\operatorname{id}}^{\star n}(1_A) &= \Bigg(\underbrace{\widetilde{\operatorname{id}}(1_A)}_{=0} \Bigg)^n \qquad (\text{by (13.43.9)}) \\ &= 0^n = 0 \qquad (\text{since } n \text{ is a positive integer}) . \end{aligned} \tag{13.43.10}$$

Applying both sides of the equality (13.43.6) to 1_A, we obtain

$$\mathfrak{e}(1_A) = \left(\sum_{n \geq 1} \frac{(-1)^{n-1}}{n} \widetilde{\operatorname{id}}^{\star n} \right) (1_A) = \sum_{n \geq 1} \frac{(-1)^{n-1}}{n} \underbrace{\widetilde{\operatorname{id}}^{\star n}(1_A)}_{\substack{=0 \\ \text{(by (13.43.10))}}} = \sum_{n \geq 1} \frac{(-1)^{n-1}}{n} 0 = 0.$$

This proves (13.43.8).]

[517] Here, we set $V^0 = \mathbf{k} \cdot 1_A$ for any $\mathbf{k}$-submodule V of A.

The map $\mathfrak{e}$ is $\mathbf{k}$-linear (since $\mathfrak{e} = \log^\star(\operatorname{id}_A) \in \mathfrak{n}(A, A) \subset \operatorname{Hom}(A, A)$). Thus,

$$\mathfrak{e}(\mathbf{k} \cdot 1_A) = \mathbf{k} \cdot \underbrace{\mathfrak{e}(1_A)}_{\substack{=0 \\ \text{(by (13.43.8))}}} = \mathbf{k} \cdot 0 = 0.$$

But Exercise 1.3.20(c) shows that $A_0 = \mathbf{k} \cdot 1_A$. Applying the map $\mathfrak{e}$ to both sides of this equality, we find

$$\mathfrak{e}(A_0) = \mathfrak{e}(\mathbf{k} \cdot 1_A) = 0. \tag{13.43.11}$$

Furthermore, recall that $\operatorname{id}_A - u_A \epsilon_A \in \mathfrak{n}(A, A)$. Hence, Proposition 1.7.18(b) (applied to $g = \operatorname{id}_A$) yields $\exp^\star(\log^\star(\operatorname{id}_A)) = \operatorname{id}_A$. Thus,

$$\exp^\star \underbrace{\mathfrak{e}}_{=\log^\star(\operatorname{id}_A)} = \exp^\star(\log^\star(\operatorname{id}_A)) = \operatorname{id}_A. \tag{13.43.12}$$

Thus, $\exp^\star \mathfrak{e} : A \to A$ is a $\mathbf{k}$-algebra homomorphism (since $\operatorname{id}_A : A \to A$ is a $\mathbf{k}$-algebra homomorphism). Hence, Proposition 1.7.26 (applied to $C = A$ and $f = \mathfrak{e}$) shows that

$$\mathfrak{e}\left((\ker \epsilon)^2\right) = 0. \tag{13.43.13}$$

The map $\mathfrak{e}$ is $\mathbf{k}$-linear. Thus,

$$\mathfrak{e}\left(\mathbf{k} \cdot 1_A + (\ker \epsilon)^2\right) = \mathbf{k} \cdot \underbrace{\mathfrak{e}(1_A)}_{\substack{=0 \\ \text{(by (13.43.8))}}} + \underbrace{\mathfrak{e}\left((\ker \epsilon)^2\right)}_{\substack{=0 \\ \text{(by (13.43.13))}}} = \mathbf{k} \cdot 0 + 0 = 0.$$

In other words,

$$\mathbf{k} \cdot 1_A + (\ker \epsilon)^2 \subset \ker \mathfrak{e}. \tag{13.43.14}$$

On the other hand, let us prove the reverse inclusion. Let us first observe that each $x \in A$ satisfies

$$\widetilde{\operatorname{id}}(x) - \mathfrak{e}(x) \in (\ker \epsilon)^2. \tag{13.43.15}$$

[*Proof of (13.43.15):* Let $x \in A$. Applying both sides of the equality (13.43.7) to x, we find

$$\mathfrak{e}(x) = \left(\widetilde{\operatorname{id}} + \sum_{n \geq 2} \frac{(-1)^{n-1}}{n} \widetilde{\operatorname{id}}^{\star n}\right)(x) = \widetilde{\operatorname{id}}(x) + \sum_{n \geq 2} \frac{(-1)^{n-1}}{n} \widetilde{\operatorname{id}}^{\star n}(x).$$

Subtracting $\widetilde{\operatorname{id}}(x)$ from both sides of this equality, we find

$$\mathfrak{e}(x) - \widetilde{\operatorname{id}}(x) = \sum_{n \geq 2} \frac{(-1)^{n-1}}{n} \widetilde{\operatorname{id}}^{\star n}\left(\underbrace{x}_{\in A}\right) \in \sum_{n \geq 2} \frac{(-1)^{n-1}}{n} \underbrace{\widetilde{\operatorname{id}}^{\star n}(A)}_{\substack{\subset (\ker \epsilon)^2 \\ \text{(by (13.43.5) (applied to } m=2\text{)} \\ \text{(since } n \geq 2 > 0\text{))}}}$$

$$\subset \sum_{n \geq 2} \frac{(-1)^{n-1}}{n} (\ker \epsilon)^2 \subset (\ker \epsilon)^2 \qquad \left(\text{since } (\ker \epsilon)^2 \text{ is a } \mathbf{k}\text{-submodule of } A\right).$$

Hence,

$$\widetilde{\operatorname{id}}(x) - \mathfrak{e}(x) = -\underbrace{\left(\mathfrak{e}(x) - \widetilde{\operatorname{id}}(x)\right)}_{\in (\ker \epsilon)^2} \in -(\ker \epsilon)^2 \subset (\ker \epsilon)^2 \qquad \left(\text{since } (\ker \epsilon)^2 \text{ is a } \mathbf{k}\text{-submodule of } A\right).$$

This proves (13.43.15).]

Now, let $x \in \ker \mathfrak{e}$. Thus, $\mathfrak{e}(x) = 0$. Now, (13.43.15) yields $\widetilde{\operatorname{id}}(x) - \mathfrak{e}(x) \in (\ker \epsilon)^2$. But

$$\widetilde{\operatorname{id}}(x) - \underbrace{\mathfrak{e}(x)}_{=0} = \underbrace{\widetilde{\operatorname{id}}}_{=\operatorname{id}_A - u_A \epsilon_A}(x) = (\operatorname{id}_A - u_A \epsilon_A)(x) = \underbrace{\operatorname{id}_A(x)}_{=x} - \underbrace{(u_A \epsilon_A)(x)}_{\substack{=u_A(\epsilon_A(x)) \\ =\epsilon_A(x) \cdot 1_A \\ \text{(by the definition of } u_A\text{)}}} = x - \epsilon_A(x) \cdot 1_A.$$

Thus,

$$x - \epsilon_A(x) \cdot 1_A = \widetilde{\operatorname{id}}(x) - \mathfrak{e}(x) \in (\ker \epsilon)^2 \qquad \text{(by (13.43.15))}.$$

Hence,

$$x \in \underbrace{\epsilon_A(x)}_{\in \mathbf{k}} \cdot 1_A + (\ker \epsilon)^2 \subset \mathbf{k} \cdot 1_A + (\ker \epsilon)^2.$$

Now, forget that we fixed x. We thus have shown that every $x \in \ker \mathfrak{e}$ satisfies $x \in \mathbf{k} \cdot 1_A + (\ker \epsilon)^2$. In other words, we have $\ker \mathfrak{e} \subset \mathbf{k} \cdot 1_A + (\ker \epsilon)^2$. Combining this with (13.43.14), we obtain

$$\ker \mathfrak{e} = \mathbf{k} \cdot 1_A + (\ker \epsilon)^2. \tag{13.43.16}$$

Thus, one part of Theorem 1.7.29(b) is proven. It remains to show that $\mathfrak{e}(A) \cong (\ker \epsilon) / (\ker \epsilon)^2$ (as $\mathbf{k}$-modules).

It is known that if V and W are two $\mathbf{k}$-modules, and if $\delta : V \to W$ is a $\mathbf{k}$-linear map, then $\delta(V) \cong V / \ker \delta$ as $\mathbf{k}$-modules. Applying this to $V = A$, $W = A$ and $\delta = \mathfrak{e}$, we obtain

$$\mathfrak{e}(A) \cong A / \underbrace{\ker \mathfrak{e}}_{\substack{= \mathbf{k} \cdot 1_A + (\ker \epsilon)^2 \\ \text{(by (13.43.16))}}} = A / \left(\mathbf{k} \cdot 1_A + (\ker \epsilon)^2\right) \cong (\ker \epsilon) / (\ker \epsilon)^2 \qquad \text{(by Lemma 13.43.3)}$$

as $\mathbf{k}$-modules. This completes the proof of Theorem 1.7.29(b).

(c) The definition of $\mathfrak{q}$ shows that

$$\mathfrak{q}(x) = \iota_{\mathfrak{e}(A)}(\mathfrak{e}(x)) \qquad \text{for each } x \in A. \tag{13.43.17}$$

Thus, each $x \in A_0$ satisfies

$$\mathfrak{q}(x) = \iota_{\mathfrak{e}(A)}\left(\mathfrak{e}\left(\underbrace{x}_{\in A_0}\right)\right) \in \iota_{\mathfrak{e}(A)}\left(\underbrace{\mathfrak{e}(A_0)}_{\substack{=0 \\ \text{(by (13.43.11))}}}\right) = \iota_{\mathfrak{e}(A)}(0) = 0.$$

In other words, we have $\mathfrak{q}(A_0) = 0$. Hence, Proposition 1.7.11(h) (applied to A, $\operatorname{Sym}(\mathfrak{e}(A))$ and $\mathfrak{q}$ instead of C, A and f) shows that $\mathfrak{q} \in \mathfrak{n}(A, \operatorname{Sym}(\mathfrak{e}(A)))$. This proves Theorem 1.7.29(c).

(d) Theorem 1.7.29(c) yields $\mathfrak{q} \in \mathfrak{n}(A, \operatorname{Sym}(\mathfrak{e}(A)))$. Hence, $\exp^\star \mathfrak{q} \in \operatorname{Hom}(A, \operatorname{Sym}(\mathfrak{e}(A)))$ is well-defined.

Recall that the $\mathbf{k}$-algebra $\operatorname{Sym} V$ is commutative whenever V is a $\mathbf{k}$-module. Applying this to $V = \mathfrak{e}(A)$, we conclude that the $\mathbf{k}$-algebra $\operatorname{Sym}(\mathfrak{e}(A))$ is commutative.

We have $\mathfrak{q}(A) = \operatorname{Sym}^1(\mathfrak{e}(A))$ [518]. Thus, $\mathfrak{q}(A)$ generates the $\mathbf{k}$-algebra $\operatorname{Sym}(\mathfrak{e}(A))$ [519].

Furthermore, Theorem 1.7.29(b) yields that $\ker \mathfrak{e} = \mathbf{k} \cdot 1_A + (\ker \epsilon)^2$. Hence, $1 = 1_A = \underbrace{1}_{\in \mathbf{k}} \cdot 1_A \in \mathbf{k} \cdot 1_A \subset \mathbf{k} \cdot 1_A + (\ker \epsilon)^2 = \ker \mathfrak{e}$ and $(\ker \epsilon)^2 \subset \mathbf{k} \cdot 1_A + (\ker \epsilon)^2 = \ker \mathfrak{e}$. Now, (13.43.17) (applied to $x = 1$) yields

$$\mathfrak{q}(1) = \iota_{\mathfrak{e}(A)}\left(\underbrace{\mathfrak{e}(1)}_{\substack{=0 \\ \text{(since } 1 \in \ker \mathfrak{e})}}\right) = \iota_{\mathfrak{e}(A)}(0) = 0 \qquad \left(\text{since the map } \iota_{\mathfrak{e}(A)} \text{ is } \mathbf{k}\text{-linear}\right).$$

[518] *Proof.* Every $\mathbf{k}$-module V satisfies $\iota_V(V) = \operatorname{Sym}^1 V$. Applying this to $V = \mathfrak{e}(A)$, we obtain $\iota_{\mathfrak{e}(A)}(\mathfrak{e}(A)) = \operatorname{Sym}^1(\mathfrak{e}(A))$. Now,

$$\mathfrak{q}(A) = \left\{ \underbrace{\mathfrak{q}(x)}_{\substack{= \iota_{\mathfrak{e}(A)}(\mathfrak{e}(x)) \\ \text{(by (13.43.17))}}} \mid x \in A \right\} = \left\{ \iota_{\mathfrak{e}(A)}(\mathfrak{e}(x)) \mid x \in A \right\} = \iota_{\mathfrak{e}(A)}\left(\underbrace{\{\mathfrak{e}(x) \mid x \in A\}}_{= \mathfrak{e}(A)}\right) = \iota_{\mathfrak{e}(A)}(\mathfrak{e}(A)) = \operatorname{Sym}^1(\mathfrak{e}(A)).$$

[519] *Proof.* It is known that if V is a $\mathbf{k}$-module, then $\operatorname{Sym}^1 V$ generates the $\mathbf{k}$-algebra $\operatorname{Sym} V$. Applying this to $V = \mathfrak{e}(A)$, we conclude that $\operatorname{Sym}^1(\mathfrak{e}(A))$ generates the $\mathbf{k}$-algebra $\operatorname{Sym}(\mathfrak{e}(A))$. In other words, $\mathfrak{q}(A)$ generates the $\mathbf{k}$-algebra $\operatorname{Sym}(\mathfrak{e}(A))$ (since $\mathfrak{q}(A) = \operatorname{Sym}^1(\mathfrak{e}(A))$).

Also, each $x \in (\ker \epsilon)^2$ satisfies

$$\mathfrak{q}(x) = \iota_{\mathfrak{e}(A)}\left(\underbrace{\mathfrak{e}(x)}_{\substack{=0 \\ (\text{since } x \in (\ker\epsilon)^2 \subset \ker \mathfrak{e})}}\right) \qquad (\text{by } (13.43.17))$$
$$= \iota_{\mathfrak{e}(A)}(0) = 0 \qquad \left(\text{since the map } \iota_{\mathfrak{e}(A)} \text{ is } \mathbf{k}\text{-linear}\right).$$

In other words, $\mathfrak{q}\left((\ker\epsilon)^2\right) \subset 0$. Hence, $\mathfrak{q}\left((\ker\epsilon)^2\right) = 0$.

Thus, Proposition 1.7.27 (applied to A, $\operatorname{Sym}(\mathfrak{e}(A))$ and $\mathfrak{q}$ instead of C, A and f) shows that $\exp^\star \mathfrak{q} : A \to \operatorname{Sym}(\mathfrak{e}(A))$ is a surjective $\mathbf{k}$-algebra homomorphism.

But let us recall that $\mathfrak{s}$ is the unique $\mathbf{k}$-algebra homomorphism $\Phi : \operatorname{Sym}(\mathfrak{e}(A)) \to A$ satisfying $\mathbf{i} = \Phi \circ \iota_{\mathfrak{e}(A)}$. Hence, $\mathfrak{s}$ is a $\mathbf{k}$-algebra homomorphism $\operatorname{Sym}(\mathfrak{e}(A)) \to A$ and satisfies $\mathbf{i} = \mathfrak{s} \circ \iota_{\mathfrak{e}(A)}$.

Thus, Proposition 1.7.11(i) (applied to A, $\operatorname{Sym}(\mathfrak{e}(A))$, A, $\mathfrak{s}$, exp and $\mathfrak{q}$ instead of C, A, B, s, u and f) shows that

$$\mathfrak{s} \circ \mathfrak{q} \in \mathfrak{n}(A, A) \qquad \text{and} \qquad \exp^\star(\mathfrak{s} \circ \mathfrak{q}) = \mathfrak{s} \circ (\exp^\star \mathfrak{q}).$$

Furthermore,

$$\mathfrak{s} \circ \mathfrak{q} = \mathfrak{e} \tag{13.43.18}$$

(since each $x \in A$ satisfies

$$(\mathfrak{s} \circ \mathfrak{q})(x) = \mathfrak{s}\left(\underbrace{\mathfrak{q}(x)}_{\substack{=\iota_{\mathfrak{e}(A)}(\mathfrak{e}(x)) \\ (\text{by } (13.43.17))}}\right) = \mathfrak{s}\left(\iota_{\mathfrak{e}(A)}(\mathfrak{e}(x))\right)$$
$$= \underbrace{\left(\mathfrak{s} \circ \iota_{\mathfrak{e}(A)}\right)}_{\substack{=\mathbf{i} \\ (\text{since } \mathbf{i} = \mathfrak{s} \circ \iota_{\mathfrak{e}(A)})}}(\mathfrak{e}(x)) = \mathbf{i}(\mathfrak{e}(x)) = \mathfrak{e}(x) \qquad (\text{since } \mathbf{i} \text{ is just an inclusion map})$$

). Hence, the equality $\exp^\star(\mathfrak{s} \circ \mathfrak{q}) = \mathfrak{s} \circ (\exp^\star \mathfrak{q})$ rewrites as $\exp^\star \mathfrak{e} = \mathfrak{s} \circ (\exp^\star \mathfrak{q})$. Comparing this with (13.43.12), we obtain $\mathfrak{s} \circ (\exp^\star \mathfrak{q}) = \operatorname{id}_A$. Thus, the map $\exp^\star \mathfrak{q}$ has a left inverse (with respect to composition), and hence is injective.

Now, the map $\exp^\star \mathfrak{q}$ is both injective and surjective. Consequently, $\exp^\star \mathfrak{q}$ is bijective, i.e., invertible. Since $\exp^\star \mathfrak{q}$ is an invertible $\mathbf{k}$-algebra homomorphism, we conclude that $\exp^\star \mathfrak{q}$ is a $\mathbf{k}$-algebra isomorphism. Its inverse must be $\mathfrak{s}$ (since $\mathfrak{s} \circ (\exp^\star \mathfrak{q}) = \operatorname{id}_A$). Hence, the maps $\exp^\star \mathfrak{q} : A \to \operatorname{Sym}(\mathfrak{e}(A))$ and $\mathfrak{s} : \operatorname{Sym}(\mathfrak{e}(A)) \to A$ are mutually inverse $\mathbf{k}$-algebra isomorphisms. This proves Theorem 1.7.29(d).

(e) Theorem 1.7.29(d) shows that the maps $\exp^\star \mathfrak{q} : A \to \operatorname{Sym}(\mathfrak{e}(A))$ and $\mathfrak{s} : \operatorname{Sym}(\mathfrak{e}(A)) \to A$ are mutually inverse $\mathbf{k}$-algebra isomorphisms. Hence, $A \cong \operatorname{Sym}(\mathfrak{e}(A))$ as $\mathbf{k}$-algebras (via these isomorphisms). But Theorem 1.7.29(b) shows that $\mathfrak{e}(A) \cong (\ker\epsilon) / (\ker\epsilon)^2$ (as $\mathbf{k}$-modules). Hence, $\operatorname{Sym}(\mathfrak{e}(A)) \cong \operatorname{Sym}\left((\ker\epsilon) / (\ker\epsilon)^2\right)$ as $\mathbf{k}$-algebras. Thus, $A \cong \operatorname{Sym}(\mathfrak{e}(A)) \cong \operatorname{Sym}\left((\ker\epsilon) / (\ker\epsilon)^2\right)$ as $\mathbf{k}$-algebras. This proves Theorem 1.7.29(e).

(f) Let $x \in A$. We have

$$\underbrace{\widetilde{\operatorname{id}}}_{=\operatorname{id}_A - u_A\epsilon_A}(x) = (\operatorname{id}_A - u_A\epsilon_A)(x) = \underbrace{\operatorname{id}_A(x)}_{=x} - \underbrace{(u_A\epsilon_A)(x)}_{\substack{=u_A(\epsilon_A(x)) \\ =\epsilon_A(x)\cdot 1_A \\ (\text{by the definition of } u_A)}} = x - \epsilon_A(x) \cdot 1_A,$$

so that $x = \epsilon_A(x) \cdot 1_A + \widetilde{\operatorname{id}}(x)$. Subtracting $\mathfrak{e}(x)$ from both sides of this equality, we obtain

$$x - \mathfrak{e}(x) = \underbrace{\epsilon_A(x)}_{\in\mathbf{k}} \cdot 1_A + \underbrace{\widetilde{\operatorname{id}}(x) - \mathfrak{e}(x)}_{\substack{\in(\ker\epsilon)^2 \\ (\text{by } (13.43.15))}} \in \mathbf{k} \cdot 1_A + (\ker\epsilon)^2 = \ker \mathfrak{e}$$

(by Theorem 1.7.29(b)). In other words, $\mathfrak{e}(x - \mathfrak{e}(x)) = 0$. Comparing this with

$$\mathfrak{e}(x - \mathfrak{e}(x)) = \mathfrak{e}(x) - \underbrace{\mathfrak{e}(\mathfrak{e}(x))}_{=(\mathfrak{e}\circ\mathfrak{e})(x)} \qquad \text{(since the map } \mathfrak{e} \text{ is } \mathbf{k}\text{-linear)}$$
$$= \mathfrak{e}(x) - (\mathfrak{e}\circ\mathfrak{e})(x),$$

we obtain $\mathfrak{e}(x) - (\mathfrak{e}\circ\mathfrak{e})(x) = 0$. In other words, $\mathfrak{e}(x) = (\mathfrak{e}\circ\mathfrak{e})(x)$.

Now, forget that we fixed x. We thus have shown that $\mathfrak{e}(x) = (\mathfrak{e}\circ\mathfrak{e})(x)$ for each $x \in A$. In other words, $\mathfrak{e} = \mathfrak{e}\circ\mathfrak{e}$. In other words, $\mathfrak{e}\circ\mathfrak{e} = \mathfrak{e}$. In other words, the map $\mathfrak{e} : A \to A$ is a projection. This proves Theorem 1.7.29(f). □

13.44. **Solution to Exercise 2.1.2.** *Solution to Exercise 2.1.2.* We have $f \in R(\mathbf{x})$. Thus, f is a formal power series of bounded degree. In other words, $f = \sum_\alpha c_\alpha \mathbf{x}^\alpha$ (with the sum ranging over all weak compositions α) for some elements c_α in $\mathbf{k}$ such that there exists a $d \in \mathbb{N}$ such that every α satisfying $\deg(\mathbf{x}^\alpha) > d$ must satisfy $c_\alpha = 0$. Consider these c_α and this d.

We have written f as the sum $\sum_\alpha c_\alpha \mathbf{x}^\alpha$. Now, substituting $a_1, a_2, \ldots, a_k, 0, 0, \ldots$ for $x_1, x_2, x_3, \ldots$ in f maps all but finitely many of the terms $c_\alpha \mathbf{x}^\alpha$ appearing in this sum to 0. In fact:

- all terms $c_\alpha \mathbf{x}^\alpha$ such that the monomial $\mathbf{x}^\alpha$ contains at least one of the variables $x_{k+1}, x_{k+2}, x_{k+3}, \ldots$ (that is, such that for some integer $i > k$, the i-th entry of α is nonzero) become 0 under our substitution (because the substitution takes each of the variables $x_{k+1}, x_{k+2}, x_{k+3}, \ldots$ to 0);
- all terms $c_\alpha \mathbf{x}^\alpha$ with $\deg(\mathbf{x}^\alpha) > d$ become 0 under our substitution (because we know that $c_\alpha = 0$ for every α satisfying $\deg(\mathbf{x}^\alpha) > d$);
- the remaining terms (that is, the terms $c_\alpha \mathbf{x}^\alpha$ satisfying neither of the preceding two conditions) might not get sent to 0, but there are only finitely many such terms[520].

Hence, our substitution maps all but finitely many of the terms $c_\alpha \mathbf{x}^\alpha$ appearing in the sum $\sum_\alpha c_\alpha \mathbf{x}^\alpha$ to 0. Since $\sum_\alpha c_\alpha \mathbf{x}^\alpha = f$, this rewrites as follows: Our substitution maps all but finitely many of the terms of f to 0. This solves Exercise 2.1.2.

13.45. **Solution to Exercise 2.2.9.** *Solution to Exercise 2.2.9.* Let us first make an auxiliary observation. Namely, let us prove the following lemma:

Lemma 13.45.1. *Let $\nu \in \operatorname{Par}$ and $k \in \mathbb{N}$. Then,*

$$(\nu^t)_1 + (\nu^t)_2 + \cdots + (\nu^t)_k = \sum_{j=1}^{\infty} \min\{\nu_j, k\}. \tag{13.45.1}$$

(Note that the right hand side of (13.45.1) is well-defined, because every sufficiently high j satisfies $\min\{\nu_j, k\} = 0$ (since every sufficiently high j satisfies $\nu_j = 0$)).

Proof of Lemma 13.45.1. This is best seen by double-counting boxes in a Ferrers diagram. The left hand side of (13.45.1) counts the boxes in the first k rows of the Ferrers diagram of ν^t. Since the Ferrers diagram of ν^t is obtained from that of ν by exchanging rows for columns (i.e., reflecting the diagram across its main diagonal), it is clear that this is the same as counting the boxes in the first k columns of the Ferrers diagram of ν. But these latter boxes can also be counted row-by-row: For every $j \in \{1, 2, 3, \ldots\}$, the number of boxes in the first k columns of the Ferrers diagram of ν that lie in row j is $\min\{\nu_j, k\}$. Thus, the total amount of boxes in the first k columns of the Ferrers diagram of ν is $\sum_{j=1}^{\infty} \min\{\nu_j, k\}$, which is precisely the right hand side of (13.45.1). Hence, the left hand side of (13.45.1) and the right hand side of (13.45.1) are equal (since they both count the same boxes). This proves (13.45.1), and thus Lemma 13.45.1 follows. □

[520] In fact, these are the terms for which $\mathbf{x}^\alpha$ is a monomial which has degree $\leq d$ and contains no other variables than $x_1, x_2, \ldots, x_k$. Clearly, there are only finitely many such monomials, and thus only finitely many such terms.

Let us now come to the solution of the exercise. We need to prove that

$$\left(\text{if } \lambda \triangleright \mu \text{, then } \mu^t \triangleright \lambda^t\right) \tag{13.45.2}$$

and

$$\left(\text{if } \mu^t \triangleright \lambda^t \text{, then } \lambda \triangleright \mu\right). \tag{13.45.3}$$

Proof of (13.45.3)*:* Assume that $\mu^t \triangleright \lambda^t$. By the definition of the dominance order, this means that

$$\left(\mu^t\right)_1 + \left(\mu^t\right)_2 + ... + \left(\mu^t\right)_k \geq \left(\lambda^t\right)_1 + \left(\lambda^t\right)_2 + ... + \left(\lambda^t\right)_k \qquad \text{for all } k \in \{1, 2, ..., n\}. \tag{13.45.4}$$

This inequality holds for $k = 0$ as well (because both sides are 0 when $k = 0$), and thus it holds for all $k \in \{0, 1, ..., n\}$. We can further rewrite this inequality by applying (13.45.1): The left hand side becomes $\sum_{j=1}^{\infty} \min\{\mu_j, k\}$, and the right hand side becomes $\sum_{j=1}^{\infty} \min\{\lambda_j, k\}$. As a result, we obtain

$$\sum_{j=1}^{\infty} \min\{\mu_j, k\} \geq \sum_{j=1}^{\infty} \min\{\lambda_j, k\} \qquad \text{for all } k \in \{0, 1, ..., n\}. \tag{13.45.5}$$

Now, let $k \in \{1, 2, ..., n\}$ be arbitrary. We are going to show that $\lambda_1 + \lambda_2 + ... + \lambda_k \geq \mu_1 + \mu_2 + ... + \mu_k$.

First of all, $\lambda \in \operatorname{Par}_n$, so that $n = |\lambda| \geq \lambda_k$. Thus, $\lambda_k \in \{0, 1, ..., n\}$. Therefore, (13.45.5) (applied to λ_k instead of k) yields that

$$\sum_{j=1}^{\infty} \min\{\mu_j, \lambda_k\} \geq \sum_{j=1}^{\infty} \min\{\lambda_j, \lambda_k\}. \tag{13.45.6}$$

But the right hand side of this inequality can be rewritten as follows:

$$\begin{aligned}
\sum_{j=1}^{\infty} \min\{\lambda_j, \lambda_k\} &= \sum_{j=1}^{k} \underbrace{\min\{\lambda_j, \lambda_k\}}_{\substack{=\lambda_k \\ (\text{since } j \leq k \text{ and thus } \lambda_j \geq \lambda_k \\ (\text{because } \lambda \text{ is a partition}))}} + \sum_{j=k+1}^{\infty} \underbrace{\min\{\lambda_j, \lambda_k\}}_{\substack{=\lambda_j \\ (\text{since } j > k \text{ and thus } \lambda_j \leq \lambda_k \\ (\text{because } \lambda \text{ is a partition}))}} \\
&= \underbrace{\sum_{j=1}^{k} \lambda_k}_{=k\lambda_k} + \underbrace{\sum_{j=k+1}^{\infty} \lambda_j}_{\substack{=\lambda_{k+1}+\lambda_{k+2}+\lambda_{k+3}+... \\ =|\lambda|-(\lambda_1+\lambda_2+...+\lambda_k)}} = k\lambda_k + \underbrace{|\lambda|}_{=n} - (\lambda_1 + \lambda_2 + ... + \lambda_k) \\
&= k\lambda_k + n - (\lambda_1 + \lambda_2 + ... + \lambda_k). \qquad (13.45.7)
\end{aligned}$$

Meanwhile, the left hand side can be bounded from above:

$$\begin{aligned}
\sum_{j=1}^{\infty} \min\{\mu_j, \lambda_k\} &= \sum_{j=1}^{k} \underbrace{\min\{\mu_j, \lambda_k\}}_{\leq \lambda_k} + \sum_{j=k+1}^{\infty} \underbrace{\min\{\mu_j, \lambda_k\}}_{\leq \mu_j} \\
&\leq \underbrace{\sum_{j=1}^{k} \lambda_k}_{=k\lambda_k} + \underbrace{\sum_{j=k+1}^{\infty} \mu_j}_{\substack{=\mu_{k+1}+\mu_{k+2}+\mu_{k+3}+... \\ =|\mu|-(\mu_1+\mu_2+...+\mu_k)}} = k\lambda_k + \underbrace{|\mu|}_{=n} - (\mu_1 + \mu_2 + ... + \mu_k) \\
&= k\lambda_k + n - (\mu_1 + \mu_2 + ... + \mu_k). \qquad (13.45.8)
\end{aligned}$$

Now, (13.45.6) yields

$$\begin{aligned}
0 &\leq \underbrace{\sum_{j=1}^{\infty} \min\{\mu_j, \lambda_k\}}_{\substack{\leq k\lambda_k+n-(\mu_1+\mu_2+...+\mu_k) \\ (\text{by } (13.45.8))}} - \underbrace{\sum_{j=1}^{\infty} \min\{\lambda_j, \lambda_k\}}_{\substack{=k\lambda_k+n-(\lambda_1+\lambda_2+...+\lambda_k) \\ (\text{by } (13.45.7))}} \\
&\leq (k\lambda_k + n - (\mu_1 + \mu_2 + ... + \mu_k)) - (k\lambda_k + n - (\lambda_1 + \lambda_2 + ... + \lambda_k)) \\
&= (\lambda_1 + \lambda_2 + ... + \lambda_k) - (\mu_1 + \mu_2 + ... + \mu_k).
\end{aligned}$$

In other words, $\lambda_1 + \lambda_2 + ... + \lambda_k \geq \mu_1 + \mu_2 + ... + \mu_k$.

Now, forget that we fixed k. We thus have shown that

$$\lambda_1 + \lambda_2 + ... + \lambda_k \geq \mu_1 + \mu_2 + ... + \mu_k \qquad \text{for all } k \in \{1, 2, ..., n\}.$$

In other words, $\lambda \triangleright \mu$. This proves (13.45.3).

Proof of (13.45.2)*:* Assume that $\lambda \triangleright \mu$. We have $(\lambda^t)^t = \lambda \triangleright \mu = (\mu^t)^t$. Thus, we can apply (13.45.3) to μ^t and λ^t instead of λ and μ. As a result, we obtain $\mu^t \triangleright \lambda^t$. This proves (13.45.2).

Combining (13.45.2) and (13.45.3), we obtain the equivalence of the two assertions $\lambda \triangleright \mu$ and $\mu^t \triangleright \lambda^t$. This solves Exercise 2.2.9.

13.46. **Solution to Exercise 2.2.13.** *Solution to Exercise 2.2.13.* Let us first introduce some terminology.

Definition 13.46.1. If $\alpha \in \mathbb{N}^\infty$ is any sequence, and if i is any positive integer, then α_i shall denote the i-th entry of α. (This generalizes the notation λ_i for the i-th entry of a partition λ.) Thus, any sequence $\alpha \in \mathbb{N}^\infty$ satisfies $\alpha = (\alpha_1, \alpha_2, \alpha_3, \ldots)$.

Definition 13.46.2. Let WC denote the set of all weak compositions. For every $f \in \mathbf{k}[[\mathbf{x}]]$ and $\mu \in \mathrm{WC}$, we let $[\mathbf{x}^\mu] f$ denote the coefficient of the monomial $\mathbf{x}^\mu$ in the power series f. (This generalizes the notation $[\mathbf{x}^\mu] f$ introduced in Exercise 2.2.13(a).)

We observe the following obvious facts:

- For any $\alpha \in \mathrm{WC}$ and $\mu \in \mathrm{WC}$, we have

$$[\mathbf{x}^\mu](\mathbf{x}^\alpha) = \delta_{\mu,\alpha}. \tag{13.46.1}$$

- If a power series $f \in \mathbf{k}[[\mathbf{x}]]$ is written in the form $f = \sum_{\alpha \in \mathrm{WC}} c_\alpha \mathbf{x}^\alpha$ for some family $(c_\alpha)_{\alpha \in \mathrm{WC}} \in \mathbf{k}^{\mathrm{WC}}$ of scalars, then

$$\text{every } \alpha \in \mathrm{WC} \text{ satisfies } [\mathbf{x}^\alpha] f = c_\alpha. \tag{13.46.2}$$

(a) Let $f \in \Lambda_n$. Thus,

$$f \in \Lambda_n \subset \Lambda = \left\{ \sum_{\alpha \in \mathrm{WC}} c_\alpha \mathbf{x}^\alpha \in R(\mathbf{x}) \mid c_\alpha = c_\beta \text{ if } \alpha, \beta \text{ lie in the same } \mathfrak{S}_{(\infty)}\text{-orbit} \right\}$$

(by the definition of Λ). In other words, f can be written in the form $f = \sum_{\alpha \in \mathrm{WC}} c_\alpha \mathbf{x}^\alpha$ for some family $(c_\alpha)_{\alpha \in \mathrm{WC}} \in \mathbf{k}^{\mathrm{WC}}$ of scalars having the property that

$$\left(c_\alpha = c_\beta \text{ if } \alpha, \beta \text{ lie in the same } \mathfrak{S}_{(\infty)}\text{-orbit}\right). \tag{13.46.3}$$

Consider this family $(c_\alpha)_{\alpha \in \mathrm{WC}}$. From (13.46.2), we conclude that every $\alpha \in \mathrm{WC}$ satisfies

$$[\mathbf{x}^\alpha] f = c_\alpha. \tag{13.46.4}$$

Now, fix $\alpha \in \mathrm{WC}$. Thus, α is a weak composition (since WC is the set of all weak compositions). Hence, there exists a **unique** partition λ that is a permutation of α [521]. Thus, the sum $\sum_{\substack{\lambda \text{ is a partition;} \\ \lambda \text{ is a permutation of } \alpha}} \mathbf{x}^\alpha$ has only one addend. Therefore, this sum simplifies as follows:

$$\sum_{\substack{\lambda \text{ is a partition;} \\ \lambda \text{ is a permutation of } \alpha}} \mathbf{x}^\alpha = \mathbf{x}^\alpha. \tag{13.46.5}$$

[521] Namely, this partition λ is the result of sorting the entries of α into decreasing order (or, more precisely: moving the positive entries of α to the left of all zero entries, and then sorting the former into decreasing order). For example, if $\alpha = (0, 2, 5, 0, 3, 1, 0, 0, 1, 0, 0, \ldots)$, then this partition λ is $(5, 3, 2, 1, 1, 0, 0, \ldots)$.

Thus,

$$\mathbf{x}^\alpha = \underbrace{\sum_{\substack{\lambda \text{ is a partition;} \\ \lambda \text{ is a permutation of } \alpha}}}_{= \sum_{\substack{\lambda \text{ is a partition;} \\ \alpha \text{ is a permutation of } \lambda}} = \sum_{\substack{\lambda \text{ is a partition;} \\ \alpha \in \mathfrak{S}_{(\infty)}\lambda}} = \sum_{\substack{\lambda \in \operatorname{Par}; \\ \alpha \in \mathfrak{S}_{(\infty)}\lambda}}} \mathbf{x}^\alpha = \sum_{\substack{\lambda \in \operatorname{Par}; \\ \alpha \in \mathfrak{S}_{(\infty)}\lambda}} \mathbf{x}^\alpha . \tag{13.46.6}$$

Now, forget that we fixed α. We thus have proven that (13.46.6) holds for each $\alpha \in \operatorname{WC}$.

Recall that each $\lambda \in \operatorname{Par}$ satisfies

$$m_\lambda = \sum_{\substack{\alpha \in \operatorname{WC}; \\ \alpha \in \mathfrak{S}_{(\infty)}\lambda}} \mathbf{x}^\alpha \tag{13.46.7}$$

(by (2.1.1)).

Now,

$$\begin{aligned} f &= \sum_{\alpha \in \operatorname{WC}} c_\alpha \underbrace{\mathbf{x}^\alpha}_{\substack{= \sum_{\substack{\lambda \in \operatorname{Par}; \\ \alpha \in \mathfrak{S}_{(\infty)}\lambda}} \mathbf{x}^\alpha \\ \text{(by (13.46.6))}}} = \sum_{\alpha \in \operatorname{WC}} c_\alpha \sum_{\substack{\lambda \in \operatorname{Par}; \\ \alpha \in \mathfrak{S}_{(\infty)}\lambda}} \mathbf{x}^\alpha \\ &= \underbrace{\sum_{\alpha \in \operatorname{WC}} \sum_{\substack{\lambda \in \operatorname{Par}; \\ \alpha \in \mathfrak{S}_{(\infty)}\lambda}}}_{= \sum_{\lambda \in \operatorname{Par}} \sum_{\substack{\alpha \in \operatorname{WC}; \\ \alpha \in \mathfrak{S}_{(\infty)}\lambda}}} \underbrace{c_\alpha}_{\substack{= c_\lambda \\ \text{(by (13.46.3), applied to } \beta = \lambda \\ \text{(since } \alpha, \lambda \text{ lie in the same } \mathfrak{S}_{(\infty)}\text{-orbit} \\ \text{(since } \alpha \in \mathfrak{S}_{(\infty)}\lambda\text{)))}}} \mathbf{x}^\alpha \\ &= \sum_{\lambda \in \operatorname{Par}} \sum_{\substack{\alpha \in \operatorname{WC}; \\ \alpha \in \mathfrak{S}_{(\infty)}\lambda}} c_\lambda \mathbf{x}^\alpha = \sum_{\lambda \in \operatorname{Par}} c_\lambda \underbrace{\sum_{\substack{\alpha \in \operatorname{WC}; \\ \alpha \in \mathfrak{S}_{(\infty)}\lambda}} \mathbf{x}^\alpha}_{\substack{= m_\lambda \\ \text{(by (13.46.7))}}} = \sum_{\lambda \in \operatorname{Par}} c_\lambda m_\lambda . \end{aligned} \tag{13.46.8}$$

Now, let $\pi_n : \mathbf{k}[[\mathbf{x}]] \to \mathbf{k}[[\mathbf{x}]]$ be the projection that maps each power series $g \in \mathbf{k}[[\mathbf{x}]]$ to its n-th homogeneous component. Then, the definition of π_n shows that the following holds:

- If $g \in \mathbf{k}[[\mathbf{x}]]$ is a homogeneous power series of degree n, then

$$\pi_n(g) = g. \tag{13.46.9}$$

- If $g \in \mathbf{k}[[\mathbf{x}]]$ is a homogeneous power series of degree $\neq n$, then

$$\pi_n(g) = 0. \tag{13.46.10}$$

We can now use this to compute $\pi_n(m_\lambda)$ for each $\lambda \in \operatorname{Par}$.

First, let us notice that

$$m_\lambda \text{ is a homogeneous power series of degree } |\lambda| \tag{13.46.11}$$

for each $\lambda \in \operatorname{Par}$. (This is clear, because each of the addends $\mathbf{x}^\alpha$ on the right hand side of (13.46.7) is a monomial of degree $|\alpha| = |\lambda|$.)

Now, we conclude the following:

- If $\lambda \in \operatorname{Par}$ satisfies $|\lambda| = n$, then

$$\pi_n(m_\lambda) = m_\lambda \tag{13.46.12}$$

[522].

[522] *Proof of (13.46.12):* Let $\lambda \in \operatorname{Par}$ be such that $|\lambda| = n$. Then, (13.46.11) shows that m_λ is a homogeneous power series of degree $|\lambda| = n$. Hence, (13.46.9) (applied to $g = m_\lambda$) shows that $\pi_n(m_\lambda) = m_\lambda$. This proves (13.46.12).

- If $\lambda \in \operatorname{Par}$ satisfies $|\lambda| \neq n$, then

$$\pi_n\left(m_\lambda\right) = 0 \tag{13.46.13}$$

[523].

Finally, let us apply the map π_n to both sides of the equality (13.46.8). We thus obtain

$$\begin{aligned}
\pi_n(f) &= \pi_n\left(\sum_{\lambda\in\operatorname{Par}} c_\lambda m_\lambda\right) = \sum_{\lambda\in\operatorname{Par}} c_\lambda \pi_n\left(m_\lambda\right) \\
&\qquad \left(\begin{array}{c}\text{since the map } \pi_n \text{ respects infinite } \mathbf{k}\text{-linear combinations} \\ \text{(because } \pi_n \text{ is } \mathbf{k}\text{-linear and continuous)}\end{array}\right) \\
&= \underbrace{\sum_{\substack{\lambda\in\operatorname{Par};\\ |\lambda|=n}}}_{\substack{=\sum_{\lambda\in\operatorname{Par}_n} \\ \text{(since } \{\lambda\in\operatorname{Par} \ \mid\ |\lambda|=n\}=\operatorname{Par}_n \text{)}}} c_\lambda \underbrace{\pi_n\left(m_\lambda\right)}_{\substack{=m_\lambda \\ \text{(by (13.46.12))}}} + \sum_{\substack{\lambda\in\operatorname{Par};\\ |\lambda|\neq n}} c_\lambda \underbrace{\pi_n\left(m_\lambda\right)}_{\substack{=0 \\ \text{(by (13.46.13))}}} \\
&= \sum_{\lambda\in\operatorname{Par}_n} c_\lambda m_\lambda + \underbrace{\sum_{\substack{\lambda\in\operatorname{Par};\\ |\lambda|\neq n}} c_\lambda 0}_{=0} = \sum_{\lambda\in\operatorname{Par}_n} c_\lambda m_\lambda .
\end{aligned}$$

But f is a homogeneous power series of degree n (since $f \in \Lambda_n$). Thus, (13.46.9) (applied to $g = f$) yields $\pi_n(f) = f$. Hence,

$$f = \pi_n(f) = \sum_{\lambda\in\operatorname{Par}_n} c_\lambda m_\lambda = \sum_{\mu\in\operatorname{Par}_n} c_\mu m_\mu$$

(here, we have renamed the summation index λ as μ). Comparing this with

$$\sum_{\mu\in\operatorname{Par}_n} \underbrace{\left(\left[\mathbf{x}^\mu\right] f\right)}_{\substack{=c_\mu \\ \text{(by (13.46.4) (applied to } \alpha=\mu\text{))}}} m_\mu = \sum_{\mu\in\operatorname{Par}_n} c_\mu m_\mu ,$$

we obtain $f = \sum_{\mu\in\operatorname{Par}_n} \left(\left[\mathbf{x}^\mu\right] f\right) m_\mu$. This solves Exercise 2.2.13(a).

(b) Let λ be a partition. Let μ be a weak composition. We must prove that the number $K_{\lambda,\mu}$ is well-defined. In other words, we must prove that there are only finitely many column-strict tableaux T of shape λ having $\operatorname{cont}(T) = \mu$.

Let F be the Ferrers diagram of λ (as a set of cells). Thus, F is a finite subset of $\{1, 2, 3, \ldots\}^2$.

Any column-strict tableau of shape λ is an assignment of entries in $\{1, 2, 3, \ldots\}$ to the cells of the Ferrers diagram of λ. In other words, any column-strict tableau of shape λ is map $F \to \{1, 2, 3, \ldots\}$ (since F is the set of all cells of the Ferrers diagram of λ).

The sequence μ is a weak composition, and thus has a finite support. In other words, the support of μ is finite. Let W be this support. Thus, W is a finite set, and satisfies

$$\begin{aligned}
W &= (\text{the support of } \mu) \\
&= (\text{the set of all positive integers } i \text{ for which } \mu_i \neq 0)
\end{aligned} \tag{13.46.14}$$

(by the definition of the support of μ). Thus, $W \subset \{1, 2, 3, \ldots\}$.

It is well-known that if X, Y and Z are three sets such that $X \subset Y$, then

$$X^Z \cong \left\{f \in Y^Z \ \mid\ f(Z) \subset X\right\} \qquad \text{as sets.}$$

Applying this to $X = W$, $Y = \{1, 2, 3, \ldots\}$ and $Z = F$, we conclude that

$$W^F \cong \left\{f \in \{1, 2, 3, \ldots\}^F \ \mid\ f(F) \subset W\right\} \qquad \text{as sets.} \tag{13.46.15}$$

But F and W are finite sets. Hence, W^F is also a finite set. Thus, $\left\{f \in \{1, 2, 3, \ldots\}^F \ \mid\ f(F) \subset W\right\}$ is a finite set (because of (13.46.15)).

[523] *Proof of (13.46.13):* Let $\lambda \in \operatorname{Par}$ be such that $|\lambda| \neq n$. Then, (13.46.11) shows that m_λ is a homogeneous power series of degree $|\lambda| \neq n$. Hence, (13.46.10) (applied to $g = m_\lambda$) shows that $\pi_n(m_\lambda) = 0$. This proves (13.46.13).

Now, let T be a column-strict tableau of shape λ having $\operatorname{cont}(T) = \mu$. We shall prove that $T \in \left\{f \in \{1,2,3,\ldots\}^F \ \mid\ f(F) \subset W\right\}$.

Indeed, T is a map $F \to \{1,2,3,\ldots\}$ (since any column-strict tableau of shape λ is a map $F \to \{1,2,3,\ldots\}$), hence an element of $\{1,2,3,\ldots\}^F$.

Furthermore, $T(F) \subset W$ [524]. Hence, T is an $f \in \{1,2,3,\ldots\}^F$ satisfying $f(F) \subset W$ (since T is an element of $\{1,2,3,\ldots\}^F$ and satisfies $T(F) \subset W$). In other words,

$$T \in \left\{f \in \{1,2,3,\ldots\}^F \ \mid\ f(F) \subset W\right\}.$$

Now, forget that we fixed T. We thus have shown that every column-strict tableau T of shape λ having $\operatorname{cont}(T) = \mu$ satisfies $T \in \left\{f \in \{1,2,3,\ldots\}^F \ \mid\ f(F) \subset W\right\}$. In other words,

$$\{\text{column-strict tableaux } T \text{ of shape } \lambda \text{ having } \operatorname{cont}(T) = \mu\} \subset \left\{f \in \{1,2,3,\ldots\}^F \ \mid\ f(F) \subset W\right\}.$$

Hence, $\{\text{column-strict tableaux } T \text{ of shape } \lambda \text{ having } \operatorname{cont}(T) = \mu\}$ is a finite set (since $\left\{f \in \{1,2,3,\ldots\}^F \ \mid\ f(F) \subset W\right\}$ is a finite set). In other words, there are only finitely many column-strict tableaux T of shape λ having $\operatorname{cont}(T) = \mu$. This solves Exercise 2.2.13(b).

(c) Let $\lambda \in \operatorname{Par}_n$. The definition of s_λ yields

$$s_\lambda = \sum_T \mathbf{x}^{\operatorname{cont}(T)},$$

where T runs through all column-strict tableaux of shape λ. In other words,

$$s_\lambda = \sum_{\substack{T \text{ is a column-strict} \\ \text{tableau of shape } \lambda}} \mathbf{x}^{\operatorname{cont}(T)}.$$

[524] *Proof.* Let $j \in T(F)$. Then, $j \in \{1,2,3,\ldots\}$. Furthermore, there exists some $c \in F$ satisfying $j = T(c)$ (since $j \in T(F)$). Consider this c. From $j = T(c)$, we obtain $c \in T^{-1}(j)$. Hence, the set $T^{-1}(j)$ contains at least one element (namely, c). Therefore, $\left|T^{-1}(j)\right| \geq 1$.

But $\mu = \operatorname{cont}(T) = \left(\left|T^{-1}(1)\right|, \left|T^{-1}(2)\right|, \left|T^{-1}(3)\right|, \ldots\right)$ (by the definition of $\operatorname{cont}(T)$). Hence, $\mu_j = \left|T^{-1}(j)\right| \geq 1 > 0$. Hence, $\mu_j \neq 0$. Thus, j belongs to the set of all positive integers i for which $\mu_i \neq 0$. In other words, $j \in (\text{the set of all positive integers } i \text{ for which } \mu_i \neq 0)$. In light of (13.46.14), this rewrites as $j \in W$.

Now, forget that we fixed j. We thus have shown that $j \in W$ for each $j \in T(F)$. In other words, $T(F) \subset W$.

Now, every $\mu \in \mathrm{Par}_n$ satisfies

$$\begin{aligned}
&[\mathbf{x}^\mu]\left(\underbrace{s_\lambda}_{=\sum_{\substack{T \text{ is a column-strict}\\ \text{tableau of shape } \lambda}} \mathbf{x}^{\mathrm{cont}(T)}}\right) \\
&= [\mathbf{x}^\mu]\left(\sum_{\substack{T \text{ is a column-strict}\\ \text{tableau of shape } \lambda}} \mathbf{x}^{\mathrm{cont}(T)}\right) = \sum_{\substack{T \text{ is a column-strict}\\ \text{tableau of shape } \lambda}} \underbrace{[\mathbf{x}^\mu]\left(\mathbf{x}^{\mathrm{cont}(T)}\right)}_{\substack{=\delta_{\mu,\mathrm{cont}(T)} \\ \text{(by (13.46.1) (applied to } \alpha=\mathrm{cont}(T)))}} \\
&= \sum_{\substack{T \text{ is a column-strict}\\ \text{tableau of shape } \lambda}} \delta_{\mu,\mathrm{cont}(T)} = \sum_{\substack{T \text{ is a column-strict}\\ \text{tableau of shape } \lambda;\\ \mu=\mathrm{cont}(T)}} \underbrace{\delta_{\mu,\mathrm{cont}(T)}}_{\substack{=1\\ \text{(since } \mu=\mathrm{cont}(T))}} + \sum_{\substack{T \text{ is a column-strict}\\ \text{tableau of shape } \lambda;\\ \mu\neq\mathrm{cont}(T)}} \underbrace{\delta_{\mu,\mathrm{cont}(T)}}_{\substack{=0\\ \text{(since } \mu\neq\mathrm{cont}(T))}} \\
&= \sum_{\substack{T \text{ is a column-strict}\\ \text{tableau of shape } \lambda;\\ \mu=\mathrm{cont}(T)}} 1 + \underbrace{\sum_{\substack{T \text{ is a column-strict}\\ \text{tableau of shape } \lambda;\\ \mu\neq\mathrm{cont}(T)}} 0}_{=0} = \sum_{\substack{T \text{ is a column-strict}\\ \text{tableau of shape } \lambda;\\ \mu=\mathrm{cont}(T)}} 1 \\
&= (\text{the number of all column-strict tableaux } T \text{ of shape } \lambda \text{ having } \mu = \mathrm{cont}(T)) \cdot 1 \\
&= (\text{the number of all column-strict tableaux } T \text{ of shape } \lambda \text{ having } \mu = \mathrm{cont}(T)) \\
&= (\text{the number of all column-strict tableaux } T \text{ of shape } \lambda \text{ having } \mathrm{cont}(T) = \mu) \\
&= K_{\lambda,\mu} \qquad (13.46.16)
\end{aligned}$$

(since $K_{\lambda,\mu}$ is defined as the number of all column-strict tableaux T of shape λ having $\mathrm{cont}(T) = \mu$).

But $\lambda \in \mathrm{Par}_n$. Hence, $|\lambda| = n$. Recall that the power series s_λ is a symmetric function (by Proposition 2.2.4), and is homogeneous of degree $|\lambda|$. Thus, $s_\lambda \in \Lambda_{|\lambda|} = \Lambda_n$ (since $|\lambda| = n$). Hence, Exercise 2.2.13(a) (applied to $f = s_\lambda$) yields

$$s_\lambda = \sum_{\mu \in \mathrm{Par}_n} \underbrace{([\mathbf{x}^\mu](s_\lambda))}_{\substack{=K_{\lambda,\mu}\\ \text{(by (13.46.16))}}} m_\mu = \sum_{\mu\in\mathrm{Par}_n} K_{\lambda,\mu} m_\mu.$$

This solves Exercise 2.2.13(c).

Before we solve Exercise 2.2.13(d), let us show a lemma:

Lemma 13.46.3. *Let $n \in \mathbb{N}$. Let $\lambda \in \mathrm{Par}_n$ and $\mu \in \mathrm{Par}_n$. Let T be a column-strict tableau of shape λ satisfying $\mathrm{cont}(T) = \mu$. Let F be the Ferrers diagram of λ (as a set of cells). For each positive integer i, we let F_i be the i-th row of F (that is, the set of all cells of F that have the form (i,j) for some $j \geq 1$).*

(a) *We have $T(i,j) \geq i$ for each $(i,j) \in F$.*
(b) *We have $T^{-1}(1) \cup T^{-1}(2) \cup \cdots \cup T^{-1}(k) \subset F_1 \cup F_2 \cup \cdots \cup F_k$ for each $k \in \mathbb{N}$.*
(c) *We have $|F_1 \cup F_2 \cup \cdots \cup F_k| = \lambda_1 + \lambda_2 + \cdots + \lambda_k$ for each $k \in \mathbb{N}$.*
(d) *We have $\left|T^{-1}(1) \cup T^{-1}(2) \cup \cdots \cup T^{-1}(k)\right| = \mu_1 + \mu_2 + \cdots + \mu_k$ for each $k \in \mathbb{N}$.*
(e) *We have $\lambda \triangleright \mu$.*
(f) *If $\mu = \lambda$, then each $(i,j) \in F$ satisfies $T(i,j) = i$.*

Proof of Lemma 13.46.3. Recall that F is the Ferrers diagram of λ. In other words, F is the set of all pairs $(i,j) \in \{1,2,3,\ldots\}^2$ satisfying $j \leq \lambda_i$ (by the definition of a Ferrers diagram). In other words,

$$F = \left\{(i,j) \in \{1,2,3,\ldots\}^2 \mid j \leq \lambda_i\right\}.$$

Hence, every $(p,q) \in F$ satisfies

$$(p',q) \in F \qquad \text{for each } p' \in \{1,2,\ldots,p\} \tag{13.46.17}$$

[525].

Any column-strict tableau of shape λ is an assignment of entries in $\{1,2,3,\ldots\}$ to the cells of the Ferrers diagram of λ. In other words, any column-strict tableau of shape λ is a map $F \to \{1,2,3,\ldots\}$ (since F is the Ferrers diagram of λ). Hence, T is a map $F \to \{1,2,3,\ldots\}$ (since T is a column-strict tableau of shape λ).

Recall that T is a column-strict tableau. Thus, the entries of T are strictly increasing top-to-bottom down columns (by the definition of a column-strict tableau).

(a) Let $(i,j) \in F$. We must prove that $T(i,j) \geq i$.

Assume the contrary. Thus, $T(i,j) < i$, so that $T(i,j) \leq i-1$ (since $T(i,j)$ and i are integers).

We have $(p',j) \in F$ for each $p' \in \{1,2,\ldots,i\}$ (by (13.46.17) (applied to $(p,q) = (i,j)$)). In other words, all of the cells $(1,j),(2,j),\ldots,(i,j)$ belong to F. These cells therefore all lie in the j-th column of F; more precisely, they are the first i cells of the j-th column of F. Hence, we have

$$T(1,j) < T(2,j) < \cdots < T(i,j) \qquad (13.46.18)$$

(since the entries of T are strictly increasing top-to-bottom down columns). Hence, $T(1,j) < T(2,j) < \cdots < T(i,j) \leq i-1$. Thus, all of the i numbers $T(1,j),T(2,j),\ldots,T(i,j)$ are elements of the set $\{1,2,\ldots,i-1\}$ (since these numbers all belong to $\{1,2,3,\ldots\}$ and are $\leq i-1$). By the pigeonhole principle, we thus conclude that two of these i numbers are equal (since the set $\{1,2,\ldots,i-1\}$ has size $i-1 < i$). This contradicts the fact that these i numbers are distinct (because of (13.46.18)).This contradiction proves that our assumption was wrong; hence, we must have $T(i,j) \geq i$. This proves Lemma 13.46.3(a).

(b) Let $k \in \mathbb{N}$. Each $p \in \{1,2,\ldots,k\}$ satisfies

$$T^{-1}(p) \subset F_1 \cup F_2 \cup \cdots \cup F_k. \qquad (13.46.19)$$

[*Proof of (13.46.19):* Let $p \in \{1,2,\ldots,k\}$. Let $c \in T^{-1}(p)$. Then, $c \in F$ (since T is a map $F \to \{1,2,3,\ldots\}$). Also, $T(c) = p$ (since $c \in T^{-1}(p)$).

Write the cell c in the form $c = (i,j)$ for some $(i,j) \in \{1,2,3,\ldots\}^2$. Then, $(i,j) = c \in F$ and $T\left(\underbrace{c}_{=(i,j)}\right) =$ $T(i,j)$, so that $T(i,j) = T(c) = p$. Hence, $p = T(i,j) \geq i$ (by Lemma 13.46.3(a)). Hence, $i \leq p \leq k$ (since $p \in \{1,2,\ldots,k\}$), so that $i \in \{1,2,\ldots,k\}$. Therefore, $F_i \subset F_1 \cup F_2 \cup \cdots \cup F_k$.

But the cell (i,j) belongs to the i-th row (since its first coordinate is i) and also belongs to F (since $(i,j) \in F$). Hence, the cell (i,j) belongs to F_i (since F_i is the i-th row of F). In other words, $(i,j) \in F_i$. Hence, $c = (i,j) \in F_i \subset F_1 \cup F_2 \cup \cdots \cup F_k$.

Now, forget that we fixed c. We thus have proven that $c \in F_1 \cup F_2 \cup \cdots \cup F_k$ for each $c \in T^{-1}(p)$. In other words, $T^{-1}(p) \subset F_1 \cup F_2 \cup \cdots \cup F_k$. This proves (13.46.19).]

Now,

$$\begin{aligned} T^{-1}(1) \cup T^{-1}(2) \cup \cdots \cup T^{-1}(k) &= \bigcup_{p \in \{1,2,\ldots,k\}} \underbrace{T^{-1}(p)}_{\substack{\subset F_1 \cup F_2 \cup \cdots \cup F_k \\ (\text{by } (13.46.19))}} \\ &\subset \bigcup_{p \in \{1,2,\ldots,k\}} (F_1 \cup F_2 \cup \cdots \cup F_k) \subset F_1 \cup F_2 \cup \cdots \cup F_k. \end{aligned}$$

This proves Lemma 13.46.3(b).

[525] *Proof of (13.46.17):* Let $(p,q) \in F$. Thus, $(p,q) \in F = \left\{(i,j) \in \{1,2,3,\ldots\}^2 \mid j \leq \lambda_i\right\}$. In other words, (p,q) is an element of $\{1,2,3,\ldots\}^2$ and satisfies $q \leq \lambda_p$.

Now, let $p' \in \{1,2,\ldots,p\}$. We must prove that $(p',q) \in F$.

From $p' \in \{1,2,\ldots,p\} \subset \{1,2,3,\ldots\}$ and $q \in \{1,2,3,\ldots\}$, we conclude that (p',q) is an element of $\{1,2,3,\ldots\}^2$. Also, $p' \leq p$ (since $p' \in \{1,2,\ldots,p\}$).

But λ is a partition. Hence, $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \cdots$. Therefore, from $p' \leq p$, we obtain $\lambda_{p'} \geq \lambda_p$. Hence, $q \leq \lambda_p \leq \lambda_{p'}$ (since $\lambda_{p'} \geq \lambda_p$).

Now, (p',q) is an element of $\{1,2,3,\ldots\}^2$ and satisfies $q \leq \lambda_{p'}$. In other words, $(p',q) \in \left\{(i,j) \in \{1,2,3,\ldots\}^2 \mid j \leq \lambda_i\right\}$. In other words, $(p',q) \in F$ (since $F = \left\{(i,j) \in \{1,2,3,\ldots\}^2 \mid j \leq \lambda_i\right\}$). This completes the proof of (13.46.17).

(c) Every positive integer i satisfies

$$\begin{aligned}|F_i| &= (\text{the size of } F_i) = (\text{the size of the } i\text{-th row of } F) \\ &\qquad (\text{since } F_i \text{ is the } i\text{-th row of } F) \\ &= (\text{the size of the } i\text{-th row of the Ferrers diagram of } \lambda) \\ &\qquad (\text{since } F \text{ is the Ferrers diagram of } \lambda) \\ &= \lambda_i. \qquad (13.46.20)\end{aligned}$$

Let $k \in \mathbb{N}$. The sets $F_1, F_2, \ldots, F_k$ are disjoint (since they are different rows of F). Hence, the size of their union equals the sum of their sizes. In other words, $|F_1 \cup F_2 \cup \cdots \cup F_k| = |F_1| + |F_2| + \cdots + |F_k|$. Thus,

$$|F_1 \cup F_2 \cup \cdots \cup F_k| = |F_1| + |F_2| + \cdots + |F_k| = \sum_{i=1}^{k} \underbrace{|F_i|}_{\substack{=\lambda_i \\ (\text{by } (13.46.20))}} = \sum_{i=1}^{k} \lambda_i$$
$$= \lambda_1 + \lambda_2 + \cdots + \lambda_k.$$

This proves Lemma 13.46.3(c).

(d) We have $\mu = \operatorname{cont}(T) = \left(\left|T^{-1}(1)\right|, \left|T^{-1}(2)\right|, \left|T^{-1}(3)\right|, \ldots\right)$ (by the definition of $\operatorname{cont}(T)$). Hence,

$$\mu_i = \left|T^{-1}(i)\right| \qquad (13.46.21)$$

for every positive integer i.

Let $k \in \mathbb{N}$. The sets $T^{-1}(1), T^{-1}(2), \ldots, T^{-1}(k)$ are disjoint (since they are different fibers of the map T). Hence, the size of their union equals the sum of their sizes. In other words, $\left|T^{-1}(1) \cup T^{-1}(2) \cup \cdots \cup T^{-1}(k)\right| = \left|T^{-1}(1)\right| + \left|T^{-1}(2)\right| + \cdots + \left|T^{-1}(k)\right|$. Thus,

$$\left|T^{-1}(1) \cup T^{-1}(2) \cup \cdots \cup T^{-1}(k)\right| = \left|T^{-1}(1)\right| + \left|T^{-1}(2)\right| + \cdots + \left|T^{-1}(k)\right| = \sum_{i=1}^{k} \underbrace{\left|T^{-1}(i)\right|}_{\substack{=\mu_i \\ (\text{by } (13.46.21))}} = \sum_{i=1}^{k} \mu_i$$
$$= \mu_1 + \mu_2 + \cdots + \mu_k.$$

This proves Lemma 13.46.3(d).

(e) Let $k \in \{1, 2, \ldots, n\}$. Then, Lemma 13.46.3(b) yields $T^{-1}(1) \cup T^{-1}(2) \cup \cdots \cup T^{-1}(k) \subset F_1 \cup F_2 \cup \cdots \cup F_k$. Thus,

$$\left|T^{-1}(1) \cup T^{-1}(2) \cup \cdots \cup T^{-1}(k)\right| \leq |F_1 \cup F_2 \cup \cdots \cup F_k|.$$

But Lemma 13.46.3(d) yields $\left|T^{-1}(1) \cup T^{-1}(2) \cup \cdots \cup T^{-1}(k)\right| = \mu_1 + \mu_2 + \cdots + \mu_k$. Hence,

$$\begin{aligned}\mu_1 + \mu_2 + \cdots + \mu_k &= \left|T^{-1}(1) \cup T^{-1}(2) \cup \cdots \cup T^{-1}(k)\right| \\ &\leq |F_1 \cup F_2 \cup \cdots \cup F_k| = \lambda_1 + \lambda_2 + \cdots + \lambda_k\end{aligned}$$

(by Lemma 13.46.3(c)). In other words, $\lambda_1 + \lambda_2 + \cdots + \lambda_k \geq \mu_1 + \mu_2 + \cdots + \mu_k$.

Now, forget that we fixed k. We thus have shown that

$$\lambda_1 + \lambda_2 + \cdots + \lambda_k \geq \mu_1 + \mu_2 + \cdots + \mu_k \qquad \text{for each } k \in \{1, 2, \ldots, n\}. \qquad (13.46.22)$$

In other words, $\lambda \triangleright \mu$ (by the definition of the dominance order). This proves Lemma 13.46.3(e).

(f) Assume that $\mu = \lambda$. We must prove that each $(i, j) \in F$ satisfies $T(i, j) = i$.

Indeed, let $(i, j) \in F$ be arbitrary. We must prove that $T(i, j) = i$.

Define two sets A and B by $A = T^{-1}(1) \cup T^{-1}(2) \cup \cdots \cup T^{-1}(i)$ and $B = F_1 \cup F_2 \cup \cdots \cup F_i$. Then,

$$\begin{aligned}A &= T^{-1}(1) \cup T^{-1}(2) \cup \cdots \cup T^{-1}(i) \\ &\subset F_1 \cup F_2 \cup \cdots \cup F_i \qquad (\text{by Lemma 13.46.3(b) (applied to } k = i)) \\ &= B.\end{aligned}$$

In other words, A is a subset of B.

Applying Lemma 13.46.3(c) to $k = i$, we find

$$|F_1 \cup F_2 \cup \cdots \cup F_i| = \lambda_1 + \lambda_2 + \cdots + \lambda_i. \qquad (13.46.23)$$

Moreover, from $A = T^{-1}(1) \cup T^{-1}(2) \cup \cdots \cup T^{-1}(i)$, we obtain

$$\begin{aligned}
|A| &= \left|T^{-1}(1) \cup T^{-1}(2) \cup \cdots \cup T^{-1}(i)\right| = \mu_1 + \mu_2 + \cdots + \mu_i \\
&\qquad \left(\text{by Lemma 13.46.3(d) (applied to } k = i)\right) \\
&= \lambda_1 + \lambda_2 + \cdots + \lambda_i \qquad (\text{since } \mu = \lambda) \\
&= \left|\underbrace{F_1 \cup F_2 \cup \cdots \cup F_i}_{=B}\right| \qquad (\text{by (13.46.23)}) \\
&= |B|.
\end{aligned}$$

Moreover, B is a finite set (since $|B| = \lambda_1 + \lambda_2 + \cdots + \lambda_i \in \mathbb{N}$).

It is well-known that if Y is a finite set, and if X is a subset of Y satisfying $|X| = |Y|$, then $X = Y$. Applying this to $X = A$ and $Y = B$, we obtain $A = B$ (since A is a subset of B and satisfies $|A| = |B|$).

Now, the cell (i,j) belongs to the i-th row of F (since it belongs to F, and since its first coordinate is i). In other words, $(i,j) \in F_i$ (since F_i is the i-th row of F). Hence,

$$\begin{aligned}
(i,j) &\in F_i \subset F_1 \cup F_2 \cup \cdots \cup F_i = B = A \qquad (\text{since } A = B) \\
&= T^{-1}(1) \cup T^{-1}(2) \cup \cdots \cup T^{-1}(i).
\end{aligned}$$

In other words, $(i,j) \in T^{-1}(p)$ for some $p \in \{1, 2, \ldots, i\}$. Consider this p.

From $(i,j) \in T^{-1}(p)$, we obtain $T(i,j) = p$. Thus, $T(i,j) = p \leq i$ (since $p \in \{1, 2, \ldots, i\}$). But Lemma 13.46.3(a) yields $T(i,j) \geq i$. Combining this with $T(i,j) \leq i$, we obtain $T(i,j) = i$.

Now, forget that we fixed (i,j). We thus have proven that each $(i,j) \in F$ satisfies $T(i,j) = i$. This proves Lemma 13.46.3(f). □

Now, let us resume the solution of Exercise 2.2.13.

(d) Let $\lambda \in \operatorname{Par}_n$ and $\mu \in \operatorname{Par}_n$ be two partitions that don't satisfy $\lambda \triangleright \mu$. We must prove that $K_{\lambda,\mu} = 0$.

Indeed, there exists no column-strict tableau T of shape λ having $\operatorname{cont}(T) = \mu$ [526]. Thus, the number of all column-strict tableaux T of shape λ having $\operatorname{cont}(T) = \mu$ equals 0. In other words, $K_{\lambda,\mu}$ equals 0 (since $K_{\lambda,\mu}$ is the number of all column-strict tableaux T of shape λ having $\operatorname{cont}(T) = \mu$). This solves Exercise 2.2.13(d).

Before we solve Exercise 2.2.13(e), let us state another simple lemma:

Lemma 13.46.4. *Let $n \in \mathbb{N}$. Let $\lambda \in \operatorname{Par}_n$. Let F be the Ferrers diagram of λ (as a set of cells). Let $T_0 : F \to \{1, 2, 3, \ldots\}$ be the map that sends each $(i,j) \in F$ to i. Then, T_0 is a column-strict tableau of shape λ and satisfies $\operatorname{cont}(T_0) = \lambda$.*

Example 13.46.5. Let $n = 8$ and $\lambda = (3, 2, 2, 1) \in \operatorname{Par}_8$. Then, the column-strict tableau T_0 defined in Lemma 13.46.4 looks as follows:

$$T_0 = \begin{array}{ccc} 1 & 1 & 1 \\ 2 & 2 & \\ 3 & 3 & \\ 4 & & \end{array}.$$

Proof of Lemma 13.46.4. Every $(i,j) \in F$ satisfies $i \in \{1, 2, 3, \ldots\}$ (since $F \subset \{1, 2, 3, \ldots\}^2$). Hence, the map T_0 is well-defined.

The map T_0 is a map from F to $\{1, 2, 3, \ldots\}$. In other words, T_0 is an assignment of entries in $\{1, 2, 3, \ldots\}$ to the cells of the Ferrers diagram of λ (since F is the Ferrers diagram of λ). We shall now show that T_0 is a column-strict tableau of shape λ.

The definition of T_0 shows that for each $i \in \{1, 2, 3, \ldots\}$, all entries in the i-th row of T_0 equal i. Therefore, the entries of T_0 are weakly increasing left-to-right in rows (because they are all equal in a given row) and strictly increasing top-to-bottom in columns (since the topmost entry is a 1, the next entry is a 2, and so

[526] *Proof.* Assume the contrary. Thus, there exists a column-strict tableau T of shape λ having $\operatorname{cont}(T) = \mu$. Consider this T.

Lemma 13.46.3(e) yields $\lambda \triangleright \mu$. This contradicts the fact that we don't have $\lambda \triangleright \mu$. This contradiction proves that our assumption was wrong. Qed.

on). Thus, T_0 is a column-strict tableau of shape λ (by the definition of a "column-strict tableau of shape λ").

It remains to prove that $\operatorname{cont}(T_0) = \lambda$.

The definition of $\operatorname{cont}(T_0)$ shows that $\operatorname{cont}(T_0) = \left(\left|(T_0)^{-1}(1)\right|, \left|(T_0)^{-1}(2)\right|, \left|(T_0)^{-1}(3)\right|, \ldots\right)$. In other words,

$$(\operatorname{cont}(T_0))_i = \left|(T_0)^{-1}(i)\right| \qquad \text{for every positive integer } i. \tag{13.46.24}$$

Let k be a positive integer. Then,

$$\begin{aligned}(T_0)^{-1}(k) &= \left\{(i,j) \in F \mid \underbrace{T_0(i,j)}_{\substack{=i \\ \text{(by the definition of } T_0)}} = k\right\} = \left\{(i,j) \in F \mid \underbrace{i = k}_{\iff ((i,j) \text{ lies in the } k\text{-th row})}\right\} \\ &= \{(i,j) \in F \mid (i,j) \text{ lies in the } k\text{-th row}\} = (\text{the set of all cells of } F \text{ that lie in the } k\text{-th row}) \\ &= (\text{the } k\text{-th row of } F).\end{aligned}$$

Hence,

$$\left|(T_0)^{-1}(k)\right| = |(\text{the } k\text{-th row of } F)| = (\text{the size of the } k\text{-th row of } F) = \lambda_k$$

(since F is the Ferrers diagram of λ). But now, (13.46.24) (applied to $i = k$) yields $(\operatorname{cont}(T_0))_k = \left|(T_0)^{-1}(k)\right| = \lambda_k$.

Now, forget that we fixed k. We thus have proven that $(\operatorname{cont}(T_0))_k = \lambda_k$ for each positive integer k. In other words, $\operatorname{cont}(T_0) = \lambda$. This completes the proof of Lemma 13.46.4. □

Now, let us resume the solution of Exercise 2.2.13.

(e) Let $\lambda \in \operatorname{Par}_n$. We must prove that $K_{\lambda,\lambda} = 1$.

Define F and T_0 as in Lemma 13.46.4. Then, Lemma 13.46.4 shows that T_0 is a column-strict tableau of shape λ and satisfies $\operatorname{cont}(T_0) = \lambda$. Hence, there exists **at least one** column-strict tableau T of shape λ having $\operatorname{cont}(T) = \lambda$ (namely, $T = T_0$).

On the other hand, using Lemma 13.46.3(f), it is easy to see that every column-strict tableau T of shape λ having $\operatorname{cont}(T) = \lambda$ must be equal to T_0 [527]. Hence, there exists **at most one** column-strict tableau T of shape λ having $\operatorname{cont}(T) = \lambda$.

We know that $K_{\lambda,\lambda}$ is the number of all column-strict tableaux T of shape λ having $\operatorname{cont}(T) = \lambda$ (by the definition of $K_{\lambda,\lambda}$). Since there exists **exactly one** such tableau T (because we have shown that there exists **at least one** such tableau T, and we have also shown that there exists **at most one** such tableau T), we thus conclude that $K_{\lambda,\lambda} = 1$. This solves Exercise 2.2.13(e).

(f) Let λ and μ be two partitions. We must prove that the number $a_{\lambda,\mu}$ is well-defined. In other words, we must prove that there are only finitely many $\{0,1\}$-matrices of size $\ell(\lambda) \times \ell(\mu)$ having row sums λ and column sums μ.

But this is easy: Any $\{0,1\}$-matrix of size $\ell(\lambda) \times \ell(\mu)$ is a map from the set $\{1,2,\ldots,\ell(\lambda)\} \times \{1,2,\ldots,\ell(\mu)\}$ to the set $\{0,1\}$. Thus,

$$\begin{aligned}&\{\{0,1\}\text{-matrices of size } \ell(\lambda) \times \ell(\mu)\} \\ &= \{\text{maps from the set } \{1,2,\ldots,\ell(\lambda)\} \times \{1,2,\ldots,\ell(\mu)\} \text{ to the set } \{0,1\}\} \\ &= \{0,1\}^{\{1,2,\ldots,\ell(\lambda)\} \times \{1,2,\ldots,\ell(\mu)\}}\end{aligned}$$

[527] *Proof.* Let T be a column-strict tableau of shape λ having $\operatorname{cont}(T) = \lambda$. We must prove that T is equal to T_0.

Lemma 13.46.3(f) (applied to $\mu = \lambda$) shows that each $(i,j) \in F$ satisfies $T(i,j) = i$ (since $\lambda = \lambda$). Thus, each $(i,j) \in F$ satisfies

$$T(i,j) = i = T_0(i,j) \qquad (\text{since } T_0(i,j) \text{ is defined to be } i).$$

Recall that any column-strict tableau of shape λ is a map $F \to \{1,2,3,\ldots\}$. Hence, T and T_0 are maps $F \to \{1,2,3,\ldots\}$ (since T and T_0 are column-strict tableaux of shape λ). Therefore, we conclude that $T = T_0$ (since each $(i,j) \in F$ satisfies $T(i,j) = T_0(i,j)$). In other words, T is equal to T_0. Qed.

is a finite set (since both $\{1, 2, \ldots, \ell(\lambda)\} \times \{1, 2, \ldots, \ell(\mu)\}$ and $\{0, 1\}$ are finite sets). In other words, there are only finitely many $\{0, 1\}$-matrices of size $\ell(\lambda) \times \ell(\mu)$. Hence, there are only finitely many $\{0, 1\}$-matrices of size $\ell(\lambda) \times \ell(\mu)$ having row sums λ and column sums μ. This solves Exercise 2.2.13(f).

[*Remark:* We can prove a slightly stronger claim: Namely, there are only finitely many matrices in $\mathbb{N}^{\ell(\lambda) \times \ell(\mu)}$ having row sums λ and column sums μ.

Let us sketch the proof of this claim. Indeed, let $N = |\lambda|$. Then, if A is any matrix in $\mathbb{N}^{\ell(\lambda) \times \ell(\mu)}$ having row sums λ and column sums μ, then the sum of all entries of A must equal $|\lambda| = N$, and therefore each entry of A must be $\leq N$ (since a sum of nonnegative integers is always $\geq$ to each of its addends); but this entails that each entry of A belongs to the finite set $\{0, 1, \ldots, N\}$, and therefore there are only finitely many choices for each entry, which leads to only finitely many possible matrices A.]

(g) Exercise 2.2.13(g) is truly not a deep fact, but its proof requires some bookkeeping. In order to make this bookkeeping more palatable, we are going to introduce various auxiliary notations.

Definition 13.46.6. Let $q \in \mathbb{N}$. Let $\overline{\mathbf{x}}_q$ denote the q-tuple $(x_1, x_2, \ldots, x_q)$ of indeterminates. Let $\mathbf{k}[\overline{\mathbf{x}}_q]$ denote the polynomial ring $\mathbf{k}[x_1, x_2, \ldots, x_q]$. Let $\eta_q : R(\mathbf{x}) \to \mathbf{k}[\overline{\mathbf{x}}_q]$ be the map that sends every power series $f \in R(\mathbf{x})$ to the polynomial $f(x_1, x_2, \ldots, x_q, 0, 0, 0, \ldots)$. (This is well-defined, because Exercise 2.1.2 (applied to $A = \mathbf{k}[\overline{\mathbf{x}}_q]$ and $k = q$) shows that substituting $x_1, x_2, \ldots, x_q, 0, 0, 0, \ldots$ for $x_1, x_2, x_3, \ldots$ in f yields an infinite sum in which all but finitely many addends are zero.)

The map η_q is an evaluation homomorphism (in an appropriate sense[528]); thus, it is a $\mathbf{k}$-algebra homomorphism.

If $\beta = (\beta_1, \beta_2, \ldots, \beta_q) \in \mathbb{N}^q$ is a q-tuple of nonnegative integers, then $\overline{\mathbf{x}}_q^{\beta}$ shall denote the monomial $x_1^{\beta_1} x_2^{\beta_2} \cdots x_q^{\beta_q}$. This is a monomial in the polynomial ring $\mathbf{k}[\overline{\mathbf{x}}_q]$.

For every $f \in \mathbf{k}[[\overline{\mathbf{x}}_q]]$ and $\beta \in \mathbb{N}^q$, we let $\left[\overline{\mathbf{x}}_q^{\beta}\right] f$ denote the coefficient of the monomial $\overline{\mathbf{x}}_q^{\beta}$ in the power series f.

Let us make the following simple observations:

- Every $q \in \mathbb{N}$ and $i \in \{1, 2, \ldots, q\}$ satisfy

$$\eta_q(x_i) = x_i(x_1, x_2, \ldots, x_q, 0, 0, 0, \ldots) \qquad \text{(by the definition of } \eta_q\text{)}$$
$$= x_i \qquad \text{(since } i \in \{1, 2, \ldots, q\}\text{)}. \tag{13.46.25}$$

- Every $q \in \mathbb{N}$ and $i \in \{q+1, q+2, q+3, \ldots\}$ satisfy

$$\eta_q(x_i) = x_i(x_1, x_2, \ldots, x_q, 0, 0, 0, \ldots) \qquad \text{(by the definition of } \eta_q\text{)}$$
$$= 0 \qquad \text{(since } i \in \{q+1, q+2, q+3, \ldots\}\text{)}. \tag{13.46.26}$$

- Any two q-tuples $\phi \in \mathbb{N}^q$ and $\psi \in \mathbb{N}^q$ satisfy

$$\left[\overline{\mathbf{x}}_q^{\phi}\right]\left(\overline{\mathbf{x}}_q^{\psi}\right) = \delta_{\phi, \psi}. \tag{13.46.27}$$

(Indeed, $\overline{\mathbf{x}}_q^{\psi}$ and $\overline{\mathbf{x}}_q^{\phi}$ are two distinct monomials if $\phi \neq \psi$, and are two identical monomials if $\phi = \psi$.)

Lemma 13.46.7. *Let $q \in \mathbb{N}$. Let β be a weak composition. Assume that $\beta_i = 0$ for every integer $i > q$. Let $f \in R(\mathbf{x})$. Then,*

$$\left[\mathbf{x}^{\beta}\right] f = \left[\overline{\mathbf{x}}_q^{(\beta_1, \beta_2, \ldots, \beta_q)}\right](\eta_q(f)).$$

Proof of Lemma 13.46.7. Let us prove that every weak composition α satisfies

$$\left[\mathbf{x}^{\beta}\right](\mathbf{x}^{\alpha}) = \left[\overline{\mathbf{x}}_q^{(\beta_1, \beta_2, \ldots, \beta_q)}\right](\eta_q(\mathbf{x}^{\alpha})). \tag{13.46.28}$$

[*Proof of (13.46.28):* Let α be a weak composition. We must prove the equality (13.46.28).

We distinguish between two cases:

Case 1: We have ($\alpha_i = 0$ for every integer $i > q$).

Case 2: We don't have ($\alpha_i = 0$ for every integer $i > q$).

Let us first consider Case 1. In this case, we have

$$(\alpha_i = 0 \text{ for every integer } i > q). \tag{13.46.29}$$

[528] i.e., it acts on a power series $f \in R(\mathbf{x})$ by substituting certain values for the indeterminates $x_1, x_2, x_3, \ldots$

Thus,

$$\prod_{i=q+1}^{\infty} \underbrace{x_i^{\alpha_i}}_{\substack{=x_i^0 \\ (\text{since } \alpha_i=0 \text{ (by (13.46.29)))}}} = \prod_{i=q+1}^{\infty} \underbrace{x_i^0}_{=1} = 1. \tag{13.46.30}$$

Now,

$$\delta_{(\beta_1,\beta_2,\ldots,\beta_q),(\alpha_1,\alpha_2,\ldots,\alpha_q)} = \delta_{\beta,\alpha} \tag{13.46.31}$$

[529].

Now, the definition of $\mathbf{x}^\alpha$ yields

$$\mathbf{x}^\alpha = \prod_{i\geq 1} x_i^{\alpha_i} = \left(\prod_{i=1}^{q} x_i^{\alpha_i}\right) \underbrace{\left(\prod_{i=q+1}^{\infty} x_i^{\alpha_i}\right)}_{\substack{=1 \\ (\text{by (13.46.30)})}} = \prod_{i=1}^{q} x_i^{\alpha_i}.$$

Applying the map η_q to both sides of this equality, we obtain

$$\begin{aligned}
\eta_q\left(\mathbf{x}^\alpha\right) = \eta_q\left(\prod_{i=1}^{q} x_i^{\alpha_i}\right) &= \prod_{i=1}^{q} \left(\underbrace{\eta_q\left(x_i\right)}_{\substack{=x_i \\ (\text{by (13.46.25)})}}\right)^{\alpha_i} \\
&\qquad (\text{since } \eta_q \text{ is a } \mathbf{k}\text{-algebra homomorphism}) \\
&= \prod_{i=1}^{q} x_i^{\alpha_i} = x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_q^{\alpha_q} = \overline{\mathbf{x}}_q^{(\alpha_1,\alpha_2,\ldots,\alpha_q)} \qquad (13.46.32)
\end{aligned}$$

(since $\overline{\mathbf{x}}_q^{(\alpha_1,\alpha_2,\ldots,\alpha_q)}$ is defined to be $x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_q^{\alpha_q}$).

Now,

$$\begin{aligned}
&\left[\overline{\mathbf{x}}_q^{(\beta_1,\beta_2,\ldots,\beta_q)}\right] \left(\underbrace{\eta_q\left(\mathbf{x}^\alpha\right)}_{\substack{=\overline{\mathbf{x}}_q^{(\alpha_1,\alpha_2,\ldots,\alpha_q)} \\ (\text{by (13.46.32)})}}\right) \\
&= \left[\overline{\mathbf{x}}_q^{(\beta_1,\beta_2,\ldots,\beta_q)}\right] \left(\overline{\mathbf{x}}_q^{(\alpha_1,\alpha_2,\ldots,\alpha_q)}\right) = \delta_{(\beta_1,\beta_2,\ldots,\beta_q),(\alpha_1,\alpha_2,\ldots,\alpha_q)} \\
&\qquad (\text{by (13.46.27) (applied to } \phi = (\beta_1,\beta_2,\ldots,\beta_q) \text{ and } \psi = (\alpha_1,\alpha_2,\ldots,\alpha_q)\,)) \\
&= \delta_{\beta,\alpha} \qquad (\text{by (13.46.31)}).
\end{aligned}$$

[529] *Proof of (13.46.31):* We are in one of the following two subcases:

Subcase 1.1: We have $\beta \neq \alpha$.

Subcase 1.2: We have $\beta = \alpha$.

Let us first consider Subcase 1.1. In this subcase, we have $\beta \neq \alpha$. In other words, $\alpha \neq \beta$. Hence, there exists some $k \in \{1, 2, 3, \ldots\}$ satisfying $\alpha_k \neq \beta_k$. Consider this k.

We claim that $k \leq q$. Indeed, assume the contrary (for the sake of contradiction). Then, $k > q$. Hence, (13.46.29) (applied to $i = k$) yields $\alpha_k = 0$. But let us recall that $\beta_i = 0$ for every integer $i > q$. Applying this to $i = k$, we find $\beta_k = 0$. Hence, $\alpha_k = 0 = \beta_k$. This contradicts $\alpha_k \neq \beta_k$.

This contradiction completes the proof of $k \leq q$. Hence, $k \in \{1, 2, \ldots, q\}$. Hence, there exists some $i \in \{1, 2, \ldots, q\}$ satisfying $\alpha_i \neq \beta_i$ (namely, $i = k$). Therefore, $(\alpha_1, \alpha_2, \ldots, \alpha_q) \neq (\beta_1, \beta_2, \ldots, \beta_q)$. In other words, $(\beta_1, \beta_2, \ldots, \beta_q) \neq (\alpha_1, \alpha_2, \ldots, \alpha_q)$. Thus, $\delta_{(\beta_1,\beta_2,\ldots,\beta_q),(\alpha_1,\alpha_2,\ldots,\alpha_q)} = 0$. Comparing this with $\delta_{\beta,\alpha} = 0$ (since $\beta \neq \alpha$), we obtain $\delta_{(\beta_1,\beta_2,\ldots,\beta_q),(\alpha_1,\alpha_2,\ldots,\alpha_q)} = \delta_{\beta,\alpha}$. Thus, (13.46.31) is proven in Subcase 1.1.

Let us now consider Subcase 1.2. In this subcase, we have $\beta = \alpha$. Hence, $(\beta_1, \beta_2, \ldots, \beta_q) = (\alpha_1, \alpha_2, \ldots, \alpha_q)$. Thus, $\delta_{(\beta_1,\beta_2,\ldots,\beta_q),(\alpha_1,\alpha_2,\ldots,\alpha_q)} = 1$. Comparing this with $\delta_{\beta,\alpha} = 1$ (since $\beta = \alpha$), we obtain $\delta_{(\beta_1,\beta_2,\ldots,\beta_q),(\alpha_1,\alpha_2,\ldots,\alpha_q)} = \delta_{\beta,\alpha}$. Thus, (13.46.31) is proven in Subcase 1.2.

We have now proven (13.46.31) in each of the two Subcases 1.1 and 1.2. Hence, (13.46.31) always holds.

Comparing this with

$$\left[\mathbf{x}^{\beta}\right]\left(\mathbf{x}^{\alpha}\right)=\delta_{\beta,\alpha} \qquad \left(\text{by (13.46.1) (applied to } \mu=\beta\text{)}\right),$$

we obtain $\left[\mathbf{x}^{\beta}\right]\left(\mathbf{x}^{\alpha}\right)=\left[\overline{\mathbf{x}}_q^{(\beta_1,\beta_2,\ldots,\beta_q)}\right]\left(\eta_q\left(\mathbf{x}^{\alpha}\right)\right)$. Hence, (13.46.28) is proven in Case 1.

Let us now consider Case 2. In this case, we don't have ($\alpha_i=0$ for every integer $i>q$). In other words, there exists some integer $i>q$ such that $\alpha_i\neq 0$. Consider this i, and denote it by k. Thus, k is an integer such that $k>q$ and $\alpha_k\neq 0$.

Recall that $\beta_i=0$ for every integer $i>q$. Applying this to $i=k$, we obtain $\beta_k=0\neq\alpha_k$ (since $\alpha_k\neq 0$). Thus, $\beta\neq\alpha$, so that $\delta_{\beta,\alpha}=0$.

On the other hand, $\alpha_k\neq 0$ and thus $\alpha_k>0$ (since $\alpha_k\in\mathbb{N}$). Hence, the monomial $\mathbf{x}^{\alpha}$ is divisible by x_k. In other words, there exists a monomial $g\in\mathbf{k}[[\mathbf{x}]]$ such that $\mathbf{x}^{\alpha}=gx_k$. Consider this g. From $k>q$, we obtain $k\in\{q+1,q+2,q+3,\ldots\}$. Thus, (13.46.26) (applied to $i=k$) yields $\eta_q(x_k)=0$. Now,

$$\eta_q\left(\underbrace{\mathbf{x}^{\alpha}}_{=gx_k}\right)=\eta_q(gx_k)=\eta_q(g)\underbrace{\eta_q(x_k)}_{=0} \qquad \left(\text{since } \eta_q \text{ is a } \mathbf{k}\text{-algebra homomorphism}\right)$$
$$=0.$$

Hence,

$$\left[\overline{\mathbf{x}}_q^{(\beta_1,\beta_2,\ldots,\beta_q)}\right]\left(\underbrace{\eta_q\left(\mathbf{x}^{\alpha}\right)}_{=0}\right)=\left[\overline{\mathbf{x}}_q^{(\beta_1,\beta_2,\ldots,\beta_q)}\right](0)=0.$$

Comparing this with

$$\left[\mathbf{x}^{\beta}\right]\left(\mathbf{x}^{\alpha}\right)=\delta_{\beta,\alpha} \qquad \left(\text{by (13.46.1) (applied to } \mu=\beta\text{)}\right)$$
$$=0,$$

we obtain $\left[\mathbf{x}^{\beta}\right]\left(\mathbf{x}^{\alpha}\right)=\left[\overline{\mathbf{x}}_q^{(\beta_1,\beta_2,\ldots,\beta_q)}\right]\left(\eta_q\left(\mathbf{x}^{\alpha}\right)\right)$. Hence, (13.46.28) is proven in Case 2.

We have now proven (13.46.28) in each of the two Cases 1 and 2. Hence, (13.46.28) always holds.]

Now, let us notice that every power series $g\in\mathbf{k}[[\mathbf{x}]]$ satisfies

$$g=\sum_{\alpha\in\mathrm{WC}}\left[\mathbf{x}^{\alpha}\right](g)\cdot\mathbf{x}^{\alpha} \tag{13.46.33}$$

(since the family of the coefficients of g is $\left(\left[\mathbf{x}^{\alpha}\right](g)\right)_{\alpha\in\mathrm{WC}}$). Applying this to $g=f$, we obtain

$$f=\sum_{\alpha\in\mathrm{WC}}\left[\mathbf{x}^{\alpha}\right](f)\cdot\mathbf{x}^{\alpha}.$$

Substituting $x_1,x_2,\ldots,x_q,0,0,0,\ldots$ for the variables $x_1,x_2,x_3,\ldots$ in this equality, we obtain

$$f(x_1,x_2,\ldots,x_q,0,0,0,\ldots)=\sum_{\alpha\in\mathrm{WC}}\left[\mathbf{x}^{\alpha}\right](f)\cdot\underbrace{\mathbf{x}^{\alpha}(x_1,x_2,\ldots,x_q,0,0,0,\ldots)}_{\substack{=\eta_q(\mathbf{x}^{\alpha})\\ \text{(since } \eta_q(\mathbf{x}^{\alpha}) \text{ is defined to be } \mathbf{x}^{\alpha}(x_1,x_2,\ldots,x_q,0,0,0,\ldots))}}$$
$$=\sum_{\alpha\in\mathrm{WC}}\left[\mathbf{x}^{\alpha}\right](f)\cdot\eta_q\left(\mathbf{x}^{\alpha}\right). \tag{13.46.34}$$

Now, the definition of η_q yields

$$\eta_q(f)=f(x_1,x_2,\ldots,x_q,0,0,0,\ldots)$$
$$=\sum_{\alpha\in\mathrm{WC}}\left[\mathbf{x}^{\alpha}\right](f)\cdot\eta_q\left(\mathbf{x}^{\alpha}\right) \qquad \left(\text{by (13.46.34)}\right).$$

Hence,

$$\begin{aligned}
&\left[\overline{\mathbf{x}}_q^{(\beta_1,\beta_2,\ldots,\beta_q)}\right]\left(\underbrace{\eta_q\left(f\right)}_{=\sum_{\alpha\in\mathrm{WC}}\left[\mathbf{x}^\alpha\right](f)\cdot\eta_q(\mathbf{x}^\alpha)}\right)\\
&=\left[\overline{\mathbf{x}}_q^{(\beta_1,\beta_2,\ldots,\beta_q)}\right]\left(\sum_{\alpha\in\mathrm{WC}}\left[\mathbf{x}^\alpha\right]\left(f\right)\cdot\eta_q\left(\mathbf{x}^\alpha\right)\right)\\
&=\sum_{\alpha\in\mathrm{WC}}\left[\mathbf{x}^\alpha\right]\left(f\right)\cdot\underbrace{\left[\overline{\mathbf{x}}_q^{(\beta_1,\beta_2,\ldots,\beta_q)}\right]\left(\eta_q\left(\mathbf{x}^\alpha\right)\right)}_{\substack{=\left[\mathbf{x}^\beta\right](\mathbf{x}^\alpha)\\ \text{(by (13.46.28))}}}\\
&=\sum_{\alpha\in\mathrm{WC}}\left[\mathbf{x}^\alpha\right]\left(f\right)\cdot\left[\mathbf{x}^\beta\right]\left(\mathbf{x}^\alpha\right).
\end{aligned}$$

Comparing this with

$$\left[\mathbf{x}^\beta\right]\underbrace{f}_{=\sum_{\alpha\in\mathrm{WC}}\left[\mathbf{x}^\alpha\right](f)\cdot\mathbf{x}^\alpha}=\left[\mathbf{x}^\beta\right]\left(\sum_{\alpha\in\mathrm{WC}}\left[\mathbf{x}^\alpha\right]\left(f\right)\cdot\mathbf{x}^\alpha\right)=\sum_{\alpha\in\mathrm{WC}}\left[\mathbf{x}^\alpha\right]\left(f\right)\cdot\left[\mathbf{x}^\beta\right]\left(\mathbf{x}^\alpha\right),$$

we obtain $\left[\mathbf{x}^\beta\right]f=\left[\overline{\mathbf{x}}_q^{(\beta_1,\beta_2,\ldots,\beta_q)}\right]\left(\eta_q\left(f\right)\right)$. This proves Lemma 13.46.7. □

We now need to introduce some more notations.

Definition 13.46.8. If $q\in\mathbb{N}$, and if $\beta=(\beta_1,\beta_2,\ldots,\beta_q)\in\mathbb{N}^q$ is a q-tuple of nonnegative integers, then $|\beta|$ shall denote the sum $\beta_1+\beta_2+\cdots+\beta_q\in\mathbb{N}$.

Definition 13.46.9. We shall use the so-called *Iverson bracket notation*: For every assertion $\mathcal{A}$, we let $[\mathcal{A}]$ denote the integer $\begin{cases}1, & \text{if } \mathcal{A} \text{ is true;}\\ 0, & \text{if } \mathcal{A} \text{ is false}\end{cases}$.

Let us state a fundamental fact in combinatorics:

Lemma 13.46.10. *Let $q\in\mathbb{N}$. Let $m\in\mathbb{N}$. Define a set $\mathfrak{I}$ by*

$$\mathfrak{I}=\left\{(i_1,i_2,\ldots,i_m)\in\{1,2,\ldots,q\}^m\ \mid\ i_1<i_2<\cdots<i_m\right\}.$$

Let $\mathfrak{P}$ be the set of all m-element subsets of $\{1,2,\ldots,q\}$. Define a set $\mathfrak{K}$ by

$$\mathfrak{K}=\left\{\beta\in\{0,1\}^q\ \mid\ |\beta|=m\right\}.$$

(a) *The map $\mathfrak{I}\to\mathfrak{P}$, $(j_1,j_2,\ldots,j_m)\mapsto\{j_1,j_2,\ldots,j_m\}$ is well-defined and bijective.*
(b) *The map $\mathfrak{P}\to\mathfrak{K}$, $T\mapsto([1\in T],[2\in T],\ldots,[q\in T])$ is well-defined and bijective.*
(c) *There exists a bijection $\Psi:\mathfrak{I}\to\mathfrak{K}$ with the property that every $(i_1,i_2,\ldots,i_m)\in\mathfrak{I}$ satisfies $\overline{\mathbf{x}}_q^{\Psi(i_1,i_2,\ldots,i_m)}=x_{i_1}x_{i_2}\cdots x_{i_m}$.*

Proof of Lemma 13.46.10. (a) The set $\mathfrak{I}$ is the set of all strictly increasing lists of m elements of $\{1,2,\ldots,q\}$. Meanwhile, the set $\mathfrak{P}$ is the set of all m-element subsets of $\{1,2,\ldots,q\}$. Hence, there are well-known bijections between these two sets: The maps

$$\mathfrak{I}\to\mathfrak{P},\qquad (j_1,j_2,\ldots,j_m)\mapsto\{j_1,j_2,\ldots,j_m\}$$

and

$$\mathfrak{P}\to\mathfrak{I},\qquad T\mapsto(\text{the increasing list of }T)$$

[530] are mutually inverse bijections. In particular, the map $\mathfrak{I}\to\mathfrak{P}$, $(j_1,j_2,\ldots,j_m)\mapsto\{j_1,j_2,\ldots,j_m\}$ is well-defined and bijective. This proves Lemma 13.46.10(a).

[530] Here, the *increasing list* of a subset T of $\{1,2,\ldots,q\}$ is defined to be the list of all elements of T in increasing order (with no repetitions).

(b) It is well-known that the map

$$\Phi : \{\text{subsets } T \text{ of } \{1,2,\ldots,q\}\} \to \{0,1\}^q , \qquad T \mapsto ([1 \in T], [2 \in T], \ldots, [q \in T])$$

is a bijection[531]. Furthermore, this bijection Φ has the property that each subset T of $\{1,2,\ldots,q\}$ satisfies $|\Phi(T)| = |T|$. Thus, in particular, a subset T of $\{1,2,\ldots,q\}$ satisfies $|\Phi(T)| = m$ if and only if it satisfies $|T| = m$. Hence, Φ restricts to a bijection

$$\begin{aligned} &\{\text{subsets } T \text{ of } \{1,2,\ldots,q\} \ \mid \ |T| = m\} \to \{\beta \in \{0,1\}^q \ \mid \ |\beta| = m\}, \\ &T \mapsto ([1 \in T], [2 \in T], \ldots, [q \in T]). \end{aligned} \tag{13.46.35}$$

Thus, the map (13.46.35) is well-defined and bijective. Since $\{\text{subsets } T \text{ of } \{1,2,\ldots,q\} \ \mid \ |T| = m\} = \mathfrak{P}$ and $\{\beta \in \{0,1\}^q \ \mid \ |\beta| = m\} = \mathfrak{K}$, this rewrites as follows: The map $\mathfrak{P} \to \mathfrak{K}$, $T \mapsto ([1 \in T], [2 \in T], \ldots, [q \in T])$ is well-defined and bijective. This proves Lemma 13.46.10(b).

(c) Let A be the map $\mathfrak{I} \to \mathfrak{P}$, $(j_1, j_2, \ldots, j_m) \mapsto \{j_1, j_2, \ldots, j_m\}$. Lemma 13.46.10(a) shows that this map A is well-defined and bijective.

Let B be the map $\mathfrak{P} \to \mathfrak{K}$, $T \mapsto ([1 \in T], [2 \in T], \ldots, [q \in T])$. Lemma 13.46.10(b) shows that this map B is well-defined and bijective.

So the maps B and A are bijective. Hence, their composition $B \circ A$ is also bijective. Thus, $B \circ A : \mathfrak{I} \to \mathfrak{K}$ is a bijection. It has the property that every $(i_1, i_2, \ldots, i_m) \in \mathfrak{I}$ satisfies $\overline{\mathbf{x}}_q^{(B\circ A)(i_1,i_2,\ldots,i_m)} = x_{i_1} x_{i_2} \cdots x_{i_m}$ [532]. Hence, there exists a bijection $\Psi : \mathfrak{I} \to \mathfrak{K}$ with the property that every $(i_1, i_2, \ldots, i_m) \in \mathfrak{I}$ satisfies $\overline{\mathbf{x}}_q^{\Psi(i_1,i_2,\ldots,i_m)} = x_{i_1} x_{i_2} \cdots x_{i_m}$ (namely, $\Psi = B \circ A$). This proves Lemma 13.46.10(c). □

[531] Indeed, its inverse is the map that sends any $(\beta_1, \beta_2, \ldots, \beta_q) \in \{0,1\}^q$ to the subset $\{i \in \{1,2,\ldots,q\} \ \mid \ \beta_i = 1\}$ of $\{1,2,\ldots,q\}$.

[532] *Proof.* Let $(i_1, i_2, \ldots, i_m) \in \mathfrak{I}$. We must prove that $\overline{\mathbf{x}}_q^{(B\circ A)(i_1,i_2,\ldots,i_m)} = x_{i_1} x_{i_2} \cdots x_{i_m}$.

From $(i_1, i_2, \ldots, i_m) \in \mathfrak{I}$, we conclude that $(i_1, i_2, \ldots, i_m)$ is an element of $\{1,2,\ldots,q\}^m$ satisfying $i_1 < i_2 < \cdots < i_m$ (by the definition of $\mathfrak{I}$).

Define $T \in \mathfrak{P}$ by $T = A(i_1, i_2, \ldots, i_m)$. Then, $T = A(i_1, i_2, \ldots, i_m) = \{i_1, i_2, \ldots, i_m\}$ (by the definition of A). Hence, $(i_1, i_2, \ldots, i_m)$ is a list of all elements of T. Furthermore, this list has no repetitions (since $i_1, i_2, \ldots, i_m$ are distinct (since $i_1 < i_2 < \cdots < i_m$)). Hence, $(i_1, i_2, \ldots, i_m)$ is a list of all elements of T with no repetitions. Therefore,

$$\prod_{i\in T} x_i = x_{i_1} x_{i_2} \cdots x_{i_m}.$$

Also, T is a subset of $\{1,2,\ldots,q\}$ (since $T \in \mathfrak{P}$). Now,

$$(B \circ A)(i_1, i_2, \ldots, i_m) = B\left(\underbrace{A(i_1, i_2, \ldots, i_m)}_{=T}\right) = B(T) = ([1 \in T], [2 \in T], \ldots, [q \in T])$$

(by the definition of B). Hence,

$$\begin{aligned}
&\overline{\mathbf{x}}_q^{(B\circ A)(i_1,i_2,\ldots,i_m)} \\
&= \overline{\mathbf{x}}_q^{([1\in T],[2\in T],\ldots,[q\in T])} = x_1^{[1\in T]} x_2^{[2\in T]} \cdots x_q^{[q\in T]} \qquad \left(\text{by the definition of } \overline{\mathbf{x}}_q^{([1\in T],[2\in T],\ldots,[q\in T])}\right) \\
&= \prod_{i\in\{1,2,\ldots,q\}} x_i^{[i\in T]} = \left(\prod_{\substack{i\in\{1,2,\ldots,q\};\\ i\in T}} \underbrace{x_i^{[i\in T]}}_{\substack{=x_i^1 \\ (\text{since } [i\in T]=1 \\ (\text{since } i\in T))}}\right)\left(\prod_{\substack{i\in\{1,2,\ldots,q\};\\ \text{not } i\in T}} \underbrace{x_i^{[i\in T]}}_{\substack{=x_i^0 \\ (\text{since } [i\in T]=0 \\ (\text{since we don't have } i\in T))}}\right) \\
&= \left(\prod_{\substack{i\in\{1,2,\ldots,q\};\\ i\in T}} \underbrace{x_i^1}_{=x_i}\right)\left(\prod_{\substack{i\in\{1,2,\ldots,q\};\\ \text{not } i\in T}} \underbrace{x_i^0}_{=1}\right) = \left(\prod_{\substack{i\in\{1,2,\ldots,q\};\\ i\in T}} x_i\right)\underbrace{\left(\prod_{\substack{i\in\{1,2,\ldots,q\};\\ \text{not } i\in T}} 1\right)}_{=1} \\
&= \underbrace{\prod_{\substack{i\in\{1,2,\ldots,q\};\\ i\in T}}}_{\substack{=\prod_{i\in T} \\ (\text{since } T \text{ is a subset of } \{1,2,\ldots,q\})}} x_i = \prod_{i\in T} x_i = x_{i_1} x_{i_2} \cdots x_{i_m}.
\end{aligned}$$

Lemma 13.46.11. *Let $q \in \mathbb{N}$. Let $m \in \mathbb{N}$. Then,*

$$\eta_q(e_m) = \sum_{\substack{\beta \in \{0,1\}^q; \\ |\beta| = m}} \overline{\mathbf{x}}_q^{\beta}.$$

Proof of Lemma 13.46.11. The definition of e_m yields

$$e_m = \underbrace{\sum_{i_1 < i_2 < \cdots < i_m}}_{= \sum_{\substack{(i_1, i_2, \ldots, i_m) \in \{1,2,3,\ldots\}^m; \\ i_1 < i_2 < \cdots < i_m}}} x_{i_1} x_{i_2} \cdots x_{i_m} = \sum_{\substack{(i_1, i_2, \ldots, i_m) \in \{1,2,3,\ldots\}^m; \\ i_1 < i_2 < \cdots < i_m}} x_{i_1} x_{i_2} \cdots x_{i_m}.$$

Substituting $x_1, x_2, \ldots, x_q, 0, 0, 0, \ldots$ for the variables $x_1, x_2, x_3, \ldots$ in this equality, we obtain

$$(13.46.36) \qquad \begin{aligned} & e_m(x_1, x_2, \ldots, x_q, 0, 0, 0, \ldots) \\ & = \sum_{\substack{(i_1, i_2, \ldots, i_m) \in \{1,2,3,\ldots\}^m; \\ i_1 < i_2 < \cdots < i_m}} (x_{i_1} x_{i_2} \cdots x_{i_m})(x_1, x_2, \ldots, x_q, 0, 0, 0, \ldots). \end{aligned}$$

Note that $\{1, 2, \ldots, q\}^m$ is a subset of $\{1, 2, 3, \ldots\}^m$. Moreover, the following holds:

- If $(i_1, i_2, \ldots, i_m) \in \{1, 2, 3, \ldots\}^m$ is an m-tuple that does not satisfy $(i_1, i_2, \ldots, i_m) \in \{1, 2, \ldots, q\}^m$, then

$$(13.46.37) \qquad (x_{i_1} x_{i_2} \cdots x_{i_m})(x_1, x_2, \ldots, x_q, 0, 0, 0, \ldots) = 0$$

[533].

- If $(i_1, i_2, \ldots, i_m) \in \{1, 2, \ldots, q\}^m$, then

$$(13.46.38) \qquad (x_{i_1} x_{i_2} \cdots x_{i_m})(x_1, x_2, \ldots, x_q, 0, 0, 0, \ldots) = x_{i_1} x_{i_2} \cdots x_{i_m}$$

[534].

[533] *Proof of (13.46.37):* Let $(i_1, i_2, \ldots, i_m) \in \{1, 2, 3, \ldots\}^m$ be an m-tuple that does not satisfy $(i_1, i_2, \ldots, i_m) \in \{1, 2, \ldots, q\}^m$.

If every $k \in \{1, 2, \ldots, m\}$ would satisfy $i_k \in \{1, 2, \ldots, q\}$, then we would have $(i_1, i_2, \ldots, i_m) \in \{1, 2, \ldots, q\}^m$, which would contradict the fact that we do not have $(i_1, i_2, \ldots, i_m) \in \{1, 2, \ldots, q\}^m$. Hence, not every $k \in \{1, 2, \ldots, m\}$ satisfies $i_k \in \{1, 2, \ldots, q\}$. In other words, there exists some $k \in \{1, 2, \ldots, m\}$ satisfying $i_k \notin \{1, 2, \ldots, q\}$. Consider this k.

The element i_k belongs to $\{1, 2, 3, \ldots\}$ but not to $\{1, 2, \ldots, q\}$ (since $i_k \notin \{1, 2, \ldots, q\}$). Thus, $i_k \in \{1, 2, 3, \ldots\} \setminus \{1, 2, \ldots, q\} = \{q+1, q+2, q+3, \ldots\}$. Hence, (13.46.26) (applied to $i = i_k$) yields $\eta_q(x_{i_k}) = 0$.

Now,

$$\begin{aligned} & \underbrace{(x_{i_1} x_{i_2} \cdots x_{i_m})}_{= \prod_{j \in \{1,2,\ldots,m\}} x_{i_j}} (x_1, x_2, \ldots, x_q, 0, 0, 0, \ldots) \\ & = \left(\prod_{j \in \{1,2,\ldots,m\}} x_{i_j} \right) (x_1, x_2, \ldots, x_q, 0, 0, 0, \ldots) = \prod_{j \in \{1,2,\ldots,m\}} \underbrace{x_{i_j}(x_1, x_2, \ldots, x_q, 0, 0, 0, \ldots)}_{\substack{= \eta_q(x_{i_j}) \\ \text{(since } \eta_q(x_{i_j}) \text{ is defined to be } x_{i_j}(x_1, x_2, \ldots, x_q, 0, 0, 0, \ldots))}} \\ & = \prod_{j \in \{1,2,\ldots,m\}} \eta_q(x_{i_j}) = \underbrace{\eta_q(x_{i_k})}_{=0} \cdot \prod_{\substack{j \in \{1,2,\ldots,m\}; \\ j \neq k}} \eta_q(x_{i_j}) \qquad \text{(here, we have split off the factor for } j = k \text{ from the product)} \\ & = 0 \cdot \prod_{\substack{j \in \{1,2,\ldots,m\}; \\ j \neq k}} \eta_q(x_{i_j}) = 0. \end{aligned}$$

This proves (13.46.37).

[534] *Proof of (13.46.38):* Let $(i_1, i_2, \ldots, i_m) \in \{1, 2, \ldots, q\}^m$.

Let $k \in \{1, 2, \ldots, m\}$. From $(i_1, i_2, \ldots, i_m) \in \{1, 2, \ldots, q\}^m$, we obtain $i_k \in \{1, 2, \ldots, q\}$. Hence, (13.46.25) (applied to $i = i_k$) yields $\eta_q(x_{i_k}) = x_{i_k}$.

Now, (13.46.36) becomes

$$\begin{aligned}
& e_m\left(x_1, x_2, \ldots, x_q, 0, 0, 0, \ldots\right) \\
& = \sum_{\substack{(i_1,i_2,\ldots,i_m)\in\{1,2,3,\ldots\}^m; \\ i_1<i_2<\cdots<i_m}} \left(x_{i_1} x_{i_2} \cdots x_{i_m}\right)\left(x_1, x_2, \ldots, x_q, 0, 0, 0, \ldots\right) \\
& = \underbrace{\sum_{\substack{(i_1,i_2,\ldots,i_m)\in\{1,2,3,\ldots\}^m; \\ i_1<i_2<\cdots<i_m; \\ (i_1,i_2,\ldots,i_m)\in\{1,2,\ldots,q\}^m}}}_{\substack{=\sum_{\substack{(i_1,i_2,\ldots,i_m)\in\{1,2,\ldots,q\}^m; \\ i_1<i_2<\cdots<i_m}} \\ \text{(since } \{1,2,\ldots,q\}^m \text{ is a subset of } \{1,2,3,\ldots\}^m\text{)}}} \underbrace{\left(x_{i_1} x_{i_2} \cdots x_{i_m}\right)\left(x_1, x_2, \ldots, x_q, 0, 0, 0, \ldots\right)}_{\substack{=x_{i_1}x_{i_2}\cdots x_{i_m} \\ \text{(by (13.46.38))}}} \\
& \qquad + \sum_{\substack{(i_1,i_2,\ldots,i_m)\in\{1,2,3,\ldots\}^m; \\ i_1<i_2<\cdots<i_m; \\ \text{not } (i_1,i_2,\ldots,i_m)\in\{1,2,\ldots,q\}^m}} \underbrace{\left(x_{i_1} x_{i_2} \cdots x_{i_m}\right)\left(x_1, x_2, \ldots, x_q, 0, 0, 0, \ldots\right)}_{\substack{=0 \\ \text{(by (13.46.37))}}} \\
& = \sum_{\substack{(i_1,i_2,\ldots,i_m)\in\{1,2,\ldots,q\}^m; \\ i_1<i_2<\cdots<i_m}} x_{i_1} x_{i_2} \cdots x_{i_m} + \underbrace{\sum_{\substack{(i_1,i_2,\ldots,i_m)\in\{1,2,3,\ldots\}^m; \\ i_1<i_2<\cdots<i_m; \\ \text{not } (i_1,i_2,\ldots,i_m)\in\{1,2,\ldots,q\}^m}} 0}_{=0} \\
& = \sum_{\substack{(i_1,i_2,\ldots,i_m)\in\{1,2,\ldots,q\}^m; \\ i_1<i_2<\cdots<i_m}} x_{i_1} x_{i_2} \cdots x_{i_m}. \qquad (13.46.39)
\end{aligned}$$

Now, let us define the sets $\mathfrak{I}$, $\mathfrak{P}$ and $\mathfrak{K}$ as in Lemma 13.46.10. Then, Lemma 13.46.10(c) shows that there exists a bijection $\Psi : \mathfrak{I} \to \mathfrak{K}$ with the property that every $(i_1, i_2, \ldots, i_m) \in \mathfrak{I}$ satisfies

$$\overline{\mathbf{x}}_q^{\Psi(i_1,i_2,\ldots,i_m)} = x_{i_1} x_{i_2} \cdots x_{i_m}. \qquad (13.46.40)$$

Consider this Ψ.

Let us forget that we fixed k. We thus have shown that $\eta_q\left(x_{i_k}\right) = x_{i_k}$ for each $k \in \{1, 2, \ldots, m\}$. Hence, $\prod_{k=1}^m \underbrace{\eta_q\left(x_{i_k}\right)}_{=x_{i_k}} = \prod_{k=1}^m x_{i_k}$. Now,

$$\begin{aligned}
& \underbrace{\left(x_{i_1} x_{i_2} \cdots x_{i_m}\right)}_{=\prod_{k=1}^m x_{i_k}}\left(x_1, x_2, \ldots, x_q, 0, 0, 0, \ldots\right) \\
& = \left(\prod_{k=1}^m x_{i_k}\right)\left(x_1, x_2, \ldots, x_q, 0, 0, 0, \ldots\right) = \prod_{k=1}^m \underbrace{x_{i_k}\left(x_1, x_2, \ldots, x_q, 0, 0, 0, \ldots\right)}_{\substack{=\eta_q\left(x_{i_k}\right) \\ \text{(since } \eta_q\left(x_{i_k}\right) \text{ is defined to be } x_{i_k}\left(x_1,x_2,\ldots,x_q,0,0,0,\ldots\right)\text{)}}} \\
& = \prod_{k=1}^m \eta_q\left(x_{i_k}\right) = \prod_{k=1}^m x_{i_k} = x_{i_1} x_{i_2} \cdots x_{i_m}.
\end{aligned}$$

This proves (13.46.38).

Now, (13.46.39) becomes

$$\begin{aligned}
& e_m\left(x_1, x_2, \ldots, x_q, 0, 0, 0, \ldots\right) \\
& = \underbrace{\sum_{\substack{(i_1, i_2, \ldots, i_m) \in \{1,2,\ldots,q\}^m; \\ i_1 < i_2 < \cdots < i_m}}}_{\substack{=\sum_{(i_1, i_2, \ldots, i_m) \in \mathfrak{I}} \\ \left(\text{since } \mathfrak{I} = \left\{(i_1, i_2, \ldots, i_m) \in \{1,2,\ldots,q\}^m \mid i_1 < i_2 < \cdots < i_m\right\}\right)}} x_{i_1} x_{i_2} \cdots x_{i_m} = \sum_{(i_1, i_2, \ldots, i_m) \in \mathfrak{I}} \underbrace{x_{i_1} x_{i_2} \cdots x_{i_m}}_{\substack{=\overline{\mathbf{x}}_q^{\Psi(i_1, i_2, \ldots, i_m)} \\ \text{(by (13.46.40))}}} \\
& = \sum_{(i_1, i_2, \ldots, i_m) \in \mathfrak{I}} \overline{\mathbf{x}}_q^{\Psi(i_1, i_2, \ldots, i_m)} = \underbrace{\sum_{\beta \in \mathfrak{K}}}_{\substack{=\sum_{\substack{\beta \in \{0,1\}^q; \\ |\beta| = m}} \\ \left(\text{since } \mathfrak{K} = \left\{\beta \in \{0,1\}^q \mid |\beta| = m\right\}\right)}} \overline{\mathbf{x}}_q^{\beta} \\
& \qquad \left(\begin{array}{c}\text{here, we have substituted } \beta \text{ for } \Psi\left(i_1, i_2, \ldots, i_m\right) \text{ in the sum,} \\ \text{since the map } \Psi : \mathfrak{I} \to \mathfrak{K} \text{ is a bijection}\end{array}\right) \\
& = \sum_{\substack{\beta \in \{0,1\}^q; \\ |\beta| = m}} \overline{\mathbf{x}}_q^{\beta}.
\end{aligned}$$

Now, the definition of η_q yields

$$\eta_q\left(e_m\right) = e_m\left(x_1, x_2, \ldots, x_q, 0, 0, 0, \ldots\right) = \sum_{\substack{\beta \in \{0,1\}^q; \\ |\beta| = m}} \overline{\mathbf{x}}_q^{\beta}.$$

This proves Lemma 13.46.11. □

Lemma 13.46.12. *Let X and Y be two sets. Let $\phi : X \to Y$ be a bijection. Let Z be a subset of X. Then, the map $Z \to \phi(Z)$, $A \mapsto \phi(A)$ is well-defined and is a bijection.*

Proof of Lemma 13.46.12. Lemma 13.46.12 is a fundamental and trivial fact of set theory. □

Lemma 13.46.13. *Let $q \in \mathbb{N}$. Let $p \in \mathbb{N}$. Let $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_p) \in \mathbb{N}^p$ be a p-tuple of nonnegative integers. Then,*

$$\eta_q\left(e_{\alpha_1} e_{\alpha_2} \cdots e_{\alpha_p}\right) = \sum_{\substack{A \in \{0,1\}^{p \times q} \text{ is a } \{0,1\}\text{-matrix} \\ \text{having row sums } \alpha}} \overline{\mathbf{x}}_q^{\operatorname{colsums} A}.$$

Here, colsums A *denotes the column sums of A.*

Proof of Lemma 13.46.13. For each matrix $A \in \{0,1\}^{p \times q}$ and each $i \in \{1, 2, \ldots, p\}$, we let $\operatorname{row}_i A$ denote the i-th row of A. This $\operatorname{row}_i A$ is an element of $\{0,1\}^q$.

Let $\Phi : \{0,1\}^{p \times q} \to \left(\{0,1\}^q\right)^p$ be the map that sends each matrix $A \in \{0,1\}^{p \times q}$ to the list $(\operatorname{row}_1 A, \operatorname{row}_2 A, \ldots, \operatorname{row}_p A)$ of all the rows of A. Then, the map Φ is a bijection (because a matrix can be viewed as a list of rows).

For each matrix $A \in \{0,1\}^{p \times q}$, we let $\operatorname{rowsums} A$ denote the row sums of A. (This is a p-tuple in $\mathbb{N}^p$.) Furthermore, for each matrix $A \in \{0,1\}^{p \times q}$, we let $\operatorname{colsums} A$ denote the column sums of A. (This is a q-tuple in $\mathbb{N}^q$.)

Every matrix $A \in \{0,1\}^{p \times q}$ satisfies

$$\prod_{i=1}^{p} \overline{\mathbf{x}}_q^{\operatorname{row}_i A} = \overline{\mathbf{x}}_q^{\operatorname{colsums} A} \tag{13.46.41}$$

[535].

[535] *Proof of (13.46.41):* Let $A \in \{0,1\}^{p \times q}$ be a matrix.

Write the $p \times q$-matrix A in the form $A = \left(a_{i,j}\right)_{1 \leq i \leq p,\ 1 \leq j \leq q}$.

Thus, every $(\beta_1, \beta_2, \ldots, \beta_p) \in (\{0,1\}^q)^p$ satisfies

$$\prod_{i=1}^{p} \overline{\mathbf{x}}_q^{\beta_i} = \overline{\mathbf{x}}_q^{\operatorname{colsums}\left(\Phi^{-1}(\beta_1,\beta_2,\ldots,\beta_p)\right)} \tag{13.46.44}$$

[536].

Let $i \in \{1,2,\ldots,p\}$. The definition of $\operatorname{row}_i A$ yields

$$\operatorname{row}_i A = (\text{the } i\text{-th row of } A) = (a_{i,1}, a_{i,2}, \ldots, a_{i,q}) \qquad \left(\text{since } A = (a_{i,j})_{1\leq i\leq p,\ 1\leq j\leq q}\right).$$

Hence,

$$\begin{aligned}\overline{\mathbf{x}}_q^{\operatorname{row}_i A} &= \overline{\mathbf{x}}_q^{(a_{i,1},a_{i,2},\ldots,a_{i,q})} = x_1^{a_{i,1}} x_2^{a_{i,2}} \cdots x_q^{a_{i,q}} \qquad \left(\text{by the definition of } \overline{\mathbf{x}}_q^{(a_{i,1},a_{i,2},\ldots,a_{i,q})}\right) \\ &= \prod_{j=1}^{q} x_j^{a_{i,j}}.\end{aligned}$$

Now, forget that we fixed i. We thus have proven $\overline{\mathbf{x}}_q^{\operatorname{row}_i A} = \prod_{j=1}^q x_j^{a_{i,j}}$ for each $i \in \{1,2,\ldots,p\}$. Hence,

$$\prod_{i=1}^{p} \underbrace{\overline{\mathbf{x}}_q^{\operatorname{row}_i A}}_{=\prod_{j=1}^q x_j^{a_{i,j}}} = \underbrace{\prod_{i=1}^{p}\prod_{j=1}^{q}}_{=\prod_{j=1}^q \prod_{i=1}^p} x_j^{a_{i,j}} = \prod_{j=1}^{q}\prod_{i=1}^{p} x_j^{a_{i,j}}. \tag{13.46.42}$$

Now, for each $j \in \{1,2,\ldots,q\}$, we let c_j be the sum of all entries in the j-th column of A. Then, the definition of $\operatorname{colsums} A$ yields

$$\operatorname{colsums} A = (\text{the column sums of } A) = (c_1, c_2, \ldots, c_q)$$

(by the definition of the column sums of A). Thus,

$$\begin{aligned}\overline{\mathbf{x}}_q^{\operatorname{colsums} A} &= \overline{\mathbf{x}}_q^{(c_1,c_2,\ldots,c_q)} = x_1^{c_1} x_2^{c_2} \cdots x_q^{c_q} \qquad \left(\text{by the definition of } \overline{\mathbf{x}}_q^{(c_1,c_2,\ldots,c_q)}\right) \\ &= \prod_{j=1}^{q} x_j^{c_j}.\end{aligned} \tag{13.46.43}$$

But each $j \in \{1,2,\ldots,q\}$ satisfies

$$\begin{aligned} c_j &= \left(\text{the sum of all entries in } \underbrace{\text{the } j\text{-th column of } A}_{\substack{=(a_{1,j},a_{2,j},\ldots,a_{p,j})^T \\ (\text{since } A=(a_{i,j})_{1\leq i\leq p,\ 1\leq j\leq q})}}\right) \qquad (\text{by the definition of } c_j) \\ &= \left(\text{the sum of all entries in } (a_{1,j}, a_{2,j}, \ldots, a_{p,j})^T\right) = a_{1,j} + a_{2,j} + \cdots + a_{p,j} = \sum_{i=1}^{p} a_{i,j},\end{aligned}$$

and thus $x_j^{c_j} = x_j^{\sum_{i=1}^p a_{i,j}} = \prod_{i=1}^p x_j^{a_{i,j}}$. Hence,

$$\prod_{j=1}^{q} \underbrace{x_j^{c_j}}_{=\prod_{i=1}^p x_j^{a_{i,j}}} = \prod_{j=1}^{q}\prod_{i=1}^{p} x_j^{a_{i,j}} = \prod_{i=1}^{p} \overline{\mathbf{x}}_q^{\operatorname{row}_i A} \qquad (\text{by (13.46.42)}).$$

Thus,

$$\prod_{i=1}^{p} \overline{\mathbf{x}}_q^{\operatorname{row}_i A} = \prod_{j=1}^{q} x_j^{c_j} = \overline{\mathbf{x}}_q^{\operatorname{colsums} A} \qquad (\text{by (13.46.43)}).$$

This proves (13.46.41).

[536] *Proof of (13.46.44):* Let $(\beta_1, \beta_2, \ldots, \beta_p) \in (\{0,1\}^q)^p$. Define $A \in \{0,1\}^{p\times q}$ by $A = \Phi^{-1}(\beta_1, \beta_2, \ldots, \beta_p)$. Thus, $\Phi(A) = (\beta_1, \beta_2, \ldots, \beta_p)$. Hence, $(\beta_1, \beta_2, \ldots, \beta_p) = \Phi(A) = (\operatorname{row}_1 A, \operatorname{row}_2 A, \ldots, \operatorname{row}_p A)$ (by the definition of Φ). In other words, $\beta_i = \operatorname{row}_i A$ for each $i \in \{1,2,\ldots,p\}$. Thus, $\overline{\mathbf{x}}_q^{\beta_i} = \overline{\mathbf{x}}_q^{\operatorname{row}_i A}$ for each $i \in \{1,2,\ldots,p\}$. Hence,

$$\begin{aligned}\prod_{i=1}^{p} \underbrace{\overline{\mathbf{x}}_q^{\beta_i}}_{=\overline{\mathbf{x}}_q^{\operatorname{row}_i A}} &= \prod_{i=1}^{p} \overline{\mathbf{x}}_q^{\operatorname{row}_i A} = \overline{\mathbf{x}}_q^{\operatorname{colsums} A} \qquad (\text{by (13.46.41)}) \\ &= \overline{\mathbf{x}}_q^{\operatorname{colsums}\left(\Phi^{-1}(\beta_1,\beta_2,\ldots,\beta_p)\right)} \qquad \left(\text{since } A = \Phi^{-1}(\beta_1, \beta_2, \ldots, \beta_p)\right).\end{aligned}$$

This proves (13.46.44).

Furthermore, every matrix $A \in \{0,1\}^{p\times q}$ satisfies

$$\operatorname{rowsums} A = (\text{the row sums of } A) = \left(\left|\operatorname{row}_1 A\right|, \left|\operatorname{row}_2 A\right|, \ldots, \left|\operatorname{row}_p A\right|\right) \tag{13.46.45}$$

(because for each $i \in \{1,2,\ldots,p\}$, the sum of the entries of the i-th row of A is precisely $\left|\operatorname{row}_i A\right|$).

Let $\mathfrak{M}_\alpha$ be the set of all matrices $A \in \{0,1\}^{p\times q}$ satisfying $\operatorname{rowsums} A = \alpha$. Thus,

$$\begin{aligned} \mathfrak{M}_\alpha &= \left\{A \in \{0,1\}^{p\times q} \ \mid \ \operatorname{rowsums} A = \alpha\right\} && (13.46.46) \\ &= \left\{A \in \{0,1\}^{p\times q} \ \mid \ A \text{ is a } \{0,1\}\text{-matrix having row sums } \alpha\right\} && (13.46.47) \\ &\subset \{0,1\}^{p\times q}. \end{aligned}$$

For each $m \in \mathbb{N}$, we define a subset $\mathfrak{K}_m$ of $\{0,1\}^q$ by

$$\mathfrak{K}_m = \left\{\beta \in \{0,1\}^q \ \mid \ |\beta| = m\right\}.$$

This set $\mathfrak{K}_m$ is finite (since it is a subset of the finite set $\{0,1\}^q$).

We have $\mathfrak{K}_{\alpha_i} \subset \{0,1\}^q$ for each $i \in \{1,2,\ldots,p\}$. Hence, $\prod_{i=1}^p \mathfrak{K}_{\alpha_i} \subset \prod_{i=1}^p \left(\{0,1\}^q\right) = \left(\{0,1\}^q\right)^p$. Thus, $\mathfrak{K}_{\alpha_1} \times \mathfrak{K}_{\alpha_2} \times \cdots \times \mathfrak{K}_{\alpha_p} = \prod_{i=1}^p \mathfrak{K}_{\alpha_i} \subset \left(\{0,1\}^q\right)^p$.

It is easy to see that for each matrix $A \in \{0,1\}^{p\times q}$, we have the following logical equivalence:

$$\left(\Phi(A) \in \mathfrak{K}_{\alpha_1} \times \mathfrak{K}_{\alpha_2} \times \cdots \times \mathfrak{K}_{\alpha_p}\right) \iff \left(\operatorname{rowsums} A = \alpha\right) \tag{13.46.48}$$

[537]. Now, applying the map Φ to the equality (13.46.46), we find

$$\begin{aligned}
\Phi\left(\mathfrak{M}_\alpha\right) &= \Phi\left(\left\{A \in \{0,1\}^{p\times q} \ \mid \ \text{rowsums } A = \alpha\right\}\right) \\
&= \left\{ \Phi\left(A\right) \ \mid \ A \in \{0,1\}^{p\times q}; \qquad \underbrace{\text{rowsums } A = \alpha}_{\substack{\iff \left(\Phi(A)\in\mathfrak{K}_{\alpha_1}\times\mathfrak{K}_{\alpha_2}\times\cdots\times\mathfrak{K}_{\alpha_p}\right) \\ \text{(by (13.46.48))}}} \right\} \\
&= \left\{ \Phi\left(A\right) \ \mid \ A \in \{0,1\}^{p\times q}; \ \Phi\left(A\right) \in \mathfrak{K}_{\alpha_1}\times\mathfrak{K}_{\alpha_2}\times\cdots\times\mathfrak{K}_{\alpha_p} \right\} \\
&= \left\{ \beta \ \mid \ \beta \in \left(\{0,1\}^q\right)^p; \ \beta \in \mathfrak{K}_{\alpha_1}\times\mathfrak{K}_{\alpha_2}\times\cdots\times\mathfrak{K}_{\alpha_p} \right\} \\
&\qquad \left(\begin{array}{c} \text{here, we have substituted } \beta \text{ for } \Phi\left(A\right) \text{, since the map} \\ \Phi : \{0,1\}^{p\times q} \to \left(\{0,1\}^q\right)^p \text{ is a bijection} \end{array}\right)
\end{aligned}$$

$$= \left\{ \beta \in \left(\{0,1\}^q\right)^p \ \mid \ \beta \in \mathfrak{K}_{\alpha_1}\times\mathfrak{K}_{\alpha_2}\times\cdots\times\mathfrak{K}_{\alpha_p} \right\} = \mathfrak{K}_{\alpha_1}\times\mathfrak{K}_{\alpha_2}\times\cdots\times\mathfrak{K}_{\alpha_p} \tag{13.46.50}$$

(since $\mathfrak{K}_{\alpha_1}\times\mathfrak{K}_{\alpha_2}\times\cdots\times\mathfrak{K}_{\alpha_p} \subset \left(\{0,1\}^q\right)^p$).

Now, $\mathfrak{M}_\alpha$ is a subset of $\{0,1\}^{p\times q}$ (since $\mathfrak{M}_\alpha \subset \{0,1\}^{p\times q}$). Hence, Lemma 13.46.12 (applied to $X = \{0,1\}^{p\times q}$, $Y = \left(\{0,1\}^q\right)^p$, $\phi = \Phi$ and $Z = \mathfrak{M}_\alpha$) yields that the map $\mathfrak{M}_\alpha \to \Phi\left(\mathfrak{M}_\alpha\right), \ A \mapsto \Phi\left(A\right)$ is well-defined and is a bijection.

[537] *Proof of (13.46.48):* Let $A \in \{0,1\}^{p\times q}$ be a matrix.

Let $i \in \{1,2,\ldots,p\}$. Recall that $\operatorname{row}_i A$ denotes the i-th row of A; this i-th row is an element of $\{0,1\}^q$. Thus, $\operatorname{row}_i A \in \{0,1\}^q$. Now, we have the following chain of equivalences:

$$\begin{aligned}
&\left(\operatorname{row}_i A \in \underbrace{\mathfrak{K}_{\alpha_i}}_{\substack{=\{\beta\in\{0,1\}^q \ \mid \ |\beta|=\alpha_i\} \\ \text{(by the definition of } \mathfrak{K}_{\alpha_i}\text{)}}}\right) \iff \left(\operatorname{row}_i A \in \left\{\beta \in \{0,1\}^q \ \mid \ |\beta| = \alpha_i\right\}\right) \\
&\iff \left(\underbrace{\operatorname{row}_i A \in \{0,1\}^q}_{\substack{\text{This is always true} \\ \text{(since we know that } \operatorname{row}_i A\in\{0,1\}^q\text{)}}} \text{ and } \left|\operatorname{row}_i A\right| = \alpha_i\right) \\
&\iff \left(\left|\operatorname{row}_i A\right| = \alpha_i\right).
\end{aligned} \tag{13.46.49}$$

Now, forget that we fixed i. We thus have proven the equivalence (13.46.49) for each $i \in \{1,2,\ldots,p\}$.

The definition of Φ yields $\Phi\left(A\right) = \left(\operatorname{row}_1 A, \operatorname{row}_2 A, \ldots, \operatorname{row}_p A\right)$. Now, we have the following chain of logical equivalences:

$$\begin{aligned}
&\left(\underbrace{\Phi\left(A\right)}_{=\left(\operatorname{row}_1 A, \operatorname{row}_2 A, \ldots, \operatorname{row}_p A\right)} \in \mathfrak{K}_{\alpha_1}\times\mathfrak{K}_{\alpha_2}\times\cdots\times\mathfrak{K}_{\alpha_p}\right) \\
&\iff \left(\left(\operatorname{row}_1 A, \operatorname{row}_2 A, \ldots, \operatorname{row}_p A\right) \in \mathfrak{K}_{\alpha_1}\times\mathfrak{K}_{\alpha_2}\times\cdots\times\mathfrak{K}_{\alpha_p}\right) \\
&\iff \left(\underbrace{\operatorname{row}_i A \in \mathfrak{K}_{\alpha_i}}_{\substack{\iff \left(\left|\operatorname{row}_i A\right|=\alpha_i\right) \\ \text{(by (13.46.49))}}} \text{ for each } i \in \{1,2,\ldots,p\}\right) \\
&\iff \left(\left|\operatorname{row}_i A\right| = \alpha_i \text{ for each } i \in \{1,2,\ldots,p\}\right) \\
&\iff \left(\underbrace{\left(\left|\operatorname{row}_1 A\right|, \left|\operatorname{row}_2 A\right|, \ldots, \left|\operatorname{row}_p A\right|\right)}_{\substack{=\text{rowsums } A \\ \text{(by (13.46.45))}}} = \underbrace{\left(\alpha_1, \alpha_2, \ldots, \alpha_p\right)}_{=\alpha}\right) \\
&\iff \left(\text{rowsums } A = \alpha\right).
\end{aligned}$$

This proves (13.46.48).

Lemma 13.46.11 yields that

$$\eta_q(e_m) = \underbrace{\sum_{\substack{\beta\in\{0,1\}^q;\\ |\beta|=m}}}_{\substack{=\sum_{\beta\in\mathfrak{K}_m}\\ (\text{since } \mathfrak{K}_m=\{\beta\in\{0,1\}^q \ \mid\ |\beta|=m\})}} \overline{\mathbf{x}}_q^{\beta} = \sum_{\beta\in\mathfrak{K}_m}\overline{\mathbf{x}}_q^{\beta} \tag{13.46.51}$$

for every $m\in\mathbb{N}$.

But applying the map η_q to the equality $e_{\alpha_1}e_{\alpha_2}\cdots e_{\alpha_p}=\prod_{i=1}^{p}e_{\alpha_i}$, we obtain

$$\begin{aligned}
&\eta_q\left(e_{\alpha_1}e_{\alpha_2}\cdots e_{\alpha_p}\right)\\
&=\eta_q\left(\prod_{i=1}^{p}e_{\alpha_i}\right)=\prod_{i=1}^{p}\underbrace{\eta_q\left(e_{\alpha_i}\right)}_{\substack{=\sum_{\beta\in\mathfrak{K}_{\alpha_i}}\overline{\mathbf{x}}_q^{\beta}\\ (\text{by (13.46.51) (applied to } m=\alpha_i))}} \qquad (\text{since } \eta_q \text{ is a } \mathbf{k}\text{-algebra homomorphism})\\
&=\prod_{i=1}^{p}\sum_{\beta\in\mathfrak{K}_{\alpha_i}}\overline{\mathbf{x}}_q^{\beta}=\sum_{(\beta_1,\beta_2,\ldots,\beta_p)\in\mathfrak{K}_{\alpha_1}\times\mathfrak{K}_{\alpha_2}\times\cdots\times\mathfrak{K}_{\alpha_p}}\underbrace{\prod_{i=1}^{p}\overline{\mathbf{x}}_q^{\beta_i}}_{\substack{=\overline{\mathbf{x}}_q^{\operatorname{colsums}\left(\Phi^{-1}(\beta_1,\beta_2,\ldots,\beta_p)\right)}\\ (\text{by (13.46.44)})\\ (\text{since } (\beta_1,\beta_2,\ldots,\beta_p)\in\mathfrak{K}_{\alpha_1}\times\mathfrak{K}_{\alpha_2}\times\cdots\times\mathfrak{K}_{\alpha_p}\subset(\{0,1\}^q)^p)}}\\
&\qquad (\text{by the product rule})\\
&=\sum_{(\beta_1,\beta_2,\ldots,\beta_p)\in\mathfrak{K}_{\alpha_1}\times\mathfrak{K}_{\alpha_2}\times\cdots\times\mathfrak{K}_{\alpha_p}}\overline{\mathbf{x}}_q^{\operatorname{colsums}\left(\Phi^{-1}(\beta_1,\beta_2,\ldots,\beta_p)\right)}=\sum_{C\in\mathfrak{K}_{\alpha_1}\times\mathfrak{K}_{\alpha_2}\times\cdots\times\mathfrak{K}_{\alpha_p}}\overline{\mathbf{x}}_q^{\operatorname{colsums}\left(\Phi^{-1}(C)\right)}\\
&\qquad (\text{here, we have renamed the summation index } (\beta_1,\beta_2,\ldots,\beta_p) \text{ as } C)\\
&=\sum_{C\in\Phi(\mathfrak{M}_\alpha)}\overline{\mathbf{x}}_q^{\operatorname{colsums}\left(\Phi^{-1}(C)\right)}\qquad \left(\text{since } \mathfrak{K}_{\alpha_1}\times\mathfrak{K}_{\alpha_2}\times\cdots\times\mathfrak{K}_{\alpha_p}=\Phi(\mathfrak{M}_\alpha) \text{ (by (13.46.50))}\right)\\
&=\underbrace{\sum_{A\in\mathfrak{M}_\alpha}}_{\substack{=\sum_{\substack{A\in\{0,1\}^{p\times q} \text{ is a } \{0,1\}\text{-matrix}\\ \text{having row sums } \alpha}}\\ (\text{because of (13.46.47)})}}\underbrace{\overline{\mathbf{x}}_q^{\operatorname{colsums}\left(\Phi^{-1}(\Phi(A))\right)}}_{\substack{=\overline{\mathbf{x}}_q^{\operatorname{colsums} A}\\ (\text{since } \Phi^{-1}(\Phi(A))=A)}}\\
&\qquad\left(\begin{array}{c}\text{here, we have substituted } \Phi(A) \text{ for } C \text{ in the sum, since}\\ \text{the map } \mathfrak{M}_\alpha\to\Phi(\mathfrak{M}_\alpha),\ A\mapsto\Phi(A) \text{ is a bijection}\end{array}\right)\\
&=\sum_{\substack{A\in\{0,1\}^{p\times q} \text{ is a } \{0,1\}\text{-matrix}\\ \text{having row sums } \alpha}}\overline{\mathbf{x}}_q^{\operatorname{colsums} A}.
\end{aligned}$$

This proves Lemma 13.46.13. □

Lemma 13.46.14. *Let* $\lambda\in\operatorname{Par}$ *and* $\mu\in\operatorname{Par}$. *Then,* $[\mathbf{x}^\mu](e_\lambda)=a_{\lambda,\mu}$.

Proof of Lemma 13.46.14. Let $p=\ell(\lambda)$. Thus, $\lambda=(\lambda_1,\lambda_2,\ldots,\lambda_p)$ with $\lambda_1\geq\lambda_2\geq\cdots\geq\lambda_p>0$. Hence, $e_\lambda=e_{\lambda_1}e_{\lambda_2}\cdots e_{\lambda_p}$ (by the definition of e_λ). Hence, $e_\lambda\in\Lambda$ (since $e_i\in\Lambda$ for each $i\in\mathbb{N}$).

Let $q=\ell(\mu)$. Thus, $\mu_{q+1}=\mu_{q+2}=\mu_{q+3}=\cdots=0$. Therefore, $\mu=(\mu_1,\mu_2,\ldots,\mu_q)$ with $\mu_1\geq\mu_2\geq\cdots\geq\mu_q>0$.

For any matrix $A\in\{0,1\}^{p\times q}$, we let $\operatorname{colsums}A$ denote the column sums of A.

The definition of $a_{\lambda,\mu}$ yields

$$\begin{aligned}
& a_{\lambda,\mu} \\
&= \left(\text{the number of all } \{0,1\}\text{-matrices of size } \underbrace{\ell(\lambda)}_{=p} \times \underbrace{\ell(\mu)}_{=q} \text{ having row sums } \lambda \text{ and column sums } \mu\right) \\
&= (\text{the number of all } \{0,1\}\text{-matrices of size } p \times q \text{ having row sums } \lambda \text{ and column sums } \mu) \\
&= \left(\text{the number of all matrices } A \in \{0,1\}^{p\times q} \text{ having row sums } \lambda \text{ and column sums } \mu\right) \\
&= \left|\left\{A \in \{0,1\}^{p\times q} \ \mid \ \text{the row sums of } A \text{ are } \lambda\text{, and } \underbrace{\text{the column sums of } A \text{ are } \mu}_{\substack{\Longleftrightarrow\ (\operatorname{colsums} A=\mu) \\ (\text{since the column sums of } A \text{ is } \operatorname{colsums} A)}}\right\}\right| \\
&= \left|\left\{A \in \{0,1\}^{p\times q} \ \mid \ \text{the row sums of } A \text{ are } \lambda\text{, and } \operatorname{colsums} A = \mu\right\}\right|.
\end{aligned} \tag{13.46.52}$$

But $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_p) \in \mathbb{N}^p$ is a p-tuple of nonnegative integers. Hence, Lemma 13.46.13 (applied to λ and λ_i instead of α and α_i) yields

$$\eta_q\left(e_{\lambda_1} e_{\lambda_2} \cdots e_{\lambda_p}\right) = \sum_{\substack{A\in\{0,1\}^{p\times q} \text{ is a } \{0,1\}\text{-matrix} \\ \text{having row sums } \lambda}} \overline{\mathbf{x}}_q^{\operatorname{colsums} A}.$$

But $e_\lambda \in \Lambda \subset R(\mathbf{x})$. Also, $\mu_i = 0$ for every integer $i > q$ (since $q = \ell(\mu)$). Hence, Lemma 13.46.7 (applied to $\beta = \mu$ and $f = e_\lambda$) yields

$$\begin{aligned}
&[\mathbf{x}^\mu](e_\lambda) \\
&= \left[\overline{\mathbf{x}}_q^{(\mu_1,\mu_2,\ldots,\mu_q)}\right]\left(\eta_q\left(\underbrace{e_\lambda}_{=e_{\lambda_1}e_{\lambda_2}\cdots e_{\lambda_p}}\right)\right) = \left[\overline{\mathbf{x}}_q^{(\mu_1,\mu_2,\ldots,\mu_q)}\right]\left(\underbrace{\eta_q\left(e_{\lambda_1}e_{\lambda_2}\cdots e_{\lambda_p}\right)}_{=\sum_{\substack{A\in\{0,1\}^{p\times q} \text{ is a } \{0,1\}\text{-matrix}\\ \text{having row sums } \lambda}} \overline{\mathbf{x}}_q^{\text{colsums } A}}\right) \\
&= \left[\overline{\mathbf{x}}_q^{(\mu_1,\mu_2,\ldots,\mu_q)}\right]\left(\sum_{\substack{A\in\{0,1\}^{p\times q} \text{ is a } \{0,1\}\text{-matrix}\\ \text{having row sums } \lambda}} \overline{\mathbf{x}}_q^{\text{colsums } A}\right) \\
&= \sum_{\substack{A\in\{0,1\}^{p\times q} \text{ is a } \{0,1\}\text{-matrix}\\ \text{having row sums } \lambda}} \underbrace{\left[\overline{\mathbf{x}}_q^{(\mu_1,\mu_2,\ldots,\mu_q)}\right]\left(\overline{\mathbf{x}}_q^{\text{colsums } A}\right)}_{\substack{=\delta_{(\mu_1,\mu_2,\ldots,\mu_q),\text{colsums } A} \\ \text{(by (13.46.27) (applied to } \phi=(\mu_1,\mu_2,\ldots,\mu_q) \\ \text{and } \psi=\text{colsums } A))}} \\
&= \sum_{\substack{A\in\{0,1\}^{p\times q} \text{ is a } \{0,1\}\text{-matrix}\\ \text{having row sums } \lambda}} \underbrace{\delta_{(\mu_1,\mu_2,\ldots,\mu_q),\text{colsums } A}}_{\substack{=\delta_{\mu,\text{colsums } A} \\ \text{(since } (\mu_1,\mu_2,\ldots,\mu_q)=\mu)}} = \sum_{\substack{A\in\{0,1\}^{p\times q} \text{ is a } \{0,1\}\text{-matrix}\\ \text{having row sums } \lambda}} \delta_{\mu,\text{colsums } A} \\
&= \underbrace{\sum_{\substack{A\in\{0,1\}^{p\times q} \text{ is a } \{0,1\}\text{-matrix}\\ \text{having row sums } \lambda; \\ \text{colsums } A=\mu}}}_{=\sum_{\substack{A\in\{0,1\}^{p\times q}; \\ \text{the row sums of } A \text{ are } \lambda; \\ \text{colsums } A=\mu}}} \underbrace{\delta_{\mu,\text{colsums } A}}_{\substack{=1 \\ \text{(since colsums } A=\mu)}} + \sum_{\substack{A\in\{0,1\}^{p\times q} \text{ is a } \{0,1\}\text{-matrix}\\ \text{having row sums } \lambda; \\ \text{colsums } A\neq\mu}} \underbrace{\delta_{\mu,\text{colsums } A}}_{\substack{=0 \\ \text{(since colsums } A\neq\mu)}} \\
&= \sum_{\substack{A\in\{0,1\}^{p\times q}; \\ \text{the row sums of } A \text{ are } \lambda; \\ \text{colsums } A=\mu}} 1 + \underbrace{\sum_{\substack{A\in\{0,1\}^{p\times q} \text{ is a } \{0,1\}\text{-matrix}\\ \text{having row sums } \lambda; \\ \text{colsums } A\neq\mu}} 0}_{=0} = \sum_{\substack{A\in\{0,1\}^{p\times q}; \\ \text{the row sums of } A \text{ are } \lambda; \\ \text{colsums } A=\mu}} 1 \\
&= \left|\left\{A \in \{0,1\}^{p\times q} \ \mid \ \text{the row sums of } A \text{ are } \lambda, \text{ and colsums } A = \mu\right\}\right| \cdot 1 \\
&= \left|\left\{A \in \{0,1\}^{p\times q} \ \mid \ \text{the row sums of } A \text{ are } \lambda, \text{ and colsums } A = \mu\right\}\right|.
\end{aligned}$$

Comparing this with (13.46.52), we obtain $[\mathbf{x}^\mu](e_\lambda) = a_{\lambda,\mu}$. This proves Lemma 13.46.14. □

Now, let $\lambda \in \operatorname{Par}_n$. Let $p = \ell(\lambda)$. We have $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_p)$ with $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p > 0$ (since $p = \ell(\lambda)$). Hence, $e_\lambda = e_{\lambda_1}e_{\lambda_2}\cdots e_{\lambda_p}$ (by the definition of e_λ).

Moreover, $\lambda \in \operatorname{Par}_n$, so that $|\lambda| = n$. Hence, $n = \left|\underbrace{\lambda}_{=(\lambda_1,\lambda_2,\ldots,\lambda_p)}\right| = |(\lambda_1, \lambda_2, \ldots, \lambda_p)| = \lambda_1 + \lambda_2 + \cdots + \lambda_p$.

Let $i \in \{1, 2, \ldots, p\}$. Then, e_{λ_i} is a homogeneous element of Λ having degree λ_i (because for each $m \in \mathbb{N}$, the element e_m is a homogeneous element of Λ having degree m).

Now, forget that we fixed i. We thus have shown that for each $i \in \{1, 2, \ldots, p\}$, the element e_{λ_i} is a homogeneous element of Λ having degree λ_i. In other words, $e_{\lambda_1}, e_{\lambda_2}, \ldots, e_{\lambda_p}$ are homogeneous elements of Λ having degrees $\lambda_1, \lambda_2, \ldots, \lambda_p$, respectively. Hence, the product $e_{\lambda_1}e_{\lambda_2}\cdots e_{\lambda_p}$ of these elements is a homogeneous element of Λ having degree $\lambda_1 + \lambda_2 + \cdots + \lambda_p$. In light of $e_\lambda = e_{\lambda_1}e_{\lambda_2}\cdots e_{\lambda_p}$ and $n = \lambda_1 + \lambda_2 + \cdots + \lambda_p$, this rewrites as follows: The element e_λ is a homogeneous element of Λ having degree n.

In other words, $e_\lambda \in \Lambda_n$. Thus, Exercise 2.2.13(a) (applied to $f = e_\lambda$) yields

$$e_\lambda = \sum_{\mu \in \operatorname{Par}_n} \underbrace{\left([\mathbf{x}^\mu] (e_\lambda)\right)}_{\substack{=a_{\lambda,\mu} \\ \text{(by Lemma 13.46.14)}}} m_\mu = \sum_{\mu \in \operatorname{Par}_n} a_{\lambda,\mu} m_\mu .$$

This solves Exercise 2.2.13(g).

Before we solve Exercise 2.2.13(h), let us show a few lemmas:

Lemma 13.46.15. *Let $p \in \mathbb{N}$. Let $a_1, a_2, \ldots, a_p$ be p integers. Let $b_1, b_2, \ldots, b_p$ be p integers. Assume that*

$$a_i \geq b_i \qquad \text{for each } i \in \{1, 2, \ldots, p\} . \tag{13.46.53}$$

Assume furthermore that $\sum_{i=1}^p a_i \leq \sum_{i=1}^p b_i$. Then, $a_i = b_i$ for each $i \in \{1, 2, \ldots, p\}$.

Proof of Lemma 13.46.15. Let $j \in \{1, 2, \ldots, p\}$. Then, (13.46.53) (applied to $i = j$) shows that $a_j \geq b_j$. But

$$\underbrace{\sum_{i=1}^p}_{=\sum_{i\in\{1,2,\ldots,p\}}} a_i = \sum_{i\in\{1,2,\ldots,p\}} a_i = a_j + \sum_{\substack{i\in\{1,2,\ldots,p\}; \\ i\neq j}} \underbrace{a_i}_{\substack{\geq b_i \\ \text{(by (13.46.53))}}}$$
$$\text{(here, we have split off the addend for } i = j \text{ from the sum)}$$
$$\geq a_j + \sum_{\substack{i\in\{1,2,\ldots,p\}; \\ i\neq j}} b_i .$$

Hence,

$$a_j + \sum_{\substack{i\in\{1,2,\ldots,p\}; \\ i\neq j}} b_i \leq \sum_{i=1}^p a_i \leq \underbrace{\sum_{i=1}^p}_{=\sum_{i\in\{1,2,\ldots,p\}}} b_i = \sum_{i\in\{1,2,\ldots,p\}} b_i = b_j + \sum_{\substack{i\in\{1,2,\ldots,p\}; \\ i\neq j}} b_i$$
$$\text{(here, we have split off the addend for } i = j \text{ from the sum)} .$$

Subtracting $\sum_{\substack{i\in\{1,2,\ldots,p\}; \\ i\neq j}} b_i$ from both sides of this inequality, we obtain $a_j \leq b_j$. Combining this with $a_j \geq b_j$, we obtain $a_j = b_j$.

Now, forget that we fixed j. We thus have shown that $a_j = b_j$ for each $j \in \{1, 2, \ldots, p\}$. Renaming the variable j as i in this statement, we conclude that $a_i = b_i$ for each $i \in \{1, 2, \ldots, p\}$. This proves Lemma 13.46.15. □

Lemma 13.46.16. *Let $k \in \mathbb{N}$. Let $a_1, a_2, \ldots, a_k$ be k integers. Let $b_1, b_2, \ldots, b_k$ be k integers. Assume that*

$$a_j \geq b_j \qquad \text{for each } j \in \{1, 2, \ldots, k\} .$$

Assume furthermore that $\sum_{j=1}^k a_j \leq \sum_{j=1}^k b_j$. Then, $a_j = b_j$ for each $j \in \{1, 2, \ldots, k\}$.

Proof of Lemma 13.46.16. Lemma 13.46.16 is obtained from Lemma 13.46.15 upon renaming the variables p and i as k and j. Thus, Lemma 13.46.16 follows from Lemma 13.46.15. □

The next few lemmas use the so-called *Iverson bracket notation*:

Definition 13.46.17. For every assertion $\mathcal{A}$, we let $[\mathcal{A}]$ denote the integer $\begin{cases} 1, & \text{if } \mathcal{A} \text{ is true;} \\ 0, & \text{if } \mathcal{A} \text{ is false} \end{cases}$.

Lemma 13.46.18. *Let $q \in \mathbb{N}$ and $r \in \mathbb{N}$ be such that $r \leq q$. Then, $\sum_{j=1}^q [j \leq r] = r$.*

Proof of Lemma 13.46.18. We have $0 \leq r \leq q$. Hence,

$$\sum_{j=1}^q [j \leq r] = \sum_{j=1}^r \underbrace{[j \leq r]}_{\substack{=1 \\ \text{(since } j\leq r)}} + \sum_{j=r+1}^q \underbrace{[j \leq r]}_{\substack{=0 \\ \text{(since we don't have } j\leq r \\ \text{(since } j\geq r+1>r))}} = \sum_{j=1}^r 1 + \underbrace{\sum_{j=r+1}^q 0}_{=0} = \sum_{j=1}^r 1 = r \cdot 1 = r.$$

This proves Lemma 13.46.18. □

Lemma 13.46.19. *Let $p \in \mathbb{N}$ and $q \in \mathbb{N}$. Let $A = (a_{i,j})_{1 \leq i \leq p,\ 1 \leq j \leq q} \in \{0,1\}^{p \times q}$ be a $\{0,1\}$-matrix. Let $(\lambda_1, \lambda_2, \ldots, \lambda_p)$ be the row sums of A. Let $(\mu_1, \mu_2, \ldots, \mu_q)$ be the column sums of A. Then:*

(a) *We have $\lambda_i = \sum_{j=1}^{q} a_{i,j}$ for each $i \in \{1, 2, \ldots, p\}$.*
(b) *We have $\mu_j = \sum_{i=1}^{p} a_{i,j}$ for each $j \in \{1, 2, \ldots, q\}$.*
(c) *We have $\sum_{i=1}^{p} \lambda_i = \sum_{j=1}^{q} \mu_j$.*
(d) *We have $\min\{\lambda_i, k\} \geq \sum_{j=1}^{k} a_{i,j}$ for each $k \in \{0, 1, \ldots, q\}$ and $i \in \{1, 2, \ldots, p\}$.*
(e) *We have $\sum_{i=1}^{p} \min\{\lambda_i, k\} \geq \sum_{j=1}^{k} \mu_j$ for each $k \in \{0, 1, \ldots, q\}$.*
(f) *If $\left(\sum_{i=1}^{p} \min\{\lambda_i, k\} = \sum_{j=1}^{k} \mu_j \text{ for each } k \in \{1, 2, \ldots, q\}\right)$, then $A = ([j \leq \lambda_i])_{1 \leq i \leq p,\ 1 \leq j \leq q}$.*

Proof of Lemma 13.46.19. We have $(a_{i,j})_{1 \leq i \leq p,\ 1 \leq j \leq q} = A \in \{0,1\}^{p \times q}$. Thus, $a_{i,j} \in \{0,1\}$ for each $i \in \{1, 2, \ldots, p\}$ and $j \in \{1, 2, \ldots, q\}$. Hence, for each $i \in \{1, 2, \ldots, p\}$ and $j \in \{1, 2, \ldots, q\}$, we have

$$a_{i,j} \geq 0 \qquad (\text{since } a_{i,j} \in \{0,1\}) \tag{13.46.54}$$

and

$$a_{i,j} \leq 1 \qquad (\text{since } a_{i,j} \in \{0,1\}) . \tag{13.46.55}$$

(a) Recall that $(\lambda_1, \lambda_2, \ldots, \lambda_p)$ is the row sums of A. By the definition of "row sums", this means that for each $i \in \{1, 2, \ldots, p\}$, the number λ_i is the sum of all entries in the i-th row of A. Thus, for each $i \in \{1, 2, \ldots, p\}$, we have

$$\lambda_i = \left(\text{the sum of all entries in } \underbrace{\text{the } i\text{-th row of } A}_{\substack{=(a_{i,1}, a_{i,2}, \ldots, a_{i,q}) \\ (\text{since } A = (a_{i,j})_{1 \leq i \leq p,\ 1 \leq j \leq q})}} \right)$$
$$= (\text{the sum of all entries in } (a_{i,1}, a_{i,2}, \ldots, a_{i,q})) = a_{i,1} + a_{i,2} + \cdots + a_{i,q} = \sum_{j=1}^{q} a_{i,j}.$$

This proves Lemma 13.46.19(a).

(b) The proof of Lemma 13.46.19(a) can easily be adapted (mutatis mutandis) to yield a proof of Lemma 13.46.19(b).

(c) Comparing

$$\sum_{i=1}^{p} \underbrace{\lambda_i}_{\substack{=\sum_{j=1}^{q} a_{i,j} \\ (\text{by Lemma 13.46.19(a)})}} = \underbrace{\sum_{i=1}^{p} \sum_{j=1}^{q}}_{=\sum_{j=1}^{q} \sum_{i=1}^{p}} a_{i,j} = \sum_{j=1}^{q} \sum_{i=1}^{p} a_{i,j}$$

with

$$\sum_{j=1}^{q} \underbrace{\mu_j}_{\substack{=\sum_{i=1}^{p} a_{i,j} \\ (\text{by Lemma 13.46.19(b)})}} = \sum_{j=1}^{q} \sum_{i=1}^{p} a_{i,j},$$

we obtain $\sum_{i=1}^{p} \lambda_i = \sum_{j=1}^{q} \mu_j$. This proves Lemma 13.46.19(c).

(d) Let $k \in \{0, 1, \ldots, q\}$ and $i \in \{1, 2, \ldots, p\}$. Lemma 13.46.19(a) yields $\lambda_i = \sum_{j=1}^{q} a_{i,j}$. But

$$\sum_{j=1}^{k} \underbrace{a_{i,j}}_{\substack{\leq 1 \\ (\text{by (13.46.55)})}} \leq \sum_{j=1}^{k} 1 = k \cdot 1 = k.$$

Furthermore, $k \in \{0, 1, \ldots, q\}$, so that $0 \leq k \leq q$. Hence,

$$\sum_{j=1}^{q} a_{i,j} = \sum_{j=1}^{k} a_{i,j} + \sum_{j=k+1}^{q} \underbrace{a_{i,j}}_{\substack{\geq 0 \\ (\text{by (13.46.54)})}} \geq \sum_{j=1}^{k} a_{i,j} + \underbrace{\sum_{j=k+1}^{q} 0}_{=0} = \sum_{j=1}^{k} a_{i,j},$$

so that

$$\sum_{j=1}^{k} a_{i,j} \leq \sum_{j=1}^{q} a_{i,j} = \lambda_i .$$

But let us recall the following basic fact about integers: If three integers α, β, γ satisfy $\alpha \leq \beta$ and $\alpha \leq \gamma$, then $\alpha \leq \min\{\beta, \gamma\}$. Applying this fact to $\alpha = \sum_{j=1}^{k} a_{i,j}$, $\beta = \lambda_i$ and $\gamma = k$, we conclude that $\sum_{j=1}^{k} a_{i,j} \leq \min\{\lambda_i, k\}$ (since $\sum_{j=1}^{k} a_{i,j} \leq \lambda_i$ and $\sum_{j=1}^{k} a_{i,j} \leq k$). In other words, $\min\{\lambda_i, k\} \geq \sum_{j=1}^{k} a_{i,j}$. This proves Lemma 13.46.19(d).

(e) Let $k \in \{0, 1, \ldots, q\}$. Now,

$$\sum_{i=1}^{p} \underbrace{\min\{\lambda_i, k\}}_{\substack{\geq \sum_{j=1}^{k} a_{i,j} \\ \text{(by Lemma 13.46.19(d))}}} \geq \underbrace{\sum_{i=1}^{p} \sum_{j=1}^{k}}_{= \sum_{j=1}^{k} \sum_{i=1}^{p}} a_{i,j} = \sum_{j=1}^{k} \underbrace{\sum_{i=1}^{p} a_{i,j}}_{\substack{= \mu_j \\ \text{(by Lemma 13.46.19(b))}}} = \sum_{j=1}^{k} \mu_j .$$

This proves Lemma 13.46.19(e).

(f) Assume that

$$\left(\sum_{i=1}^{p} \min\{\lambda_i, k\} = \sum_{j=1}^{k} \mu_j \text{ for each } k \in \{1, 2, \ldots, q\} \right) . \tag{13.46.56}$$

We want to show that $A = ([j \leq \lambda_i])_{1 \leq i \leq p,\ 1 \leq j \leq q}$.

Let $k \in \{1, 2, \ldots, q\}$. Thus, $k \in \{1, 2, \ldots, q\} \subset \{0, 1, \ldots, q\}$. But (13.46.56) yields

$$\sum_{i=1}^{p} \min\{\lambda_i, k\} = \sum_{j=1}^{k} \underbrace{\mu_j}_{\substack{= \sum_{i=1}^{p} a_{i,j} \\ \text{(by Lemma 13.46.19(b))}}} = \underbrace{\sum_{j=1}^{k} \sum_{i=1}^{p}}_{= \sum_{i=1}^{p} \sum_{j=1}^{k}} a_{i,j} = \sum_{i=1}^{p} \sum_{j=1}^{k} a_{i,j} \leq \sum_{i=1}^{p} \sum_{j=1}^{k} a_{i,j} .$$

Furthermore, we know that $\min\{\lambda_i, k\} \geq \sum_{j=1}^{k} a_{i,j}$ for each $i \in \{1, 2, \ldots, p\}$ (by Lemma 13.46.19(d)). Hence, Lemma 13.46.15 (applied to $\min\{\lambda_i, k\}$ and $\sum_{j=1}^{k} a_{i,j}$ instead of a_i and b_i) shows that

$$\min\{\lambda_i, k\} = \sum_{j=1}^{k} a_{i,j} \qquad \text{for each } i \in \{1, 2, \ldots, p\} . \tag{13.46.57}$$

Now, forget that we fixed k. We thus have proven (13.46.57) for each $k \in \{1, 2, \ldots, q\}$.

Now, pick $i \in \{1, 2, \ldots, p\}$. We shall show that

$$a_{i,k} \geq [k \leq \lambda_i] \qquad \text{for each } k \in \{1, 2, \ldots, q\} . \tag{13.46.58}$$

[*Proof of (13.46.58):* Let $k \in \{1, 2, \ldots, q\}$. We must prove the inequality $a_{i,k} \geq [k \leq \lambda_i]$.

Indeed, we are in one of the following two cases:

Case 1: We have $k \leq \lambda_i$.

Case 2: We don't have $k \leq \lambda_i$.

Let us first consider Case 1. In this case, we have $k \leq \lambda_i$. Thus, $[k \leq \lambda_i] = 1$. Also, $k \geq 1$ (since $k \in \{1, 2, \ldots, q\}$), so that $k \in \{1, 2, \ldots, k\}$. But from $k \leq \lambda_i$, we also obtain $\min\{\lambda_i, k\} = k$. Thus,

$$\sum_{j=1}^{k} 1 = k \cdot 1 = k \leq k = \min\{\lambda_i, k\} = \sum_{j=1}^{k} a_{i,j}$$

(by (13.46.57)). Furthermore, $1 \geq a_{i,j}$ for each $j \in \{1, 2, \ldots, k\}$ [538]. Hence, Lemma 13.46.16 (applied to 1 and $a_{i,j}$ instead of a_j and b_j) shows that $1 = a_{i,j}$ for each $j \in \{1, 2, \ldots, k\}$. Applying this to $j = k$, we obtain $1 = a_{i,k}$ (since $k \in \{1, 2, \ldots, k\}$). Hence, $a_{i,k} = 1 = [k \leq \lambda_i]$, so that $a_{i,k} \geq [k \leq \lambda_i]$. Thus, the inequality $a_{i,k} \geq [k \leq \lambda_i]$ is proven in Case 1.

[538] *Proof.* Let $j \in \{1, 2, \ldots, k\}$. Then, $j \in \{1, 2, \ldots, k\} \subset \{1, 2, \ldots, q\}$ (since $k \in \{1, 2, \ldots, q\}$). Hence, (13.46.55) yields $a_{i,j} \leq 1$. In other words, $1 \geq a_{i,j}$. Qed.

Next, let us consider Case 2. In this case, we don't have $k \leq \lambda_i$. Thus, $[k \leq \lambda_i] = 0$. But (13.46.54) (applied to $j = k$) yields $a_{i,k} \geq 0$. Thus, $a_{i,k} \geq 0 = [k \leq \lambda_i]$. Hence, the inequality $a_{i,k} \geq [k \leq \lambda_i]$ is proven in Case 2.

We have now proven the inequality $a_{i,k} \geq [k \leq \lambda_i]$ in each of the two Cases 1 and 2. Thus, the inequality $a_{i,k} \geq [k \leq \lambda_i]$ always holds. This proves (13.46.58).]

Lemma 13.46.19(a) yields $\lambda_i = \sum_{j=1}^{q} a_{i,j}$. But Lemma 13.46.19(d) (applied to $k = q$) yields $\min\{\lambda_i, q\} \geq \sum_{j=1}^{q} a_{i,j}$. Hence, $\sum_{j=1}^{q} a_{i,j} \leq \min\{\lambda_i, q\}$, so that $\lambda_i = \sum_{j=1}^{q} a_{i,j} \leq \min\{\lambda_i, q\}$. If we had $\lambda_i > q$, then we would have $\min\{\lambda_i, q\} = q < \lambda_i \leq \min\{\lambda_i, q\}$, which would be absurd. Thus, we cannot have $\lambda_i > q$. We therefore must have $\lambda_i \leq q$. Hence, Lemma 13.46.18 (applied to $r = \lambda_i$) yields

$$\sum_{j=1}^{q} [j \leq \lambda_i] = \lambda_i = \sum_{j=1}^{q} a_{i,j}.$$

Thus,

$$\sum_{j=1}^{q} a_{i,j} = \sum_{j=1}^{q} [j \leq \lambda_i] \leq \sum_{j=1}^{q} [j \leq \lambda_i].$$

Also,

$$a_{i,j} \geq [j \leq \lambda_i] \qquad \text{for each } j \in \{1, 2, \ldots, q\}$$

(by (13.46.58), applied to $k = j$). Hence, Lemma 13.46.16 (applied to q, $a_{i,j}$ and $[j \leq \lambda_i]$ instead of k, a_j and b_j) shows that

$$a_{i,j} = [j \leq \lambda_i] \qquad \text{for each } j \in \{1, 2, \ldots, q\}. \tag{13.46.59}$$

Now, forget that we fixed i. We thus have proven (13.46.59) for each $i \in \{1, 2, \ldots, p\}$.

Now,

$$A = \left(\underbrace{a_{i,j}}_{\substack{=[j \leq \lambda_i] \\ (\text{by } (13.46.59))}} \right)_{1 \leq i \leq p,\ 1 \leq j \leq q} = ([j \leq \lambda_i])_{1 \leq i \leq p,\ 1 \leq j \leq q}.$$

This proves Lemma 13.46.19(f). □

Lemma 13.46.20. *Let $n \in \mathbb{N}$ and $q \in \mathbb{N}$. Let $\lambda \in \operatorname{Par}_n$ and $\mu \in \operatorname{Par}_n$ be such that $\ell(\mu) \leq q$. Assume that*

$$\lambda_1 + \lambda_2 + \cdots + \lambda_k \geq \mu_1 + \mu_2 + \cdots + \mu_k \qquad \text{for each } k \in \{1, 2, \ldots, q\}.$$

Then, $\lambda \triangleright \mu$.

Proof of Lemma 13.46.20. We have assumed that

$$\lambda_1 + \lambda_2 + \cdots + \lambda_k \geq \mu_1 + \mu_2 + \cdots + \mu_k \qquad \text{for each } k \in \{1, 2, \ldots, q\}. \tag{13.46.60}$$

Hence,

$$\lambda_1 + \lambda_2 + \cdots + \lambda_q \geq \mu_1 + \mu_2 + \cdots + \mu_q \tag{13.46.61}$$

[539].

Both λ and μ are partitions of n (since $\lambda \in \operatorname{Par}_n$ and $\mu \in \operatorname{Par}_n$). Thus, we have $\lambda \triangleright \mu$ if and only if

$$\lambda_1 + \lambda_2 + \cdots + \lambda_k \geq \mu_1 + \mu_2 + \cdots + \mu_k \qquad \text{for all } k \in \{1, 2, \ldots, n\} \tag{13.46.62}$$

(by the definition of the dominance order).

Now, we are going to prove (13.46.62).

[*Proof of (13.46.62):* Let $k \in \{1, 2, \ldots, n\}$. We must prove that $\lambda_1 + \lambda_2 + \cdots + \lambda_k \geq \mu_1 + \mu_2 + \cdots + \mu_k$. If $k \in \{1, 2, \ldots, q\}$, then this follows immediately from (13.46.60). Hence, for the rest of this proof, we WLOG assume that $k \notin \{1, 2, \ldots, q\}$. Combining $k \in \{1, 2, \ldots, n\} \subset \{1, 2, 3, \ldots\}$ with $k \notin \{1, 2, \ldots, q\}$, we

[539] *Proof of (13.46.61):* If $q = 0$, then the inequality (13.46.61) holds because both of its sides equal 0 (in fact, an empty sum is 0). Hence, for the rest of this proof, we WLOG assume that $q \neq 0$. Thus, q is a positive integer (since $q \in \mathbb{N}$). Thus, $q \in \{1, 2, \ldots, q\}$. Therefore, (13.46.60) (applied to $k = q$) yields $\lambda_1 + \lambda_2 + \cdots + \lambda_q \geq \mu_1 + \mu_2 + \cdots + \mu_q$. This proves (13.46.61).

obtain $k \in \{1,2,3,\ldots\} \setminus \{1,2,\ldots,q\} = \{q+1,q+2,q+3,\ldots\}$. Thus, $k \geq q+1 > q$. Hence, $q < k$. Recall also that $\ell(\mu) \leq q$, hence $q \geq \ell(\mu)$. Now,

$$\mu_1 + \mu_2 + \cdots + \mu_k = \sum_{i=1}^{k} \mu_i = \sum_{i=1}^{q} \mu_i + \sum_{i=q+1}^{k} \underbrace{\mu_i}_{\substack{=0 \\ (\text{since } i \geq q+1 > q \geq \ell(\mu))}} \qquad (\text{since } 0 \leq q < k)$$

$$= \sum_{i=1}^{q} \mu_i + \underbrace{\sum_{i=q+1}^{k} 0}_{=0} = \sum_{i=1}^{q} \mu_i = \mu_1 + \mu_2 + \cdots + \mu_q. \tag{13.46.63}$$

On the other hand,

$$\begin{aligned}
\lambda_1 + \lambda_2 + \cdots + \lambda_k &= \sum_{i=1}^{k} \lambda_i = \sum_{i=1}^{q} \lambda_i + \sum_{i=q+1}^{k} \underbrace{\lambda_i}_{\geq 0} \qquad (\text{since } 0 \leq q < k) \\
&\geq \sum_{i=1}^{q} \lambda_i + \underbrace{\sum_{i=q+1}^{k} 0}_{=0} = \sum_{i=1}^{q} \lambda_i = \lambda_1 + \lambda_2 + \cdots + \lambda_q \\
&\geq \mu_1 + \mu_2 + \cdots + \mu_q \qquad (\text{by (13.46.61)}) \\
&= \mu_1 + \mu_2 + \cdots + \mu_k \qquad (\text{by (13.46.63)}) .
\end{aligned}$$

This proves $\lambda_1 + \lambda_2 + \cdots + \lambda_k \geq \mu_1 + \mu_2 + \cdots + \mu_k$. Thus, (13.46.62) is proven.]

Recall that we have $\lambda \triangleright \mu$ if and only if (13.46.62) holds. Thus, we have $\lambda \triangleright \mu$ (since (13.46.62) holds). This proves Lemma 13.46.20. □

Lemma 13.46.21. *Let $n \in \mathbb{N}$ and $q \in \mathbb{N}$. Let $\lambda \in \operatorname{Par}_n$ and $\mu \in \operatorname{Par}_n$ be such that $\ell(\lambda) \leq q$. Assume that*

$$\lambda_1 + \lambda_2 + \cdots + \lambda_k \geq \mu_1 + \mu_2 + \cdots + \mu_k \qquad \text{for each } k \in \{1,2,\ldots,q\}.$$

Then, $\lambda \triangleright \mu$.

Proof of Lemma 13.46.21. We have assumed that

$$\lambda_1 + \lambda_2 + \cdots + \lambda_k \geq \mu_1 + \mu_2 + \cdots + \mu_k \qquad \text{for each } k \in \{1,2,\ldots,q\}. \tag{13.46.64}$$

Both λ and μ are partitions of n (since $\lambda \in \operatorname{Par}_n$ and $\mu \in \operatorname{Par}_n$). Thus, we have $\lambda \triangleright \mu$ if and only if

$$\lambda_1 + \lambda_2 + \cdots + \lambda_k \geq \mu_1 + \mu_2 + \cdots + \mu_k \qquad \text{for all } k \in \{1,2,\ldots,n\} \tag{13.46.65}$$

(by the definition of the dominance order).

Now, we are going to prove (13.46.65).

[*Proof of (13.46.65):* Let $k \in \{1,2,\ldots,n\}$. We must prove that $\lambda_1 + \lambda_2 + \cdots + \lambda_k \geq \mu_1 + \mu_2 + \cdots + \mu_k$.

If $k \in \{1,2,\ldots,q\}$, then this follows immediately from (13.46.64). Hence, for the rest of this proof of $\lambda_1+\lambda_2+\cdots+\lambda_k \geq \mu_1+\mu_2+\cdots+\mu_k$, we WLOG assume that $k \notin \{1,2,\ldots,q\}$. Combining $k \in \{1,2,\ldots,n\} \subset \{1,2,3,\ldots\}$ with $k \notin \{1,2,\ldots,q\}$, we obtain $k \in \{1,2,3,\ldots\} \setminus \{1,2,\ldots,q\} = \{q+1,q+2,q+3,\ldots\}$. Thus, $k \geq q+1 > q \geq \ell(\lambda)$ (since $\ell(\lambda) \leq q$).

From $\mu \in \operatorname{Par}_n$, we obtain $|\mu| = n$. Hence, $n = |\mu| = \mu_1 + \mu_2 + \mu_3 + \cdots$ (by the definition of $|\mu|$).

But the definition of $|\lambda|$ yields

$$|\lambda| = \lambda_1 + \lambda_2 + \lambda_3 + \cdots = \sum_{i=1}^{\infty} \lambda_i = \sum_{i=1}^{k} \lambda_i + \sum_{i=k+1}^{\infty} \underbrace{\lambda_i}_{\substack{=0 \\ (\text{since } i \geq k+1 > k > \ell(\lambda))}} = \sum_{i=1}^{k} \lambda_i + \underbrace{\sum_{i=k+1}^{\infty} 0}_{=0}$$

$$= \sum_{i=1}^{k} \lambda_i = \lambda_1 + \lambda_2 + \cdots + \lambda_k.$$

Thus,

$$\begin{aligned}\lambda_1+\lambda_2+\cdots+\lambda_k=|\lambda|=n \qquad &(\text{since } \lambda\in\operatorname{Par}_n)\\ &=\mu_1+\mu_2+\mu_3+\cdots=\sum_{i=1}^{\infty}\mu_i=\sum_{i=1}^{k}\mu_i+\sum_{i=k+1}^{\infty}\underbrace{\mu_i}_{\geq 0}\\ &\geq\sum_{i=1}^{k}\mu_i+\underbrace{\sum_{i=k+1}^{\infty}0}_{=0}=\sum_{i=1}^{k}\mu_i=\mu_1+\mu_2+\cdots+\mu_k.\end{aligned}$$

This proves $\lambda_1+\lambda_2+\cdots+\lambda_k\geq\mu_1+\mu_2+\cdots+\mu_k$. Thus, (13.46.65) is proven.]

Recall that we have $\lambda\triangleright\mu$ if and only if (13.46.65) holds. Thus, we have $\lambda\triangleright\mu$ (since (13.46.65) holds). This proves Lemma 13.46.21. □

Lemma 13.46.22. *Let $n\in\mathbb{N}$, $p\in\mathbb{N}$ and $q\in\mathbb{N}$. Let $\lambda\in\operatorname{Par}_n$ and $\mu\in\operatorname{Par}_n$. Let $A\in\{0,1\}^{p\times q}$ be a $\{0,1\}$-matrix having row sums λ and column sums μ. Then:*

(a) *We have $\lambda^t\triangleright\mu$.*
(b) *If $\mu=\lambda^t$, then $A=([j\leq\lambda_i])_{1\leq i\leq p,\ 1\leq j\leq q}$.*

Proof of Lemma 13.46.22. The row sums of A is λ (since A has row sums λ). Thus, λ is a p-tuple (since the row sums of A is a p-tuple). Therefore, $\lambda=(\lambda_1,\lambda_2,\ldots,\lambda_p)$. Thus, $(\lambda_1,\lambda_2,\ldots,\lambda_p)$ is the row sums of A (since λ is the row sums of A).

The column sums of A is μ (since A has column sums μ). Thus, μ is a q-tuple (since the column sums of A is a q-tuple). Therefore, $\mu=(\mu_1,\mu_2,\ldots,\mu_q)$. Thus, $(\mu_1,\mu_2,\ldots,\mu_q)$ is the column sums of A (since μ is the column sums of A). From $\mu=(\mu_1,\mu_2,\ldots,\mu_q)$, we obtain $\ell(\mu)\leq q$.

We have $|\lambda^t|=|\lambda|$ (since the transpose of a partition always has the same size as the partition itself). But $|\lambda|=n$ (since $\lambda\in\operatorname{Par}_n$). Hence, $|\lambda^t|=|\lambda|=n$, so that $\lambda^t\in\operatorname{Par}_n$.

Write the $p\times q$-matrix A in the form $A=(a_{i,j})_{1\leq i\leq p,\ 1\leq j\leq q}$. Hence, Lemma 13.46.19(e) shows that we have

$$\sum_{i=1}^{p}\min\{\lambda_i,k\}\geq\sum_{j=1}^{k}\mu_j \tag{13.46.66}$$

for each $k\in\{0,1,\ldots,q\}$.

But $\lambda\in\operatorname{Par}_n\subset\operatorname{Par}$. Hence, Lemma 13.45.1 (applied to $\nu=\lambda$) shows that

$$\left(\lambda^t\right)_1+\left(\lambda^t\right)_2+\cdots+\left(\lambda^t\right)_k=\sum_{j=1}^{\infty}\min\{\lambda_j,k\} \tag{13.46.67}$$

for every $k \in \mathbb{N}$. Hence, every $k \in \{1, 2, \ldots, q\}$ satisfies

$$\begin{aligned}
\left(\lambda^t\right)_1 + \left(\lambda^t\right)_2 + \cdots + \left(\lambda^t\right)_k &= \sum_{j=1}^{\infty} \min\{\lambda_j, k\} \qquad \text{(by (13.46.67))} \\
&= \sum_{j=1}^{p} \min\{\lambda_j, k\} + \sum_{j=p+1}^{\infty} \underbrace{\min\{\lambda_j, k\}}_{\geq 0} \\
&\geq \sum_{j=1}^{p} \min\{\lambda_j, k\} + \underbrace{\sum_{j=p+1}^{\infty} 0}_{=0} = \sum_{j=1}^{p} \min\{\lambda_j, k\} \\
&= \sum_{i=1}^{p} \min\{\lambda_i, k\} \qquad (13.46.68) \\
&\qquad \text{(here, we have renamed the summation index } j \text{ as } i\text{)} \\
&\geq \sum_{j=1}^{k} \mu_j \qquad \text{(by (13.46.66))} \\
&= \mu_1 + \mu_2 + \cdots + \mu_k.
\end{aligned}$$

Hence, Lemma 13.46.20 (applied to λ^t instead of λ) shows that $\lambda^t \rhd \mu$. This proves Lemma 13.46.22(a).

(b) Assume that $\mu = \lambda^t$.

Let $k \in \{1, 2, \ldots, q\}$. Then, (13.46.68) shows that

$$\left(\lambda^t\right)_1 + \left(\lambda^t\right)_2 + \cdots + \left(\lambda^t\right)_k \geq \sum_{i=1}^{p} \min\{\lambda_i, k\}.$$

Hence,

$$\sum_{i=1}^{p} \min\{\lambda_i, k\} \leq \left(\lambda^t\right)_1 + \left(\lambda^t\right)_2 + \cdots + \left(\lambda^t\right)_k = \sum_{j=1}^{k} \Bigg(\underbrace{\lambda^t}_{=\mu}\Bigg)_j = \sum_{j=1}^{k} \mu_j.$$

Combining this with (13.46.66), we obtain $\sum_{i=1}^{p} \min\{\lambda_i, k\} = \sum_{j=1}^{k} \mu_j$.

Now, forget that we fixed k. We thus have proven that $\left(\sum_{i=1}^{p} \min\{\lambda_i, k\} = \sum_{j=1}^{k} \mu_j \text{ for each } k \in \{1, 2, \ldots, q\}\right)$. Hence, Lemma 13.46.19(f) shows that $A = ([j \leq \lambda_i])_{1 \leq i \leq p,\ 1 \leq j \leq q}$. This proves Lemma 13.46.22(b). □

Lemma 13.46.23. *Let $p \in \mathbb{N}$ and $q \in \mathbb{N}$. Let λ be a partition satisfying $\ell(\lambda) \leq p$ and $\lambda_1 \leq q$.*

Let B be the $p \times q$-matrix $([j \leq \lambda_i])_{1 \leq i \leq p,\ 1 \leq j \leq q}$.

Then, B is a $\{0,1\}$-matrix having row sums λ and column sums λ^t.

Proof of Lemma 13.46.23. We have $[j \leq \lambda_i] \in \{0,1\}$ for each $i \in \{1, 2, \ldots, p\}$ and $j \in \{1, 2, \ldots, q\}$ (since $[\mathcal{A}] \in \{0,1\}$ for any logical statement $\mathcal{A}$). Hence, $([j \leq \lambda_i])_{1 \leq i \leq p,\ 1 \leq j \leq q} \in \{0,1\}^{p \times q}$.

Thus, $B = ([j \leq \lambda_i])_{1 \leq i \leq p,\ 1 \leq j \leq q} \in \{0,1\}^{p \times q}$. Therefore, B is a $\{0,1\}$-matrix.

From $\ell(\lambda) \leq p$, we obtain $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_p)$.

Let $\left(\widetilde{\lambda}_1, \widetilde{\lambda}_2, \ldots, \widetilde{\lambda}_p\right)$ be the row sums of B. Let $(\widetilde{\mu}_1, \widetilde{\mu}_2, \ldots, \widetilde{\mu}_q)$ be the column sums of B.

Lemma 13.46.19(a) (applied to B, $\widetilde{\lambda}_i$, $\widetilde{\mu}_j$ and $[j \leq \lambda_i]$ instead of A, λ_i, μ_j and $a_{i,j}$) shows that we have

$$\widetilde{\lambda}_i = \sum_{j=1}^{q} [j \leq \lambda_i] \qquad \text{for each } i \in \{1, 2, \ldots, p\} \tag{13.46.69}$$

(since $B = ([j \leq \lambda_i])_{1 \leq i \leq p,\ 1 \leq j \leq q}$).

Lemma 13.46.19(b) (applied to B, $\widetilde{\lambda}_i$, $\widetilde{\mu}_j$ and $[j \leq \lambda_i]$ instead of A, λ_i, μ_j and $a_{i,j}$) shows that we have

$$\widetilde{\mu}_j = \sum_{i=1}^{p} [j \leq \lambda_i] \qquad \text{for each } j \in \{1, 2, \ldots, q\} \tag{13.46.70}$$

(since $B = ([j \leq \lambda_i])_{1 \leq i \leq p,\ 1 \leq j \leq q}$).

We have

$$\lambda_i \leq q \tag{13.46.71}$$

for each $i \in \{1, 2, 3, \ldots\}$ [540].

Let $i \in \{1, 2, \ldots, p\}$. Then, $\lambda_i \leq q$ (by (13.46.71)). Hence, Lemma 13.46.18 (applied to $r = \lambda_i$) yields $\sum_{j=1}^{q} [j \leq \lambda_i] = \lambda_i$. Hence, (13.46.69) becomes $\widetilde{\lambda}_i = \sum_{j=1}^{q} [j \leq \lambda_i] = \lambda_i$.

Now, forget that we fixed i. We thus have shown that $\widetilde{\lambda}_i = \lambda_i$ for each $i \in \{1, 2, \ldots, p\}$. In other words, $\left(\widetilde{\lambda}_1, \widetilde{\lambda}_2, \ldots, \widetilde{\lambda}_p\right) = (\lambda_1, \lambda_2, \ldots, \lambda_p) = \lambda$ (since $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_p)$). Thus, the row sums of B is λ (since the row sums of B is $\left(\widetilde{\lambda}_1, \widetilde{\lambda}_2, \ldots, \widetilde{\lambda}_p\right)$). In other words, the matrix B has row sums λ.

On the other hand, the definition (2.2.7) of the conjugate partition λ^t of λ shows that

$$\left(\lambda^t\right)_i = |\{j \in \{1, 2, 3, \ldots\} \mid \lambda_j \geq i\}| \tag{13.46.72}$$

for every positive integer i.

Now, let $k \in \{1, 2, \ldots, q\}$. Applying (13.46.72) to $i = k$, we obtain

$$\left(\lambda^t\right)_k = |\{j \in \{1, 2, 3, \ldots\} \mid \lambda_j \geq k\}| . \tag{13.46.73}$$

But (13.46.70) (applied to $j = k$) shows that

$$\begin{aligned}
\widetilde{\mu}_k &= \underbrace{\sum_{i=1}^{p}}_{=\sum_{i \in \{1,2,\ldots,p\}}} \Bigg[\underbrace{k \leq \lambda_i}_{\iff (\lambda_i \geq k)}\Bigg] = \sum_{i \in \{1,2,\ldots,p\}} [\lambda_i \geq k] \\
&= \sum_{\substack{i \in \{1,2,\ldots,p\}; \\ \lambda_i \geq k}} \underbrace{[\lambda_i \geq k]}_{\substack{=1 \\ (\text{since } \lambda_i \geq k)}} + \sum_{\substack{i \in \{1,2,\ldots,p\}; \\ \text{not } \lambda_i \geq k}} \underbrace{[\lambda_i \geq k]}_{\substack{=0 \\ (\text{since we don't have } \lambda_i \geq k)}} \\
&= \sum_{\substack{i \in \{1,2,\ldots,p\}; \\ \lambda_i \geq k}} 1 + \underbrace{\sum_{\substack{i \in \{1,2,\ldots,p\}; \\ \text{not } \lambda_i \geq k}} 0}_{=0} = \sum_{\substack{i \in \{1,2,\ldots,p\}; \\ \lambda_i \geq k}} 1 \\
&= |\{i \in \{1, 2, \ldots, p\} \mid \lambda_i \geq k\}| \cdot 1 = |\{i \in \{1, 2, \ldots, p\} \mid \lambda_i \geq k\}| \\
&= |\{j \in \{1, 2, \ldots, p\} \mid \lambda_j \geq k\}| \qquad (13.46.74)
\end{aligned}$$

(here, we have renamed the index i as j). But

$$\{j \in \{1, 2, 3, \ldots\} \mid \lambda_j \geq k\} = \{j \in \{1, 2, \ldots, p\} \mid \lambda_j \geq k\} \tag{13.46.75}$$

[540] *Proof of (13.46.71):* Let $i \in \{1, 2, 3, \ldots\}$. Thus, $i \geq 1$, so that $1 \leq i$.

The sequence λ is a partition, and thus is weakly decreasing. In other words, $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \cdots$. Hence, $\lambda_1 \geq \lambda_i$ (since $1 \leq i$). Thus, $\lambda_i \leq \lambda_1 \leq q$. This proves (13.46.71).

[541]. Hence, (13.46.73) becomes

$$\left(\lambda^t\right)_k = \left| \underbrace{\left\{j \in \{1,2,3,\ldots\} \mid \lambda_j \geq k\right\}}_{\substack{=\{j\in\{1,2,\ldots,p\} \mid \lambda_j \geq k\} \\ \text{(by (13.46.75))}}} \right| = \left|\left\{j \in \{1,2,\ldots,p\} \mid \lambda_j \geq k\right\}\right| = \widetilde{\mu}_k$$

(by (13.46.74)).

Now, forget that we fixed k. We thus have shown that $\left(\lambda^t\right)_k = \widetilde{\mu}_k$ for each $k \in \{1,2,\ldots,q\}$. In other words,

$$\left(\left(\lambda^t\right)_1, \left(\lambda^t\right)_2, \ldots, \left(\lambda^t\right)_q\right) = \left(\widetilde{\mu}_1, \widetilde{\mu}_2, \ldots, \widetilde{\mu}_q\right).$$

But $\left(\lambda^t\right)_i = 0$ for each integer $i > q$ [542]. In other words, $\left(\lambda^t\right)_{q+1} = \left(\lambda^t\right)_{q+2} = \left(\lambda^t\right)_{q+3} = \cdots = 0$. Hence, $\lambda^t = \left(\left(\lambda^t\right)_1, \left(\lambda^t\right)_2, \ldots, \left(\lambda^t\right)_q\right)$ (since we omit trailing zeroes from a partition). Thus,

$$\lambda^t = \left(\left(\lambda^t\right)_1, \left(\lambda^t\right)_2, \ldots, \left(\lambda^t\right)_q\right) = \left(\widetilde{\mu}_1, \widetilde{\mu}_2, \ldots, \widetilde{\mu}_q\right) = (\text{the column sums of } B)$$

(since $\left(\widetilde{\mu}_1, \widetilde{\mu}_2, \ldots, \widetilde{\mu}_q\right)$ is the column sums of B). In other words, the matrix B has column sums λ^t.

Hence, we know that the matrix B has row sums λ and column sums λ^t. Thus, B is a $\{0,1\}$-matrix having row sums λ and column sums λ^t (since B is a $\{0,1\}$-matrix). This proves Lemma 13.46.23. □

Lemma 13.46.24. *Let λ be a partition. Then:*

(a) *We have $\left(\lambda^t\right)_1 = \ell(\lambda)$.*
(b) *We have $\lambda_1 = \ell\left(\lambda^t\right)$.*

Proof of Lemma 13.46.24. (a) Write the partition λ in the form $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_k, 0, 0, 0, \ldots)$ with $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k > 0$. Then, $\ell(\lambda) = k$ (by the definition of $\ell(\lambda)$). Thus, we have

$$\lambda_i = 0 \qquad \text{for each integer } i > \ell(\lambda) \tag{13.46.76}$$

(by the definition of $\ell(\lambda)$). Thus,

$$\left\{j \in \{1,2,3,\ldots\} \mid \lambda_j \geq 1\right\} = \{1,2,\ldots,k\} \tag{13.46.77}$$

[543].

[541] *Proof of (13.46.75):* Let $i \in \left\{j \in \{1,2,3,\ldots\} \mid \lambda_j \geq k\right\}$. Thus, i is an element of $\{1,2,3,\ldots\}$ and satisfies $\lambda_i \geq k$.

We have $\lambda_i \geq k > 0$ (since $k \in \{1,2,\ldots,q\}$). Hence, $\lambda_i \neq 0$.

Assume (for the sake of contradiction) that $i > p$. Hence, $i > p \geq \ell(\lambda)$ (since $\ell(\lambda) \leq p$). But each integer $j > \ell(\lambda)$ satisfies $\lambda_j = 0$ (by the definition of $\ell(\lambda)$). Applying this to $j = i$, we obtain $\lambda_i = 0$ (since $i > \ell(\lambda)$). This contradicts $\lambda_i \neq 0$. This contradiction shows that our assumption (that $i > p$) was wrong. Hence, we must have $i \leq p$. Thus, $i \in \{1,2,\ldots,p\}$.

Now, forget that we fixed i. We thus have shown that each $i \in \left\{j \in \{1,2,3,\ldots\} \mid \lambda_j \geq k\right\}$ satisfies $i \in \{1,2,\ldots,p\}$. In other words, $\left\{j \in \{1,2,3,\ldots\} \mid \lambda_j \geq k\right\} \subset \{1,2,\ldots,p\}$. Hence,

$$\begin{aligned}\left\{j \in \{1,2,3,\ldots\} \mid \lambda_j \geq k\right\} &= \{1,2,\ldots,p\} \cap \left\{j \in \{1,2,3,\ldots\} \mid \lambda_j \geq k\right\} \\ &= \left\{j \in \underbrace{\{1,2,\ldots,p\} \cap \{1,2,3,\ldots\}}_{=\{1,2,\ldots,p\}} \mid \lambda_j \geq k\right\} = \left\{j \in \{1,2,\ldots,p\} \mid \lambda_j \geq k\right\}.\end{aligned}$$

This proves (13.46.75).

[542] *Proof.* Let $i > q$ be an integer. Then, $q < i$. But each $j \in \{1,2,3,\ldots\}$ satisfies $\lambda_j \leq q$ (by (13.46.71), applied to j instead of i). Hence, each $j \in \{1,2,3,\ldots\}$ satisfies $\lambda_j < i$ (since $\lambda_j \leq q < i$). In other words, no $j \in \{1,2,3,\ldots\}$ satisfies $\lambda_j \geq i$. In other words, $\left\{j \in \{1,2,3,\ldots\} \mid \lambda_j \geq i\right\} = \varnothing$. Hence, (13.46.72) becomes $\left(\lambda^t\right)_i = \left|\underbrace{\left\{j \in \{1,2,3,\ldots\} \mid \lambda_j \geq i\right\}}_{=\varnothing}\right| = |\varnothing| = 0$.

[543] *Proof of (13.46.77):* Let $i \in \left\{j \in \{1,2,3,\ldots\} \mid \lambda_j \geq 1\right\}$. Thus, i is an element of $\{1,2,3,\ldots\}$ and satisfies $\lambda_i \geq 1$. Thus, $\lambda_i \geq 1 > 0$, so that $\lambda_i \neq 0$.

If we had $i > \ell(\lambda)$, then we would have $\lambda_i = 0$ (by (13.46.76)), which would contradict $\lambda_i \neq 0$. Hence, we cannot have $i > \ell(\lambda)$. Thus, we must have $i \leq \ell(\lambda)$. Hence, $i \leq \ell(\lambda) = k$, so that $i \in \{1,2,\ldots,k\}$.

Now, forget that we fixed i. We thus have shown that $i \in \{1,2,\ldots,k\}$ for each $i \in \left\{j \in \{1,2,3,\ldots\} \mid \lambda_j \geq 1\right\}$. In other words, $\left\{j \in \{1,2,3,\ldots\} \mid \lambda_j \geq 1\right\} \subset \{1,2,\ldots,k\}$.

The definition (2.2.7) of the conjugate partition λ^t of λ shows that $(\lambda^t)_i = |\{j \in \{1,2,3,\ldots\} \mid \lambda_j \geq i\}|$ for each positive integer i. Applying this to $i = 1$, we obtain

$$\left(\lambda^t\right)_1 = \left| \underbrace{\{j \in \{1,2,3,\ldots\} \mid \lambda_j \geq 1\}}_{\substack{=\{1,2,\ldots,k\} \\ \text{(by (13.46.77))}}} \right| = |\{1,2,\ldots,k\}| = k = \ell(\lambda).$$

This proves Lemma 13.46.24(a).

(b) It is known that $(\lambda^t)^t = \lambda$. But Lemma 13.46.24(a) (applied to λ^t instead of λ) yields $\left((\lambda^t)^t\right)_1 = \ell(\lambda^t)$. In light of $(\lambda^t)^t = \lambda$, this rewrites as $\lambda_1 = \ell(\lambda^t)$. This proves Lemma 13.46.24(b). □

Now, let us resume the solution of Exercise 2.2.13.

(h) Let $\lambda \in \operatorname{Par}_n$ and $\mu \in \operatorname{Par}_n$ be two partitions that don't satisfy $\lambda^t \rhd \mu$. We must prove that $a_{\lambda,\mu} = 0$.

Indeed, there exists no $\{0,1\}$-matrix of size $\ell(\lambda) \times \ell(\mu)$ having row sums λ and column sums μ [544]. Thus, the number of all $\{0,1\}$-matrices of size $\ell(\lambda) \times \ell(\mu)$ having row sums λ and column sums μ equals 0. In other words, $a_{\lambda,\mu}$ equals 0 (since $a_{\lambda,\mu}$ is the number of all $\{0,1\}$-matrices of size $\ell(\lambda) \times \ell(\mu)$ having row sums λ and column sums μ). In other words, $a_{\lambda,\mu} = 0$. This solves Exercise 2.2.13(h).

(i) Let $\lambda \in \operatorname{Par}_n$. We shall prove that $a_{\lambda,\lambda^t} = 1$.

Let $p = \ell(\lambda)$ and $q = \lambda_1$. Then, $\ell(\lambda) = p \leq p$ and $\lambda_1 = q \leq q$. Also, $q = \lambda_1 = \ell(\lambda^t)$ (by Lemma 13.46.24(b)).

Define the $p \times q$-matrix B as in Lemma 13.46.23. Then, Lemma 13.46.23 shows that B is a $\{0,1\}$-matrix having row sums λ and column sums λ^t. Furthermore, B is a $\{0,1\}$-matrix of size $p \times q$. In other words, B is a $\{0,1\}$-matrix of size $\ell(\lambda) \times \ell(\lambda^t)$ (since $p = \ell(\lambda)$ and $q = \ell(\lambda^t)$). Hence, there exists **at least one** $\{0,1\}$-matrix of size $\ell(\lambda) \times \ell(\lambda^t)$ having row sums λ and column sums λ^t (namely, B).

On the other hand, using Lemma 13.46.22(b), it is easy to see that every $\{0,1\}$-matrix of size $\ell(\lambda) \times \ell(\lambda^t)$ having row sums λ and column sums λ^t must be equal to B [545]. Hence, there exists **at most one** $\{0,1\}$-matrix of size $\ell(\lambda) \times \ell(\lambda^t)$ having row sums λ and column sums λ^t.

We know that a_{λ,λ^t} is the number of all $\{0,1\}$-matrices of size $\ell(\lambda) \times \ell(\lambda^t)$ having row sums λ and column sums λ^t (by the definition of a_{λ,λ^t}). Since there exists **exactly one** such matrix (because we have shown that there exists **at least one** such matrix, and we have also shown that there exists **at most one** such matrix), we thus conclude that $a_{\lambda,\lambda^t} = 1$.

Now, forget that we fixed λ. We thus have shown that

$$a_{\lambda,\lambda^t} = 1 \qquad \text{for every } \lambda \in \operatorname{Par}_n. \tag{13.46.78}$$

Now, let $\lambda \in \operatorname{Par}_n$. We shall show that $a_{\lambda^t,\lambda} = 1$.

It is known that $(\lambda^t)^t = \lambda$. But $|\lambda^t| = |\lambda|$ (since the transpose of a partition always has the same size as the partition itself). But $|\lambda| = n$ (since $\lambda \in \operatorname{Par}_n$). Hence, $|\lambda^t| = |\lambda| = n$, so that $\lambda^t \in \operatorname{Par}_n$. Thus,

On the other hand, let $h \in \{1,2,\ldots,k\}$. Then, $\lambda_h > 0$ (since $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k > 0$). Hence, $\lambda_h \geq 1$ (since λ_h is an integer). Hence, h is an element of $\{1,2,3,\ldots\}$ and satisfies $\lambda_h \geq 1$. In other words, $h \in \{j \in \{1,2,3,\ldots\} \mid \lambda_j \geq 1\}$.

Now, forget that we fixed h. We thus have shown that $h \in \{j \in \{1,2,3,\ldots\} \mid \lambda_j \geq 1\}$ for each $h \in \{1,2,\ldots,k\}$. In other words, $\{1,2,\ldots,k\} \subset \{j \in \{1,2,3,\ldots\} \mid \lambda_j \geq 1\}$. Combining this with $\{j \in \{1,2,3,\ldots\} \mid \lambda_j \geq 1\} \subset \{1,2,\ldots,k\}$, we obtain $\{j \in \{1,2,3,\ldots\} \mid \lambda_j \geq 1\} = \{1,2,\ldots,k\}$. This proves (13.46.77).

[544] *Proof.* Let A be a $\{0,1\}$-matrix of size $\ell(\lambda) \times \ell(\mu)$ having row sums λ and column sums μ. We shall derive a contradiction.

We have $A \in \{0,1\}^{\ell(\lambda) \times \ell(\mu)}$ (since A is a $\{0,1\}$-matrix of size $\ell(\lambda) \times \ell(\mu)$). Thus, Lemma 13.46.22(a) shows that $\lambda^t \rhd \mu$. This contradicts the fact that we don't have $\lambda^t \rhd \mu$.

Now, forget that we fixed A. We thus have found a contradiction for each $\{0,1\}$-matrix A of size $\ell(\lambda) \times \ell(\mu)$ having row sums λ and column sums μ. Thus, there exists no $\{0,1\}$-matrix of size $\ell(\lambda) \times \ell(\mu)$ having row sums λ and column sums μ. Qed.

[545] *Proof.* Let A be a $\{0,1\}$-matrix of size $\ell(\lambda) \times \ell(\lambda^t)$ having row sums λ and column sums λ^t. We must prove that $A = B$.

We know that A is a $\{0,1\}$-matrix of size $\ell(\lambda) \times \ell(\lambda^t)$. In other words, A is a $\{0,1\}$-matrix of size $p \times q$ (since $\ell(\lambda) = p$ and $\ell(\lambda^t) = q$). In other words, $A \in \{0,1\}^{p \times q}$.

We have $|\lambda^t| = |\lambda|$ (since the transpose of a partition always has the same size as the partition itself). But $|\lambda| = n$ (since $\lambda \in \operatorname{Par}_n$). Hence, $|\lambda^t| = |\lambda| = n$, so that $\lambda^t \in \operatorname{Par}_n$. Hence, Lemma 13.46.22(b) (applied to $\mu = \lambda^t$) shows that $A = ([j \leq \lambda_i])_{1 \leq i \leq p,\ 1 \leq j \leq q}$. But the definition of B yields $B = ([j \leq \lambda_i])_{1 \leq i \leq p,\ 1 \leq j \leq q}$. Comparing this with $A = ([j \leq \lambda_i])_{1 \leq i \leq p,\ 1 \leq j \leq q}$, we obtain $A = B$. This proves $A = B$, qed.

(13.46.78) (applied to λ^t instead of λ) yields $a_{\lambda^t,(\lambda^t)^t} = 1$. In light of $(\lambda^t)^t = \lambda$, this rewrites as $a_{\lambda^t,\lambda} = 1$. This solves Exercise 2.2.13(i).

Before we solve Exercise 2.2.13(j), let us introduce some terminology and prove some lemmas.

Definition 13.46.25. Let $m \in \mathbb{N}$. Then, $[m]$ shall denote the subset $\{1, 2, \ldots, m\}$ of $\{1, 2, 3, \ldots\}$. Notice that $|[m]| = m$ for each $m \in \mathbb{N}$.

Definition 13.46.26. Let $p \in \mathbb{N}$ and $q \in \mathbb{N}$. Let $\varphi : [p] \to [q]$ be any map. Let $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_p) \in \mathbb{N}^p$ be a p-tuple. Then, we define a q-tuple $\varphi_* \alpha \in \mathbb{N}^q$ by

$$\varphi_* \alpha = \left(\sum_{i \in \varphi^{-1}\{1\}} \alpha_i, \sum_{i \in \varphi^{-1}\{2\}} \alpha_i, \ldots, \sum_{i \in \varphi^{-1}\{q\}} \alpha_i \right).$$

(Recall that if U is any subset of $[q]$, then $\varphi^{-1}U$ denotes the subset $\{i \in [p] \mid \varphi(i) \in U\}$ of $[p]$.)

Example 13.46.27. Let $p = 5$ and $q = 4$, and let $\varphi : [p] \to [q]$ be the map that sends $1, 2, 3, 4, 5$ to $1, 4, 4, 2, 2$, respectively. Let $\alpha = (\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5) \in \mathbb{N}^5$ be a 5-tuple. Then, the 4-tuple $\varphi_* \alpha \in \mathbb{N}^4$ is

$$\varphi_* \alpha = \left(\sum_{i \in \varphi^{-1}\{1\}} \alpha_i, \sum_{i \in \varphi^{-1}\{2\}} \alpha_i, \sum_{i \in \varphi^{-1}\{3\}} \alpha_i, \sum_{i \in \varphi^{-1}\{4\}} \alpha_i \right) = (\alpha_1, \alpha_4 + \alpha_5, 0, \alpha_2 + \alpha_3)$$

(since $\varphi^{-1}\{1\} = \{1\}$, $\varphi^{-1}\{2\} = \{4, 5\}$, $\varphi^{-1}\{3\} = \varnothing$ and $\varphi^{-1}\{4\} = \{2, 3\}$).

Proposition 13.46.28. *Let p, q and r be three elements of $\mathbb{N}$. Let $\varphi : [p] \to [q]$ and $\psi : [q] \to [r]$ be any maps. Let $\alpha \in \mathbb{N}^p$. Then, $(\psi \circ \varphi)_* \alpha = \psi_* (\varphi_* \alpha)$.*

Proof of Proposition 13.46.28. Write the p-tuple $\alpha \in \mathbb{N}^p$ in the form $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_p)$ for some elements $\alpha_1, \alpha_2, \ldots, \alpha_p$ of $\mathbb{N}$. Thus,

$$(\psi \circ \varphi)_* \alpha = \left(\sum_{i \in (\psi \circ \varphi)^{-1}\{1\}} \alpha_i, \sum_{i \in (\psi \circ \varphi)^{-1}\{2\}} \alpha_i, \ldots, \sum_{i \in (\psi \circ \varphi)^{-1}\{r\}} \alpha_i \right) \tag{13.46.79}$$

(by the definition of $(\psi \circ \varphi)_* \alpha$).

Write the q-tuple $\varphi_* \alpha \in \mathbb{N}^q$ in the form $\varphi_* \alpha = (\beta_1, \beta_2, \ldots, \beta_q)$ for some elements $\beta_1, \beta_2, \ldots, \beta_q$ of $\mathbb{N}$. Then,

$$(\beta_1, \beta_2, \ldots, \beta_q) = \varphi_* \alpha = \left(\sum_{i \in \varphi^{-1}\{1\}} \alpha_i, \sum_{i \in \varphi^{-1}\{2\}} \alpha_i, \ldots, \sum_{i \in \varphi^{-1}\{q\}} \alpha_i \right)$$

(by the definition of $\varphi_* \alpha$, because $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_p)$). Thus,

$$\beta_j = \sum_{i \in \varphi^{-1}\{j\}} \alpha_i \qquad \text{for each } j \in [q]. \tag{13.46.80}$$

Write the r-tuple $\psi_* (\varphi_* \alpha) \in \mathbb{N}^r$ in the form $\psi_* (\varphi_* \alpha) = (\gamma_1, \gamma_2, \ldots, \gamma_r)$ for some elements $\gamma_1, \gamma_2, \ldots, \gamma_r$ of $\mathbb{N}$. Then,

$$(\gamma_1, \gamma_2, \ldots, \gamma_r) = \psi_* (\varphi_* \alpha) = \left(\sum_{i \in \psi^{-1}\{1\}} \beta_i, \sum_{i \in \psi^{-1}\{2\}} \beta_i, \ldots, \sum_{i \in \psi^{-1}\{r\}} \beta_i \right)$$

(by the definition of $\psi_* (\varphi_* \alpha)$, because $\varphi_* \alpha = (\beta_1, \beta_2, \ldots, \beta_q)$). Thus, for each $m \in [r]$, we have

$$\begin{aligned}
\gamma_m &= \underbrace{\sum_{i \in \psi^{-1}\{m\}}}_{= \sum_{\substack{i \in [q];\\ \psi(i)=m}}} \beta_i = \sum_{\substack{i \in [q];\\ \psi(i)=m}} \beta_i = \sum_{\substack{j \in [q];\\ \psi(j)=m}} \underbrace{\beta_j}_{\substack{= \sum_{i \in \varphi^{-1}\{j\}} \alpha_i \\ \text{(by (13.46.80))}}} \qquad \left(\begin{array}{c} \text{here, we have renamed the} \\ \text{summation index } i \text{ as } j \end{array} \right) \\
&= \sum_{\substack{j \in [q];\\ \psi(j)=m}} \underbrace{\sum_{i \in \varphi^{-1}\{j\}}}_{= \sum_{\substack{i \in [p];\\ \varphi(i)=j}}} \alpha_i = \underbrace{\sum_{\substack{j \in [q];\\ \psi(j)=m}} \sum_{\substack{i \in [p];\\ \varphi(i)=j}}}_{= \sum_{j \in [q]} \sum_{\substack{i \in [p];\\ \psi(j)=m;\\ \varphi(i)=j}}} \alpha_i \\
&= \sum_{j \in [q]} \underbrace{\sum_{\substack{i \in [p];\\ \psi(j)=m;\\ \varphi(i)=j}}}_{\substack{= \sum_{\substack{i \in [p];\\ \varphi(i)=j;\\ \psi(j)=m}} = \sum_{\substack{i \in [p];\\ \varphi(i)=j;\\ \psi(\varphi(i))=m}} \\ \text{(because for any } i \in [p] \text{ satisfying } \varphi(i)=j, \\ \text{the condition } (\psi(j)=m) \text{ is equivalent to } (\psi(\varphi(i))=m) \\ \text{(since } j=\varphi(i)))}} \alpha_i = \underbrace{\sum_{j \in [q]} \sum_{\substack{i \in [p];\\ \varphi(i)=j;\\ \psi(\varphi(i))=m}}}_{\substack{= \sum_{i \in [p]} \sum_{\substack{j \in [q];\\ \varphi(i)=j;\\ \psi(\varphi(i))=m}} \\ = \sum_{\substack{i \in [p];\\ \psi(\varphi(i))=m}} \sum_{\substack{j \in [q];\\ \varphi(i)=j}}}} \alpha_i \\
&= \underbrace{\sum_{\substack{i \in [p];\\ \psi(\varphi(i))=m}}}_{= \sum_{\substack{i \in [p];\\ (\psi \circ \varphi)(i)=m}} = \sum_{i \in (\psi \circ \varphi)^{-1}\{m\}}} \underbrace{\sum_{\substack{j \in [q];\\ \varphi(i)=j}} \alpha_i}_{\substack{= \alpha_i \\ \text{(since there is a} \\ \text{unique } j \in [q] \text{ satisfying } \varphi(i)=j)}} = \sum_{i \in (\psi \circ \varphi)^{-1}\{m\}} \alpha_i .
\end{aligned}$$

In other words, we have

$$(\gamma_1, \gamma_2, \ldots, \gamma_r) = \left(\sum_{i \in (\psi \circ \varphi)^{-1}\{1\}} \alpha_i, \sum_{i \in (\psi \circ \varphi)^{-1}\{2\}} \alpha_i, \ldots, \sum_{i \in (\psi \circ \varphi)^{-1}\{r\}} \alpha_i \right).$$

Comparing this with (13.46.79), we obtain

$$(\psi \circ \varphi)_* \alpha = (\gamma_1, \gamma_2, \ldots, \gamma_r) = \psi_* (\varphi_* \alpha)$$

(since $\psi_* (\varphi_* \alpha) = (\gamma_1, \gamma_2, \ldots, \gamma_r)$). This proves Proposition 13.46.28. □

Proposition 13.46.29. *Let $p \in \mathbb{N}$. Let $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_p) \in \mathbb{N}^p$. Let $\sigma \in \mathfrak{S}_p$. Then:*

(a) *The p-tuple $\sigma_* \alpha$ is well-defined and satisfies $\sigma_* \alpha = \left(\alpha_{\sigma^{-1}(1)}, \alpha_{\sigma^{-1}(2)}, \ldots, \alpha_{\sigma^{-1}(p)} \right)$.*

(b) *We have $\sigma_* \alpha = \alpha$ if and only if $\left(\alpha_{\sigma(i)} = \alpha_i \text{ for all } i \in [p] \right)$.*

Proof of Proposition 13.46.29. We have $\sigma \in \mathfrak{S}_p$. In other words, σ is a permutation of $[p]$ (since $\mathfrak{S}_p$ is the set of all permutations of $[p]$ (because $[p] = \{1, 2, \ldots, p\}$)). In other words, σ is a bijection $[p] \to [p]$. Thus, the p-tuple $\sigma_* \alpha$ is well-defined (since $\alpha \in \mathbb{N}^p$). The definition of $\sigma_* \alpha$ yields

$$\sigma_* \alpha = \left(\sum_{i \in \sigma^{-1}\{1\}} \alpha_i, \sum_{i \in \sigma^{-1}\{2\}} \alpha_i, \ldots, \sum_{i \in \sigma^{-1}\{p\}} \alpha_i \right). \tag{13.46.81}$$

But each $j \in [p]$ satisfies

$$\begin{aligned}
\sum_{i \in \sigma^{-1}\{j\}} \alpha_i &= \sum_{i \in \{\sigma^{-1}(j)\}} \alpha_i \qquad \left(\text{since } \sigma^{-1}\{j\} = \left\{ \sigma^{-1}(j) \right\} \text{ (because } \sigma \text{ is a bijection)} \right) \\
&= \alpha_{\sigma^{-1}(j)}.
\end{aligned}$$

In other words,

$$\left(\sum_{i\in\sigma^{-1}\{1\}}\alpha_i,\sum_{i\in\sigma^{-1}\{2\}}\alpha_i,\ldots,\sum_{i\in\sigma^{-1}\{p\}}\alpha_i\right)=\left(\alpha_{\sigma^{-1}(1)},\alpha_{\sigma^{-1}(2)},\ldots,\alpha_{\sigma^{-1}(p)}\right).$$

Hence, (13.46.81) becomes

$$\sigma_*\alpha=\left(\sum_{i\in\sigma^{-1}\{1\}}\alpha_i,\sum_{i\in\sigma^{-1}\{2\}}\alpha_i,\ldots,\sum_{i\in\sigma^{-1}\{p\}}\alpha_i\right)=\left(\alpha_{\sigma^{-1}(1)},\alpha_{\sigma^{-1}(2)},\ldots,\alpha_{\sigma^{-1}(p)}\right).$$

This completes the proof of Proposition 13.46.29(a).

(b) We have the following chain of logical equivalences:

$$\begin{aligned}
&\left(\underbrace{\sigma_*\alpha}_{=\left(\alpha_{\sigma^{-1}(1)},\alpha_{\sigma^{-1}(2)},\ldots,\alpha_{\sigma^{-1}(p)}\right)}=\underbrace{\alpha}_{=(\alpha_1,\alpha_2,\ldots,\alpha_p)}\right)\\
\Longleftrightarrow\ &\left(\left(\alpha_{\sigma^{-1}(1)},\alpha_{\sigma^{-1}(2)},\ldots,\alpha_{\sigma^{-1}(p)}\right)=(\alpha_1,\alpha_2,\ldots,\alpha_p)\right)\\
\Longleftrightarrow\ &\left(\alpha_{\sigma^{-1}(j)}=\alpha_j\text{ for all }j\in[p]\right)\\
\Longleftrightarrow\ &\left(\underbrace{\alpha_{\sigma^{-1}(\sigma(i))}}_{\substack{=\alpha_i\\(\text{since }\sigma^{-1}(\sigma(i))=i)}}=\alpha_{\sigma(i)}\text{ for all }i\in[p]\right)\\
&\qquad(\text{here, we have substituted }\sigma(i)\text{ for }j\text{, since the map }\sigma:[p]\to[p]\text{ is a bijection})\\
\Longleftrightarrow\ &\left(\alpha_i=\alpha_{\sigma(i)}\text{ for all }i\in[p]\right)\\
\Longleftrightarrow\ &\left(\alpha_{\sigma(i)}=\alpha_i\text{ for all }i\in[p]\right).
\end{aligned}$$

This proves Proposition 13.46.29(b). □

Proposition 13.46.30. *Let $p\in\mathbb{N}$. Let $\alpha\in\mathbb{N}^p$. Then, $\left(\mathrm{id}_{[p]}\right)_*\alpha=\alpha$.*

Proof of Proposition 13.46.30. Clearly, $\mathrm{id}_{[p]}\in\mathfrak{S}_p$. Write the p-tuple $\alpha\in\mathbb{N}^p$ in the form $\alpha=(\alpha_1,\alpha_2,\ldots,\alpha_p)$ for some elements $\alpha_1,\alpha_2,\ldots,\alpha_p$ of $\mathbb{N}$. Thus, Proposition 13.46.29(a) (applied to $\sigma=\mathrm{id}_{[p]}$) shows that the p-tuple $\left(\mathrm{id}_{[p]}\right)_*\alpha$ is well-defined and satisfies $\left(\mathrm{id}_{[p]}\right)_*\alpha=\left(\alpha_{\left(\mathrm{id}_{[p]}\right)^{-1}(1)},\alpha_{\left(\mathrm{id}_{[p]}\right)^{-1}(2)},\ldots,\alpha_{\left(\mathrm{id}_{[p]}\right)^{-1}(p)}\right)$. Hence,

$$\begin{aligned}
\left(\mathrm{id}_{[p]}\right)_*\alpha&=\left(\alpha_{\left(\mathrm{id}_{[p]}\right)^{-1}(1)},\alpha_{\left(\mathrm{id}_{[p]}\right)^{-1}(2)},\ldots,\alpha_{\left(\mathrm{id}_{[p]}\right)^{-1}(p)}\right)\\
&=(\alpha_1,\alpha_2,\ldots,\alpha_p)\qquad\left(\text{since }\left(\mathrm{id}_{[p]}\right)^{-1}(i)=i\text{ for all }i\in[p]\right)\\
&=\alpha.
\end{aligned}$$

This proves Proposition 13.46.30. □

Lemma 13.46.31. *Let $p\in\mathbb{N}$ and $q\in\mathbb{N}$. Let $\varphi:[p]\to[q]$ be any map. Let $\alpha=(\alpha_1,\alpha_2,\ldots,\alpha_p)\in\mathbb{N}^p$ be a p-tuple. Let $\beta=(\beta_1,\beta_2,\ldots,\beta_q)\in\mathbb{N}^q$ be a q-tuple such that ($\beta_i>0$ for each $i\in[q]$) and $\beta=\varphi_*\alpha$. Then, the map φ is surjective.*

Proof of Lemma 13.46.31. We have

$$(\beta_1,\beta_2,\ldots,\beta_q)=\beta=\varphi_*\alpha=\left(\sum_{i\in\varphi^{-1}\{1\}}\alpha_i,\sum_{i\in\varphi^{-1}\{2\}}\alpha_i,\ldots,\sum_{i\in\varphi^{-1}\{q\}}\alpha_i\right)$$

(by the definition of $\varphi_*\alpha$, since $\alpha=(\alpha_1,\alpha_2,\ldots,\alpha_p)$). In other words,

$$\beta_j=\sum_{i\in\varphi^{-1}\{j\}}\alpha_i\qquad\text{for each }j\in[q].\tag{13.46.82}$$

Let $j \in [q]$. Assume (for the sake of contradiction) that $\varphi^{-1}\{j\} = \varnothing$. Thus, $\sum_{i \in \varphi^{-1}\{j\}} \alpha_i = \sum_{i \in \varnothing} \alpha_i = (\text{empty sum}) = 0$. Thus, (13.46.82) becomes $\beta_j = \sum_{i \in \varphi^{-1}\{j\}} \alpha_i = 0$. But recall that ($\beta_i > 0$ for each $i \in [q]$). Applying this to $i = j$, we obtain $\beta_j > 0$. This contradicts $\beta_j = 0$.

This contradiction shows that our assumption (that $\varphi^{-1}\{j\} = \varnothing$) was wrong. Hence, we have $\varphi^{-1}\{j\} \neq \varnothing$. Hence, there exists some $k \in \varphi^{-1}\{j\}$. Consider this k.

From $k \in \varphi^{-1}\{j\}$, we obtain $\varphi(k) \in \{j\}$, so that $\varphi(k) = j$. Hence, $j = \varphi\left(\underbrace{k}_{\in \varphi^{-1}\{j\} \subset [p]}\right) \in \varphi([p])$.

Now, forget that we fixed j. We thus have shown that $j \in \varphi([p])$ for each $j \in [q]$. In other words, $[q] \subset \varphi([p])$. In other words, the map φ is surjective. This proves Lemma 13.46.31. □

Lemma 13.46.32. *Let $(\beta_1, \beta_2, \beta_3, \ldots) \in \mathbb{N}^\infty$ be such that $\beta_1 \geq \beta_2 \geq \beta_3 \geq \cdots$. Let $k \in \mathbb{N}$. Let R be a finite subset of $\{1, 2, 3, \ldots\}$ such that $|R| \leq k$. Then,*

$$\sum_{r \in R} \beta_r \leq \beta_1 + \beta_2 + \cdots + \beta_k.$$

Proof of Lemma 13.46.32. We have $\beta_1 \geq \beta_2 \geq \beta_3 \geq \cdots$. In other words, if u and v are two elements of $\{1, 2, 3, \ldots\}$ satisfying $u \leq v$, then

$$\beta_u \geq \beta_v. \tag{13.46.83}$$

Let $(r_1, r_2, \ldots, r_p)$ be a list of all elements of R in increasing order (with no repetitions). Thus, $\sum_{r \in R} \beta_r = \beta_{r_1} + \beta_{r_2} + \cdots + \beta_{r_p}$ and $R = \{r_1, r_2, \ldots, r_p\}$ and $|R| = p$ and $r_1 < r_2 < \cdots < r_p$.

Hence, $p = |R| \leq k$, so that $0 \leq p \leq k$. Furthermore, $\{r_1, r_2, \ldots, r_p\} = R \subset \{1, 2, 3, \ldots\}$.

But each $j \in \{1, 2, \ldots, p-1\}$ satisfies

$$r_{j+1} - r_j \geq 1 \tag{13.46.84}$$

[546]. Thus, each $i \in \{1, 2, \ldots, p\}$ satisfies

$$r_i \geq i \tag{13.46.85}$$

[547]. Hence, each $i \in \{1, 2, \ldots, p\}$ satisfies

$$\beta_{r_i} \leq \beta_i \tag{13.46.86}$$

[548].

Now,

$$\sum_{r \in R} \beta_r = \beta_{r_1} + \beta_{r_2} + \cdots + \beta_{r_p} = \sum_{i=1}^{p} \underbrace{\beta_{r_i}}_{\substack{\leq \beta_i \\ (\text{by } (13.46.86))}} \leq \sum_{i=1}^{p} \beta_i. \tag{13.46.87}$$

[546] *Proof of (13.46.84):* Let $j \in \{1, 2, \ldots, p-1\}$. Then, $r_j < r_{j+1}$ (since $r_1 < r_2 < \cdots < r_p$). Hence, $r_j \leq r_{j+1} - 1$ (since r_j and r_{j+1} are integers). In other words, $r_{j+1} - r_j \geq 1$. This proves (13.46.84).

[547] *Proof of (13.46.85):* Let $i \in \{1, 2, \ldots, p\}$. Then, $\sum_{j=1}^{i-1} (r_{j+1} - r_j) = r_i - r_1$ (by the telescope principle). Hence,

$$r_i - r_1 = \sum_{j=1}^{i-1} \underbrace{(r_{j+1} - r_j)}_{\substack{\geq 1 \\ (\text{by } (13.46.84))}} \geq \sum_{j=1}^{i-1} 1 = i - 1.$$

Hence, $r_i \geq (i-1) + r_1$. But $r_1 \geq 1$ (since $r_1 \in \{r_1, r_2, \ldots, r_p\} \subset \{1, 2, 3, \ldots\}$). Thus, $r_i \geq (i-1) + \underbrace{r_1}_{\geq 1} \geq (i-1) + 1 = i$. This proves (13.46.85).

[548] *Proof of (13.46.86):* Let $i \in \{1, 2, \ldots, p\}$. From (13.46.85), we obtain $r_i \geq i$. Thus, $i \leq r_i$. Hence, (13.46.83) (applied to r_i and i instead of u and v) shows that $\beta_i \geq \beta_{r_i}$. In other words, $\beta_{r_i} \leq \beta_i$. This proves (13.46.86).

But

$$\beta_1 + \beta_2 + \cdots + \beta_k = \sum_{i=1}^{k} \beta_i = \sum_{i=1}^{p} \beta_i + \sum_{i=p+1}^{k} \underbrace{\beta_i}_{\substack{\geq 0 \\ (\text{since } \beta_i \in \mathbb{N})}}$$

(here, we have split the sum into two at $i = p$ (since $0 \leq p \leq k$))

$$\geq \sum_{i=1}^{p} \beta_i + \underbrace{\sum_{i=p+1}^{k} 0}_{=0} = \sum_{i=1}^{p} \beta_i \geq \sum_{r \in R} \beta_r \qquad (\text{by } (13.46.87)).$$

In other words, $\sum_{r \in R} \beta_r \leq \beta_1 + \beta_2 + \cdots + \beta_k$. This proves Lemma 13.46.32. □

We now introduce some notations:

Definition 13.46.33. Let $p \in \mathbb{N}$ and $q \in \mathbb{N}$. Let $\alpha \in \mathbb{N}^p$ and $\beta \in \mathbb{N}^q$. Then, we define a set $\mathfrak{B}_{\alpha,\beta,p,q}$ by

$$\mathfrak{B}_{\alpha,\beta,p,q} = \{\varphi : [p] \to [q] \mid \beta = \varphi_* \alpha\}.$$

Note that this set $\mathfrak{B}_{\alpha,\beta,p,q}$ is a subset of $\{\varphi : [p] \to [q]\} = [q]^{[p]}$, and therefore is a finite set (since $[q]^{[p]}$ is a finite set).

Example 13.46.34. Let $p = 3$, $q = 4$, $\alpha = (2, 1, 2)$ and $\beta = (3, 0, 2, 0)$. Then, $\mathfrak{B}_{\alpha,\beta,p,q} = \{\varphi_1, \varphi_2\}$, where φ_1 and φ_2 are the two maps $[3] \to [4]$ defined by

$$\begin{aligned} &\varphi_1(1) = 1, \qquad &&\varphi_1(2) = 1, \qquad &&\varphi_1(3) = 3; \\ &\varphi_2(1) = 3, &&\varphi_2(2) = 1, &&\varphi_2(3) = 1. \end{aligned}$$

Notice that $\mathfrak{B}_{\alpha,\beta,p,q}$ depends nontrivially on p and q. Indeed, recall that we are identifying any k-tuple $(a_1, a_2, \ldots, a_k) \in \mathbb{N}^k$ with the weak composition $(a_1, a_2, \ldots, a_k, 0, 0, 0, \ldots)$; therefore, any two tuples of nonnegative integers that differ only in trailing zeroes are equated with each other (for example, $(2, 3)$ is equated with $(2, 3, 0)$), because they are being identified with one and the same weak composition. Thus, for example, the 3-tuple $\alpha = (2, 1, 2)$ is equated with the 4-tuple $\alpha' = (2, 1, 2, 0)$. But the set $\mathfrak{B}_{\alpha,\beta,p,q} = \mathfrak{B}_{\alpha,\beta,3,4}$ cannot be equated with $\mathfrak{B}_{\alpha',\beta,4,4}$; indeed, the set $\mathfrak{B}_{\alpha',\beta,4,4}$ has more than two elements (due to the extra choice in picking the image of 4 under the map $\varphi \in \mathfrak{B}_{\alpha',\beta,4,4}$), and so we have $|\mathfrak{B}_{\alpha,\beta,3,4}| \neq |\mathfrak{B}_{\alpha',\beta,4,4}|$.

Remark 13.46.35. We recall that every k-tuple $(a_1, a_2, \ldots, a_k) \in \mathbb{N}^k$ is identified with the weak composition $(a_1, a_2, \ldots, a_k, 0, 0, 0, \ldots)$. Thus, conversely, any weak composition $\alpha = (\alpha_1, \alpha_2, \alpha_3, \ldots)$ is identified with any k-tuple that is obtained from it by removing trailing zeroes. For example, the weak composition $(2, 0, 3, 0, 0, 0, \ldots)$ is identified with the 3-tuple $(2, 0, 3)$, with the 4-tuple $(2, 0, 3, 0)$, and so on.

These identifications have an important consequence: If λ and μ are two weak compositions, then there exist nonnegative integers p and q for which λ can be identified with a p-tuple (namely, with $(\lambda_1, \lambda_2, \ldots, \lambda_p)$) and μ can be identified with a q-tuple (namely, with $(\mu_1, \mu_2, \ldots, \mu_q)$). Any two such integers p and q give rise to a well-defined set $\mathfrak{B}_{\lambda,\mu,p,q}$, defined by regarding λ as a p-tuple and regarding μ as a q-tuple.

Definition 13.46.36. Let $\lambda \in \operatorname{Par}$ and $\mu \in \mathbb{N}^\infty$. Let $\ell = \ell(\lambda)$. Then, we define a set $\mathfrak{B}'_{\lambda,\mu}$ by

$$\mathfrak{B}'_{\lambda,\mu} = \left\{ \varphi : [\ell] \to \{1, 2, 3, \ldots\} \mid \mu_j = \sum_{\substack{i \in [\ell]; \\ \varphi(i) = j}} \lambda_i \text{ for all } j \in \{1, 2, 3, \ldots\} \right\}.$$

Definition 13.46.37. Let X, Y and Z be three sets such that $X \subset Y$. Let $g : Z \to X$ be any map. Then, we define a map $g\mid^Y : Z \to Y$ by

$$\left(\left(g \mid^Y \right)(z) = g(z) \qquad \text{for each } z \in Z \right).$$

(This is well-defined, since each $z \in Z$ satisfies $g(z) \in X \subset Y$.)

The map $g\mid^Y$ is identical to the map g except for the fact that its target is Y rather than X.

If X, Y and Z are three sets such that $X \subset Y$, then the maps of the form $g \mid^{Y}$ (with g being a map $Z \to X$) are exactly those maps $Z \to Y$ whose image is contained in X. More precisely, the following lemma holds:

Lemma 13.46.38. *Let X, Y and Z be three sets such that $X \subset Y$. Then, the map*

$$X^Z \to \left\{ f \in Y^Z \mid f(Z) \subset X \right\}, \\ g \mapsto g \mid^{Y}$$

is a bijection.

Proof of Lemma 13.46.38. Lemma 13.46.38 is a fundamental fact about sets. We thus omit its proof. □

Proposition 13.46.39. *Let $\lambda \in \operatorname{Par}$ and $\mu \in \operatorname{WC}$. Let $p = \ell(\lambda)$. From $p = \ell(\lambda)$, we obtain $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_p)$; thus, λ is a p-tuple in $\mathbb{N}^p$. Let $q \in \mathbb{N}$ be such that $\mu = (\mu_1, \mu_2, \ldots, \mu_q)$ (that is, $\mu_i = 0$ for all integers $i > q$). Thus, μ is a q-tuple in $\mathbb{N}^q$. Hence, the set $\mathfrak{B}_{\lambda,\mu,p,q}$ is well-defined (since $\lambda \in \mathbb{N}^p$ and $\mu \in \mathbb{N}^q$).*

Now, $\mathfrak{B}'_{\lambda,\mu} \cong \mathfrak{B}_{\lambda,\mu,p,q}$ as sets.

Proof of Proposition 13.46.39. We have $\mu = (\mu_1, \mu_2, \ldots, \mu_q)$. In other words,

$$\mu_j = 0 \qquad \text{for each } j \in \{q+1, q+2, q+3, \ldots\}. \tag{13.46.88}$$

We have $p = \ell(\lambda)$. Hence, $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p > 0$ (by the definition of $\ell(\lambda)$). Therefore,

$$\lambda_k > 0 \qquad \text{for each } k \in [p]. \tag{13.46.89}$$

The definition of $\mathfrak{B}_{\lambda,\mu,p,q}$ shows that

$$\mathfrak{B}_{\lambda,\mu,p,q} = \{\varphi : [p] \to [q] \mid \mu = \varphi_* \lambda\}$$

(where, of course, λ and μ are regarded as the tuples $(\lambda_1, \lambda_2, \ldots, \lambda_p) \in \mathbb{N}^p$ and $(\mu_1, \mu_2, \ldots, \mu_q) \in \mathbb{N}^q$). Thus,

$$\mathfrak{B}_{\lambda,\mu,p,q} = \{\varphi : [p] \to [q] \mid \mu = \varphi_* \lambda\} = \{\zeta : [p] \to [q] \mid \mu = \zeta_* \lambda\} \tag{13.46.90}$$

(here, we have renamed the index φ as ζ).

On the other hand, $p = \ell(\lambda)$. Hence, the definition of $\mathfrak{B}'_{\lambda,\mu}$ yields

$$\begin{aligned} \mathfrak{B}'_{\lambda,\mu} &= \left\{ \varphi : [p] \to \{1,2,3,\ldots\} \mid \mu_j = \sum_{\substack{i \in [p]; \\ \varphi(i) = j}} \lambda_i \text{ for all } j \in \{1,2,3,\ldots\} \right\} \\ &= \left\{ \zeta : [p] \to \{1,2,3,\ldots\} \mid \mu_j = \sum_{\substack{i \in [p]; \\ \zeta(i) = j}} \lambda_i \text{ for all } j \in \{1,2,3,\ldots\} \right\} \end{aligned} \tag{13.46.91}$$

(here, we have renamed the index φ as ζ).

But $[q] = \{1, 2, \ldots, q\} \subset \{1, 2, 3, \ldots\}$. Thus, Lemma 13.46.38 (applied to $X = [q]$, $Y = \{1, 2, 3, \ldots\}$ and $Z = [p]$) shows that the map

$$[q]^{[p]} \to \left\{ f \in \{1,2,3,\ldots\}^{[p]} \mid f([p]) \subset [q] \right\}, \\ g \mapsto g \mid^{\{1,2,3,\ldots\}}$$

is a bijection. Denote this bijection by Φ.

We have $\mathfrak{B}_{\lambda,\mu,p,q} = \{\varphi : [p] \to [q] \mid \mu = \varphi_* \lambda\} \subset \{\varphi : [p] \to [q]\} = [q]^{[p]}$. In other words, $\mathfrak{B}_{\lambda,\mu,p,q}$ is a subset of $[q]^{[p]}$. Thus, the image $\Phi(\mathfrak{B}_{\lambda,\mu,p,q})$ is well-defined.

The map Φ is bijective (since it is a bijection), and therefore injective.

If U and V are two finite sets, and if T is a subset of U, and if $\Psi : U \to V$ is an injective map, then $\Psi(T) \cong T$ as sets.[549] Applying this to $U = [q]^{[p]}$, $V = \left\{ f \in \{1,2,3,\ldots\}^{[p]} \mid f([p]) \subset [q] \right\}$, $T = \mathfrak{B}_{\lambda,\mu,p,q}$ and $\Psi = \Phi$, we conclude that

$$\Phi(\mathfrak{B}_{\lambda,\mu,p,q}) \cong \mathfrak{B}_{\lambda,\mu,p,q} \qquad \text{as sets} \tag{13.46.92}$$

[549] This is a basic fact about sets.

(since Φ is injective).

On the other hand, let $\varphi : [p] \to [q]$ be any map. Then,

$$\mu_j = \sum_{\substack{i\in[p];\\ \varphi(i)=j}} \lambda_i \text{ for all } j \in \{q+1, q+2, q+3, \ldots\} \tag{13.46.93}$$

[550].

Recall that $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_p)$. Hence, the definition of $\varphi_* \lambda$ shows that

$$\varphi_* \lambda = \left(\sum_{i\in\varphi^{-1}\{1\}} \lambda_i, \sum_{i\in\varphi^{-1}\{2\}} \lambda_i, \ldots, \sum_{i\in\varphi^{-1}\{q\}} \lambda_i \right).$$

Thus, we have the following chain of logical equivalences:

$$\begin{aligned} &\left(\underbrace{\mu}_{=(\mu_1,\mu_2,\ldots,\mu_q)} = \underbrace{\varphi_*\lambda}_{=\left(\sum_{i\in\varphi^{-1}\{1\}}\lambda_i, \sum_{i\in\varphi^{-1}\{2\}}\lambda_i, \ldots, \sum_{i\in\varphi^{-1}\{q\}}\lambda_i\right)} \right) \\ \Longleftrightarrow\ & \left((\mu_1, \mu_2, \ldots, \mu_q) = \left(\sum_{i\in\varphi^{-1}\{1\}} \lambda_i, \sum_{i\in\varphi^{-1}\{2\}} \lambda_i, \ldots, \sum_{i\in\varphi^{-1}\{q\}} \lambda_i \right) \right) \\ \Longleftrightarrow\ & \left(\mu_j = \underbrace{\sum_{i\in\varphi^{-1}\{j\}}}_{=\sum_{\substack{i\in[p];\\ \varphi(i)=j}}} \lambda_i \text{ for all } j \in \{1, 2, \ldots, q\} \right) \\ \Longleftrightarrow\ & \left(\mu_j = \sum_{\substack{i\in[p];\\ \varphi(i)=j}} \lambda_i \text{ for all } j \in \{1, 2, \ldots, q\} \right). \end{aligned} \tag{13.46.94}$$

Now, we have the following logical implication:

$$(\mu = \varphi_* \lambda) \implies \left(\mu_j = \sum_{\substack{i\in[p];\\ \varphi(i)=j}} \lambda_i \text{ for all } j \in \{1, 2, 3, \ldots\} \right) \tag{13.46.95}$$

[551].

Now, forget that we fixed φ. We thus have proven the implication (13.46.95) for each map $\varphi : [p] \to [q]$.

[550] *Proof of (13.46.93):* Let $j \in \{q+1, q+2, q+3, \ldots\}$. Thus, $j \geq q+1$ and $\mu_j = 0$ (by (13.46.88)).

Let $i \in [p]$ be such that $\varphi(i) = j$. Thus, $j = \varphi(i) \in [q] = \{1, 2, \ldots, q\}$, so that $j \leq q$. This contradicts $j \geq q+1 > q$.

Now, forget that we fixed i. We thus have found a contradiction for each $i \in [p]$ satisfying $\varphi(i) = j$. Hence, there exists no $i \in [p]$ satisfying $\varphi(i) = j$. Thus, the sum $\sum_{\substack{i\in[p];\\ \varphi(i)=j}} \lambda_i$ is empty. Hence, $\sum_{\substack{i\in[p];\\ \varphi(i)=j}} \lambda_i = (\text{empty sum}) = 0$. Comparing this with $\mu_j = 0$, we obtain $\mu_j = \sum_{\substack{i\in[p];\\ \varphi(i)=j}} \lambda_i$. This proves (13.46.93).

[551] *Proof of (13.46.95):* Assume that $(\mu = \varphi_* \lambda)$ holds. We must show that $\left(\mu_j = \sum_{\substack{i\in[p];\\ \varphi(i)=j}} \lambda_i \text{ for all } j \in \{1, 2, 3, \ldots\} \right)$ holds.

From the equivalence (13.46.94), we conclude that

$$\left(\mu_j = \sum_{\substack{i\in[p];\\ \varphi(i)=j}} \lambda_i \text{ for all } j \in \{1,2,\ldots,q\}\right) \tag{13.46.96}$$

holds (because $(\mu = \varphi_* \lambda)$ holds).

Now, let $j \in \{1,2,3,\ldots\}$. Then, we want to prove that $\mu_j = \sum_{\substack{i\in[p];\\ \varphi(i)=j}} \lambda_i$. If $j \in \{q+1, q+2, q+3, \ldots\}$, then this follows immediately from (13.46.93). Hence, for the rest of this proof, we WLOG assume that $j \notin \{q+1, q+2, q+3, \ldots\}$. Combining this with $j \in \{1,2,3,\ldots\}$, we obtain $j \in \{1,2,3,\ldots\} \setminus \{q+1, q+2, q+3, \ldots\} = \{1,2,\ldots,q\}$. Hence, from (13.46.96), we obtain $\mu_j = \sum_{\substack{i\in[p];\\ \varphi(i)=j}} \lambda_i$.

Now, forget that we fixed j. We thus have shown that $\left(\mu_j = \sum_{\substack{i\in[p];\\ \varphi(i)=j}} \lambda_i \text{ for all } j \in \{1,2,3,\ldots\}\right)$ holds. This concludes the proof of the implication (13.46.95).

Now, it is easy to see that $\Phi\left(\mathfrak{B}_{\lambda,\mu,p,q}\right) \subset \mathfrak{B}'_{\lambda,\mu}$ [552] and $\mathfrak{B}'_{\lambda,\mu} \subset \Phi\left(\mathfrak{B}_{\lambda,\mu,p,q}\right)$ [553]. Combining these two inclusions, we obtain $\Phi\left(\mathfrak{B}_{\lambda,\mu,p,q}\right) = \mathfrak{B}'_{\lambda,\mu}$. Hence,

$$\mathfrak{B}'_{\lambda,\mu} = \Phi\left(\mathfrak{B}_{\lambda,\mu,p,q}\right) \cong \mathfrak{B}_{\lambda,\mu,p,q} \qquad \text{as sets}$$

(by (13.46.92)). This proves Proposition 13.46.39. □

[552] *Proof.* Let $\psi \in \Phi\left(\mathfrak{B}_{\lambda,\mu,p,q}\right)$. Thus, $\psi \in \Phi\left(\mathfrak{B}_{\lambda,\mu,p,q}\right) \subset \left\{f \in \{1,2,3,\ldots\}^{[p]} \mid f([p]) \subset [q]\right\}$. In other words, ψ is an element of $\{1,2,3,\ldots\}^{[p]}$ and satisfies $\psi([p]) \subset [q]$.

But $\psi \in \Phi\left(\mathfrak{B}_{\lambda,\mu,p,q}\right)$. In other words, there exists some $\varphi \in \mathfrak{B}_{\lambda,\mu,p,q}$ such that $\psi = \Phi(\varphi)$. Consider this φ. We have $\psi = \Phi(\varphi) = \varphi \mid^{\{1,2,3,\ldots\}}$ (by the definition of Φ).

We have $\varphi \in \mathfrak{B}_{\lambda,\mu,p,q} = \{\zeta : [p] \to [q] \mid \mu = \zeta_* \lambda\}$ (by (13.46.90)). In other words, φ is a map $[p] \to [q]$ and satisfies $\mu = \varphi_* \lambda$.

Each $i \in [p]$ satisfies

$$\underbrace{\psi}_{=\varphi\mid^{\{1,2,3,\ldots\}}}(i) = \left(\varphi \mid^{\{1,2,3,\ldots\}}\right)(i) = \varphi(i) \tag{13.46.97}$$

(by the definition of $\varphi \mid^{\{1,2,3,\ldots\}}$).

We have $(\mu = \varphi_* \lambda)$. Hence, from the implication (13.46.95), we conclude that

$$\left(\mu_j = \sum_{\substack{i\in[p];\\ \varphi(i)=j}} \lambda_i \text{ for all } j \in \{1,2,3,\ldots\}\right). \tag{13.46.98}$$

Now, each $j \in \{1,2,3,\ldots\}$ satisfies

$$\mu_j = \underbrace{\sum_{\substack{i\in[p];\\ \varphi(i)=j}}}_{\substack{=\sum_{\substack{i\in[p];\\ \psi(i)=j}} \\ \text{(because every } i\in[p] \text{ satisfies } \varphi(i)=\psi(i) \\ \text{(by (13.46.97)))}}} \lambda_i \qquad \text{(by (13.46.98))}$$
$$= \sum_{\substack{i\in[p];\\ \psi(i)=j}} \lambda_i.$$

In other words, we have $\left(\mu_j = \sum_{\substack{i\in[p];\\ \psi(i)=j}} \lambda_i \text{ for all } j \in \{1,2,3,\ldots\}\right)$.

Now, ψ is a map $[p] \to \{1,2,3,\ldots\}$ (since ψ is an element of $\{1,2,3,\ldots\}^{[p]}$) and satisfies $\left(\mu_j = \sum_{\substack{i\in[p];\\ \psi(i)=j}} \lambda_i \text{ for all } j \in \{1,2,3,\ldots\}\right)$. In other words,

$$\psi \in \left\{\zeta : [p] \to \{1,2,3,\ldots\} \mid \mu_j = \sum_{\substack{i\in[p];\\ \zeta(i)=j}} \lambda_i \text{ for all } j \in \{1,2,3,\ldots\}\right\}.$$

In light of (13.46.91), this rewrites as $\psi \in \mathfrak{B}'_{\lambda,\mu}$.

Now, forget that we fixed ψ. We thus have proven that $\psi \in \mathfrak{B}'_{\lambda,\mu}$ for each $\psi \in \Phi\left(\mathfrak{B}_{\lambda,\mu,p,q}\right)$. In other words, $\Phi\left(\mathfrak{B}_{\lambda,\mu,p,q}\right) \subset \mathfrak{B}'_{\lambda,\mu}$. Qed.

[553] *Proof.* Let $\psi \in \mathfrak{B}'_{\lambda,\mu}$. Thus, $\psi \in \mathfrak{B}'_{\lambda,\mu} = \left\{\zeta : [p] \to \{1,2,3,\ldots\} \mid \mu_j = \sum_{\substack{i\in[p];\\ \zeta(i)=j}} \lambda_i \text{ for all } j \in \{1,2,3,\ldots\}\right\}$ (by (13.46.91)). In other words, ψ is a map $[p] \to \{1,2,3,\ldots\}$ and satisfies

$$\left(\mu_j = \sum_{\substack{i\in[p];\\ \psi(i)=j}} \lambda_i \text{ for all } j \in \{1,2,3,\ldots\}\right). \tag{13.46.99}$$

Now, let $k \in [p]$. We shall show that $\psi(k) \in [q]$.

Proposition 13.46.40. *Let $n \in \mathbb{N}$. Let $\lambda \in \operatorname{Par}_n$ and $\mu \in \operatorname{Par}_n$. Assume that $\mathfrak{B}'_{\lambda,\mu} \neq \varnothing$. Then, $\mu \rhd \lambda$.*

Proof of Proposition 13.46.40. Let $\ell = \ell(\lambda)$. We have $\mathfrak{B}'_{\lambda,\mu} \neq \varnothing$. In other words, there exists some $\psi \in \mathfrak{B}'_{\lambda,\mu}$. Consider this ψ. We have

$$\psi = \mathfrak{B}'_{\lambda,\mu} = \left\{ \varphi : [\ell] \to \{1,2,3,\ldots\} \ \mid \ \mu_j = \sum_{\substack{i \in [\ell];\\ \varphi(i)=j}} \lambda_i \text{ for all } j \in \{1,2,3,\ldots\} \right\}$$

Indeed, assume the contrary. Thus, $\psi(k) \notin [q]$. Set $j = \psi(k)$. Combining $j = \psi(k) \in \{1,2,3,\ldots\}$ with $j = \psi(k) \notin [q] = \{1,2,\ldots,q\}$, we obtain $j \in \{1,2,3,\ldots\} \setminus \{1,2,\ldots,q\} = \{q+1, q+2, q+3, \ldots\}$. Hence, $j \geq q+1$ and $\mu_j = 0$ (by (13.46.88)).

But k is an element of $[p]$ and satisfies $\psi(k) = j$ (since $j = \psi(k)$). Hence, the sum $\sum_{\substack{i \in [p];\\ \psi(i)=j}} \lambda_i$ has an addend for $i = k$. If we split off this addend from this sum, then we obtain

$$\sum_{\substack{i \in [p];\\ \psi(i)=j}} \lambda_i = \lambda_k + \sum_{\substack{i \in [p];\\ \psi(i)=j;\\ i \neq k}} \underbrace{\lambda_i}_{\substack{\geq 0\\ (\text{since } \lambda_i \in \mathbb{N})}} \geq \lambda_k + \underbrace{\sum_{\substack{i \in [p];\\ \psi(i)=j;\\ i \neq k}} 0}_{=0} = \lambda_k > 0 \qquad (\text{by (13.46.89)}).$$

Hence, (13.46.99) yields $\mu_j = \sum_{\substack{i \in [p];\\ \psi(i)=j}} \lambda_i > 0$. This contradicts $\mu_j = 0$.

This contradiction completes our proof of $\psi(k) \in [q]$.

Now, forget that we fixed k. We thus have proven that $\psi(k) \in [q]$ for each $k \in [p]$. In other words, $\psi([p]) \subset [q]$.

Altogether, we know that ψ is an element of $\{1,2,3,\ldots\}^{[p]}$ (since ψ is a map $[p] \to \{1,2,3,\ldots\}$) and satisfies $\psi([p]) \subset [q]$. In other words, $\psi \in \left\{ f \in \{1,2,3,\ldots\}^{[p]} \ \mid \ f([p]) \subset [q] \right\}$. Hence, an element $\Phi^{-1}(\psi) \in [q]^{[p]}$ is well-defined (since Φ is a bijection $[q]^{[p]} \to \left\{ f \in \{1,2,3,\ldots\}^{[p]} \ \mid \ f([p]) \subset [q] \right\}$). Let us denote this element $\Phi^{-1}(\psi)$ by φ. Thus, $\varphi = \Phi^{-1}(\psi) \in [q]^{[p]}$. In other words, φ is a map $[p] \to [q]$.

From $\varphi = \Phi^{-1}(\psi)$, we obtain $\psi = \Phi(\varphi) = \varphi\mid^{\{1,2,3,\ldots\}}$ (by the definition of Φ).

Each $i \in [p]$ satisfies

$$\underbrace{\psi}_{=\varphi\mid^{\{1,2,3,\ldots\}}} (i) = \left(\varphi \mid^{\{1,2,3,\ldots\}} \right)(i) = \varphi(i) \tag{13.46.100}$$

(by the definition of $\varphi\mid^{\{1,2,3,\ldots\}}$).

Now, each $j \in \{1,2,\ldots,q\}$ satisfies

$$\begin{aligned} \mu_j &= \underbrace{\sum_{\substack{i \in [p];\\ \psi(i)=j}}}_{\substack{=\sum_{\substack{i\in[p];\\ \varphi(i)=j}}\\ (\text{because every } i \in [p] \text{ satisfies } \psi(i)=\varphi(i)\\ (\text{by } (13.46.100)))}} \lambda_i \qquad (\text{by (13.46.99)}) \\ &= \sum_{\substack{i \in [p];\\ \varphi(i)=j}} \lambda_i . \end{aligned}$$

In other words, we have $\left(\mu_j = \sum_{\substack{i \in [p];\\ \varphi(i)=j}} \lambda_i \text{ for all } j \in \{1,2,\ldots,q\} \right)$. Hence, the equivalence (13.46.94) shows that $(\mu = \varphi_* \lambda)$.

Altogether, we now know that φ is a map $[p] \to [q]$ and satisfies $\mu = \varphi_* \lambda$. In other words, $\varphi \in \{\zeta : [p] \to [q] \ \mid \ \mu = \zeta_* \lambda\}$. In light of (13.46.90), this rewrites as $\varphi \in \mathfrak{B}_{\lambda,\mu,p,q}$. Hence, $\psi = \Phi\left(\underbrace{\varphi}_{\in \mathfrak{B}_{\lambda,\mu,p,q}} \right) \in \Phi\left(\mathfrak{B}_{\lambda,\mu,p,q}\right)$.

Now, forget that we fixed ψ. We thus have proven that $\psi \in \Phi\left(\mathfrak{B}_{\lambda,\mu,p,q}\right)$ for each $\psi \in \mathfrak{B}'_{\lambda,\mu}$. In other words, $\mathfrak{B}'_{\lambda,\mu} \subset \Phi\left(\mathfrak{B}_{\lambda,\mu,p,q}\right)$. Qed.

(by the definition of $\mathfrak{B}'_{\lambda,\mu}$ (since $\ell = \ell(\lambda)$)). In other words, ψ is a map $[\ell] \to \{1, 2, 3, \ldots\}$ and satisfies

$$\left(\mu_j = \sum_{\substack{i\in[\ell];\\ \psi(i)=j}} \lambda_i \text{ for all } j \in \{1, 2, 3, \ldots\}\right). \tag{13.46.101}$$

Let $k \in \{1, 2, \ldots, \ell\}$. Then, $1 \leq k \leq \ell$. Hence, $[k]$ is a subset of $[\ell]$.

Define a subset R of $\{1, 2, 3, \ldots\}$ by $R = \psi([k])$. This set $R = \psi([k])$ is finite (since $[k]$ is finite).

Each $i \in \{1, 2, \ldots, k\}$ satisfies

$$\sum_{\substack{j\in R;\\ \psi(i)=j}} \lambda_i = \lambda_i \tag{13.46.102}$$

[554].

It is well-known that if X and Y are two sets, if $f : X \to Y$ is any map, and if T is a finite subset of X, then $|f(T)| \leq |T|$. Applying this to $X = [\ell]$, $Y = \{1, 2, 3, \ldots\}$, $f = \psi$ and $T = [k]$, we conclude that $|\psi([k])| \leq |[k]| = k$. In light of $R = \psi([k])$, this rewrites as $|R| \leq k$.

We know that μ is a partition; thus, $\mu_1 \geq \mu_2 \geq \mu_3 \geq \cdots$. Moreover, $\mu \in \operatorname{Par} \subset \mathbb{N}^\infty$. Hence, $(\mu_1, \mu_2, \mu_3, \ldots) = \mu \in \mathbb{N}^\infty$. Thus, Lemma 13.46.32 (applied to $\beta_i = \mu_i$) yields $\sum_{r\in R} \mu_r \leq \mu_1 + \mu_2 + \cdots + \mu_k$. Hence,

$$\begin{aligned}
\mu_1 + \mu_2 + \cdots + \mu_k &\geq \sum_{r\in R} \mu_r \\
&= \sum_{j\in R} \underbrace{\mu_j}_{\substack{=\sum_{\substack{i\in[\ell];\\ \psi(i)=j}} \lambda_i \\ \text{(by (13.46.101))}}} \qquad \text{(here, we have renamed the summation index } r \text{ as } j\text{)} \\
&= \underbrace{\sum_{j\in R} \sum_{\substack{i\in[\ell];\\ \psi(i)=j}}}_{=\sum_{i\in[\ell]} \sum_{\substack{j\in R;\\ \psi(i)=j}}} \lambda_i = \underbrace{\sum_{i\in[\ell]}}_{=\sum_{i=1}^{\ell}} \sum_{\substack{j\in R;\\ \psi(i)=j}} \lambda_i \\
&= \sum_{i=1}^{\ell} \sum_{\substack{j\in R;\\ \psi(i)=j}} \lambda_i = \sum_{i=1}^{k} \underbrace{\sum_{\substack{j\in R;\\ \psi(i)=j}} \lambda_i}_{\substack{=\lambda_i \\ \text{(by (13.46.102))}}} + \sum_{i=k+1}^{\ell} \sum_{\substack{j\in R;\\ \psi(i)=j}} \underbrace{\lambda_i}_{\substack{\geq 0 \\ \text{(since } \lambda_i\in\mathbb{N})}} \\
&\qquad \text{(here, we have split the outer sum at } i = k\text{, because } 1 \leq k \leq \ell\text{)} \\
&\geq \sum_{i=1}^{k} \lambda_i + \underbrace{\sum_{i=k+1}^{\ell} \sum_{\substack{j\in R;\\ \psi(i)=j}} 0}_{=0} = \sum_{i=1}^{k} \lambda_i = \lambda_1 + \lambda_2 + \cdots + \lambda_k.
\end{aligned}$$

[554] *Proof of (13.46.102):* Let $i \in \{1, 2, \ldots, k\}$. Thus, $i \in \{1, 2, \ldots, k\} = [k]$. Hence, $\psi\left(\underbrace{i}_{\in[k]}\right) \in \psi([k]) = R$. Thus, there exists a $j \in R$ satisfying $j = \psi(i)$ (namely, $j = \psi(i)$). This j is furthermore unique (since the condition $j = \psi(i)$ determines j uniquely). Thus, there exists **exactly one** $j \in R$ satisfying $\psi(i) = j$ (namely, $j = \psi(i)$). Hence, the sum $\sum_{\substack{j\in R;\\ \psi(i)=j}} \lambda_i$ has exactly one addend (namely, the addend for $j = \psi(i)$). Thus, this sum simplifies as follows: $\sum_{\substack{j\in R;\\ \psi(i)=j}} \lambda_i = \lambda_i$. This proves (13.46.102).

Now, let us forget that we fixed k. We thus have shown that $\mu_1+\mu_2+\cdots+\mu_k \geq \lambda_1+\lambda_2+\cdots+\lambda_k$ for each $k \in \{1,2,\ldots,\ell\}$. Therefore, Lemma 13.46.20 (applied to ℓ, λ and μ instead of q, μ and λ) yields $\mu \trianglerighteq \lambda$ (since $\ell(\lambda)=\ell\leq\ell$). This proves Proposition 13.46.40. □

Proposition 13.46.41. *Let $\mu \in \operatorname{Par}$. Let $k=\ell(\mu)$. Then:*

(a) *We have $\mathfrak{B}_{\mu,\mu,k,k}=\left\{\sigma\in\mathfrak{S}_k \mid \mu_{\sigma(i)}=\mu_i \text{ for each } i\in[k]\right\}$.*

(b) *The set $\mathfrak{B}_{\mu,\mu,k,k}$ is a subgroup of $\mathfrak{S}_k$.*

Proof of Proposition 13.46.41. We know that $\mathfrak{S}_k$ is the set of all permutations of $\{1,2,\ldots,k\}$. In other words, $\mathfrak{S}_k$ is the set of all permutations of $[k]$ (since $[k]=\{1,2,\ldots,k\}$).

Recall that $\ell(\mu)=k$. Thus, $\mu=(\mu_1,\mu_2,\ldots,\mu_k)\in\mathbb{N}^k$. Also, from $\ell(\mu)=k$, we obtain $\mu_1\geq\mu_2\geq\cdots\geq\mu_k>0$ (by the definition of $\ell(\mu)$). Hence,

$$\mu_i>0 \qquad \text{for each } i\in[k]. \tag{13.46.103}$$

The definition of $\mathfrak{B}_{\mu,\mu,k,k}$ yields

$$\mathfrak{B}_{\mu,\mu,k,k}=\{\varphi:[k]\to[k] \mid \mu=\varphi_*\mu\}, \tag{13.46.104}$$

where μ is regarded as the k-tuple $(\mu_1,\mu_2,\ldots,\mu_k)\in\mathbb{N}^k$.

We have $\mathfrak{B}_{\mu,\mu,k,k}\subset\mathfrak{S}_k$ [555]. Hence, we have the two inclusions $\mathfrak{B}_{\mu,\mu,k,k}\subset\{\sigma\in\mathfrak{S}_k \mid \sigma_*\mu=\mu\}$ [556] and $\{\sigma\in\mathfrak{S}_k \mid \sigma_*\mu=\mu\}\subset\mathfrak{B}_{\mu,\mu,k,k}$ [557]. Combining these two inclusions, we obtain

$$\mathfrak{B}_{\mu,\mu,k,k}=\{\sigma\in\mathfrak{S}_k \mid \sigma_*\mu=\mu\}. \tag{13.46.105}$$

But let $\sigma\in\mathfrak{S}_k$ be arbitrary. Then, Proposition 13.46.29(b) (applied to $p=k$, $\alpha=\mu$ and $\alpha_i=\mu_i$) shows that we have $\sigma_*\mu=\mu$ if and only if $\left(\mu_{\sigma(i)}=\mu_i \text{ for all } i\in[k]\right)$. In other words, we have the following logical equivalence:

$$(\sigma_*\mu=\mu)\iff\left(\mu_{\sigma(i)}=\mu_i \text{ for all } i\in[k]\right). \tag{13.46.106}$$

Now, forget that we fixed σ. We thus have proven the equivalence (13.46.106) for each $\sigma\in\mathfrak{S}_k$.

Now, the equality (13.46.105) becomes

$$\mathfrak{B}_{\mu,\mu,k,k}=\left\{\sigma\in\mathfrak{S}_k \mid \underbrace{\sigma_*\mu=\mu}_{\substack{\iff\left(\mu_{\sigma(i)}=\mu_i \text{ for all } i\in[k]\right)\\ \text{(by (13.46.106))}}}\right\}=\left\{\sigma\in\mathfrak{S}_k \mid \mu_{\sigma(i)}=\mu_i \text{ for each } i\in[k]\right\}.$$

This proves Proposition 13.46.41(a).

(b) The neutral element of the group $\mathfrak{S}_k$ is $\operatorname{id}_{\{1,2,\ldots,k\}}=\operatorname{id}_{[k]}$ (since $\{1,2,\ldots,k\}=[k]$).

The following four observations hold:

[555] *Proof.* Let $\psi\in\mathfrak{B}_{\mu,\mu,k,k}$. Thus, $\psi\in\mathfrak{B}_{\mu,\mu,k,k}=\{\varphi:[k]\to[k] \mid \mu=\varphi_*\mu\}$. In other words, ψ is a map $[k]\to[k]$ and satisfies $\mu=\psi_*\mu$. Thus, $\psi_*\mu=\mu$.

But Lemma 13.46.31 (applied to $p=k$, $q=k$, $\varphi=\psi$, $\alpha=\mu$, $\alpha_i=\mu_i$, $\beta=\mu$ and $\beta_i=\mu_i$) shows that the map ψ is surjective (because (13.46.103) shows that ($\mu_i>0$ for each $i\in[k]$)). Hence, ψ is a surjective map $[k]\to[k]$.

Now, recall the following known fact about finite sets: If T is a finite set, then each surjective map $T\to T$ is bijective. Applying this fact to $T=[k]$, we conclude that each surjective map $[k]\to[k]$ is bijective (since $[k]$ is a finite set). Thus, the map ψ is bijective (since ψ is a surjective map $[k]\to[k]$). Thus, ψ is a bijection $[k]\to[k]$. In other words, ψ is a permutation of $[k]$. In other words, $\psi\in\mathfrak{S}_k$ (since $\mathfrak{S}_k$ is the set of all permutations of $[k]$).

Now, forget that we fixed ψ. We thus have shown that $\psi\in\mathfrak{S}_k$ for each $\psi\in\mathfrak{B}_{\mu,\mu,k,k}$. In other words, $\mathfrak{B}_{\mu,\mu,k,k}\subset\mathfrak{S}_k$. Qed.

[556] *Proof.* Let $\psi\in\mathfrak{B}_{\mu,\mu,k,k}$. Thus, $\psi\in\mathfrak{B}_{\mu,\mu,k,k}=\{\varphi:[k]\to[k] \mid \mu=\varphi_*\mu\}$. In other words, ψ is a map $[k]\to[k]$ and satisfies $\mu=\psi_*\mu$. Thus, $\psi_*\mu=\mu$. Also, $\psi\in\mathfrak{B}_{\mu,\mu,k,k}\subset\mathfrak{S}_k$. Hence, ψ is an element of $\mathfrak{S}_k$ and satisfies $\psi_*\mu=\mu$. In other words, $\psi\in\{\sigma\in\mathfrak{S}_k \mid \sigma_*\mu=\mu\}$.

Now, forget that we fixed ψ. We thus have shown that $\psi\in\{\sigma\in\mathfrak{S}_k \mid \sigma_*\mu=\mu\}$ for each $\psi\in\mathfrak{B}_{\mu,\mu,k,k}$. In other words, $\mathfrak{B}_{\mu,\mu,k,k}\subset\{\sigma\in\mathfrak{S}_k \mid \sigma_*\mu=\mu\}$. Qed.

[557] *Proof.* Let $\psi\in\{\sigma\in\mathfrak{S}_k \mid \sigma_*\mu=\mu\}$. Thus, ψ is an element of $\mathfrak{S}_k$ and satisfies $\psi_*\mu=\mu$.

We know that ψ is an element of $\mathfrak{S}_k$. In other words, ψ is a permutation of $[k]$ (since $\mathfrak{S}_k$ is the set of all permutations of $[k]$). In other words, ψ is a bijection $[k]\to[k]$.

Thus, ψ is a map $[k]\to[k]$ and satisfies $\mu=\psi_*\mu$ (since $\psi_*\mu=\mu$). In other words, $\psi\in\{\varphi:[k]\to[k] \mid \mu=\varphi_*\mu\}$. In light of (13.46.104), this rewrites as $\psi\in\mathfrak{B}_{\mu,\mu,k,k}$.

Now, forget that we fixed ψ. We thus have proven that $\psi\in\mathfrak{B}_{\mu,\mu,k,k}$ for each $\psi\in\{\sigma\in\mathfrak{S}_k \mid \sigma_*\mu=\mu\}$. In other words, $\{\sigma\in\mathfrak{S}_k \mid \sigma_*\mu=\mu\}\subset\mathfrak{B}_{\mu,\mu,k,k}$. Qed.

- The set $\mathfrak{B}_{\mu,\mu,k,k}$ is a subset of $\mathfrak{S}_k$ (since $\mathfrak{B}_{\mu,\mu,k,k} \subset \mathfrak{S}_k$).
- If γ and δ are two elements of $\mathfrak{B}_{\mu,\mu,k,k}$, then $\gamma \circ \delta \in \mathfrak{B}_{\mu,\mu,k,k}$ [558].
- We have $\operatorname{id}_{[k]} \in \mathfrak{B}_{\mu,\mu,k,k}$ [559].
- If $\gamma \in \mathfrak{B}_{\mu,\mu,k,k}$, then $\gamma^{-1} \in \mathfrak{B}_{\mu,\mu,k,k}$ [560].

Combining these four observations, we conclude that $\mathfrak{B}_{\mu,\mu,k,k}$ is a subgroup of $\mathfrak{S}_k$ (since the neutral element of the group $\mathfrak{S}_k$ is $\operatorname{id}_{[k]}$). This proves Proposition 13.46.41(b). □

Proposition 13.46.42. *Let $p \in \mathbb{N}$. Let $\varphi : [p] \to \{1, 2, 3, \ldots\}$ be any map. Let $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_p) \in \mathbb{N}^p$ be a p-tuple. Then,*

$$\left(\sum_{i \in \varphi^{-1}\{1\}} \alpha_i, \sum_{i \in \varphi^{-1}\{2\}} \alpha_i, \sum_{i \in \varphi^{-1}\{3\}} \alpha_i, \ldots \right)$$

is a weak composition.

Proof of Proposition 13.46.42. For each $j \in \{1, 2, 3, \ldots\}$, the sum $\sum_{i \in \varphi^{-1}\{j\}} \alpha_i$ is a well-defined element of $\mathbb{N}$ [561]. Hence, the sequence

$$\left(\sum_{i \in \varphi^{-1}\{1\}} \alpha_i, \sum_{i \in \varphi^{-1}\{2\}} \alpha_i, \sum_{i \in \varphi^{-1}\{3\}} \alpha_i, \ldots \right)$$

is a well-defined element of $\mathbb{N}^\infty$. Denote this sequence by β. We shall now show that this sequence β is a weak composition.

Write the sequence β in the form $\beta = (\beta_1, \beta_2, \beta_3, \ldots)$. Thus,

$$(\beta_1, \beta_2, \beta_3, \ldots) = \beta = \left(\sum_{i \in \varphi^{-1}\{1\}} \alpha_i, \sum_{i \in \varphi^{-1}\{2\}} \alpha_i, \sum_{i \in \varphi^{-1}\{3\}} \alpha_i, \ldots \right).$$

In other words,

$$\beta_j = \sum_{i \in \varphi^{-1}\{j\}} \alpha_i \qquad \text{for each } j \in \{1, 2, 3, \ldots\}. \tag{13.46.107}$$

[558] *Proof.* Let γ and δ be two elements of $\mathfrak{B}_{\mu,\mu,k,k}$. We must prove that $\gamma \circ \delta \in \mathfrak{B}_{\mu,\mu,k,k}$.

We have $\gamma \in \mathfrak{B}_{\mu,\mu,k,k} = \{\varphi : [k] \to [k] \mid \mu = \varphi_* \mu\}$ (by (13.46.104)). In other words, γ is a map $[k] \to [k]$ and satisfies $\mu = \gamma_* \mu$. The same argument (applied to δ instead of γ) shows that δ is a map $[k] \to [k]$ and satisfies $\mu = \delta_* \mu$. Now, Proposition 13.46.28 (applied to $p = k$, $q = k$, $r = k$, $\psi = \gamma$, $\varphi = \delta$ and $\alpha = \mu$) yields $(\gamma \circ \delta)_* \mu = \gamma_* \left(\underbrace{\delta_* \mu}_{= \mu} \right) = \gamma_* \mu = \mu$. In other words, $\mu = (\gamma \circ \delta)_* \mu$. Hence, $\gamma \circ \delta$ is a map $[k] \to [k]$ (since γ and δ are maps $[k] \to [k]$) and satisfies $\mu = (\gamma \circ \delta)_* \mu$. In other words, $\gamma \circ \delta \in \{\varphi : [k] \to [k] \mid \mu = \varphi_* \mu\}$. In light of (13.46.104), this rewrites as $\gamma \circ \delta \in \mathfrak{B}_{\mu,\mu,k,k}$. Qed.

[559] *Proof.* Proposition 13.46.30 (applied to $p = k$ and $\alpha = \mu$) yields $\left(\operatorname{id}_{[k]}\right)_* \mu = \mu$. Hence, $\mu = \left(\operatorname{id}_{[k]}\right)_* \mu$.

Thus, $\operatorname{id}_{[k]}$ is a map $[k] \to [k]$ and satisfies $\mu = \left(\operatorname{id}_{[k]}\right)_* \mu$. In other words, $\operatorname{id}_{[k]} \in \{\varphi : [k] \to [k] \mid \mu = \varphi_* \mu\}$. In light of (13.46.104), this rewrites as $\operatorname{id}_{[k]} \in \mathfrak{B}_{\mu,\mu,k,k}$. Qed.

[560] *Proof.* Let $\gamma \in \mathfrak{B}_{\mu,\mu,k,k}$. We must prove that $\gamma^{-1} \in \mathfrak{B}_{\mu,\mu,k,k}$.

We have $\gamma \in \mathfrak{B}_{\mu,\mu,k,k} = \{\varphi : [k] \to [k] \mid \mu = \varphi_* \mu\}$ (by (13.46.104)). In other words, γ is a map $[k] \to [k]$ and satisfies $\mu = \gamma_* \mu$. Also, $\gamma \in \mathfrak{B}_{\mu,\mu,k,k} \subset \mathfrak{S}_k$; therefore, γ has an inverse γ^{-1} (since $\mathfrak{S}_k$ is a group). Proposition 13.46.28 (applied to $p = k$, $q = k$, $r = k$, $\psi = \gamma^{-1}$, $\varphi = \gamma$ and $\alpha = \mu$) yields $\left(\gamma^{-1} \circ \gamma\right)_* \mu = \left(\gamma^{-1}\right)_* \left(\underbrace{\gamma_* \mu}_{= \mu} \right) = \left(\gamma^{-1}\right)_* \mu$. Hence, $\left(\gamma^{-1}\right)_* \mu = \left(\underbrace{\gamma^{-1} \circ \gamma}_{= \operatorname{id}_{[k]}} \right)_* \mu = \left(\operatorname{id}_{[k]}\right)_* \mu$.

But Proposition 13.46.30 (applied to $p = k$ and $\alpha = \mu$) yields $\left(\operatorname{id}_{[k]}\right)_* \mu = \mu$. Hence, $\mu = \left(\operatorname{id}_{[k]}\right)_* \mu$. Comparing this with $\left(\gamma^{-1}\right)_* \mu = \left(\operatorname{id}_{[k]}\right)_* \mu$, we obtain $\mu = \left(\gamma^{-1}\right)_* \mu$. Hence, γ^{-1} is a map $[k] \to [k]$ (since γ is a map $[k] \to [k]$) and satisfies $\mu = \left(\gamma^{-1}\right)_* \mu$. In other words, $\gamma^{-1} \in \{\varphi : [k] \to [k] \mid \mu = \varphi_* \mu\}$. In light of (13.46.104), this rewrites as $\gamma^{-1} \in \mathfrak{B}_{\mu,\mu,k,k}$. Qed.

[561] *Proof.* Let $j \in \{1, 2, 3, \ldots\}$. Then, $\varphi^{-1}\{j\}$ is a subset of $[p]$, and thus a finite set (since $[p]$ is a finite set). Hence, the sum $\sum_{i \in \varphi^{-1}\{j\}} \alpha_i$ is well-defined. Moreover, this sum belongs to $\mathbb{N}$ (since each of its addends α_i belongs to $\mathbb{N}$). Thus, the sum $\sum_{i \in \varphi^{-1}\{j\}} \alpha_i$ is a well-defined element of $\mathbb{N}$. Qed.

Let Z be the support of the sequence β. Then,

$$\begin{aligned}Z &= (\text{the support of the sequence } \beta)\\ &= (\text{the set of all positive integers } i \text{ for which } \beta_i \neq 0)\\ &\qquad (\text{by the definition of the support of a sequence})\\ &= \{i \in \{1,2,3,\ldots\} \ \mid \ \beta_i \neq 0\}.\end{aligned}$$

Every $j \in Z$ satisfies $j \in \varphi([p])$ [562]. In other words, we have $Z \subset \varphi([p])$. Hence, the set Z is finite (since the set $\varphi([p])$ is finite (since the set $[p]$ is finite)). In other words, the support of the sequence β is finite (since Z is the support of the sequence β).

Hence, we know that β is a sequence in $\mathbb{N}^\infty$ having finite support. In other words, β is a weak composition (since a weak composition is defined as a sequence in $\mathbb{N}^\infty$ having finite support). In other words,

$$\left(\sum_{i\in\varphi^{-1}\{1\}} \alpha_i, \sum_{i\in\varphi^{-1}\{2\}} \alpha_i, \sum_{i\in\varphi^{-1}\{3\}} \alpha_i, \ldots\right)$$

is a weak composition (since $\beta = \left(\sum_{i\in\varphi^{-1}\{1\}} \alpha_i, \sum_{i\in\varphi^{-1}\{2\}} \alpha_i, \sum_{i\in\varphi^{-1}\{3\}} \alpha_i, \ldots\right)$). This proves Proposition 13.46.42. □

Proposition 13.46.42 allows us to make the following definition (which is similar to Definition 13.46.26, but uses the infinite set $\{1,2,3,\ldots\}$ instead of $[q]$):

Definition 13.46.43. Let $p \in \mathbb{N}$. Let $\varphi : [p] \to \{1,2,3,\ldots\}$ be any map. Let $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_p) \in \mathbb{N}^p$ be a p-tuple. Then, we define a weak composition $\varphi_*\alpha \in \operatorname{WC}$ by

$$\varphi_*\alpha = \left(\sum_{i\in\varphi^{-1}\{1\}} \alpha_i, \sum_{i\in\varphi^{-1}\{2\}} \alpha_i, \sum_{i\in\varphi^{-1}\{3\}} \alpha_i, \ldots\right).$$

(This is indeed a weak composition, because of Proposition 13.46.42.)

Proposition 13.46.44. *Let $p \in \mathbb{N}$. Let $\varphi : [p] \to \{1,2,3,\ldots\}$ be any map. Let $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_p) \in \mathbb{N}^p$ be a p-tuple. Then, $\mathbf{x}^{\varphi_*\alpha} = x_{\varphi(1)}^{\alpha_1} x_{\varphi(2)}^{\alpha_2} \cdots x_{\varphi(p)}^{\alpha_p}$.*

Proof of Proposition 13.46.44. The definition of $\varphi_*\alpha$ yields

$$\varphi_*\alpha = \left(\sum_{i\in\varphi^{-1}\{1\}} \alpha_i, \sum_{i\in\varphi^{-1}\{2\}} \alpha_i, \sum_{i\in\varphi^{-1}\{3\}} \alpha_i, \ldots\right).$$

[562] *Proof.* Let $j \in Z$. We must show that $j \in \varphi([p])$.

We have $j \in Z = \{i \in \{1,2,3,\ldots\} \ \mid \ \beta_i \neq 0\}$. In other words, j is an element of $\{1,2,3,\ldots\}$ and satisfies $\beta_j \neq 0$.

Assume (for the sake of contradiction) that $\varphi^{-1}\{j\} = \varnothing$. Thus, $\sum_{i\in\varphi^{-1}\{j\}} \alpha_i = \sum_{i\in\varnothing} \alpha_i = (\text{empty sum}) = 0$. Hence, (13.46.107) yields $\beta_j = \sum_{i\in\varphi^{-1}\{j\}} \alpha_i = 0$. This contradicts $\beta_j \neq 0$.

This contradiction shows that our assumption (that $\varphi^{-1}\{j\} = \varnothing$) was false. Hence, we must have $\varphi^{-1}\{j\} \neq \varnothing$. In other words, there exists some $k \in \varphi^{-1}\{j\}$. Consider this k.

From $k \in \varphi^{-1}\{j\}$, we obtain $\varphi(k) \in \{j\}$. Hence, $\varphi(k) = j$. Therefore, $j = \varphi\left(\underbrace{k}_{\in\varphi^{-1}\{j\}\subset[p]}\right) \in \varphi([p])$. This completes our proof.

Hence, the definition of $\mathbf{x}^{\varphi_*\alpha}$ shows that

$$\mathbf{x}^{\varphi_*\alpha} = x_1^{\sum_{i\in\varphi^{-1}\{1\}}\alpha_i} x_2^{\sum_{i\in\varphi^{-1}\{2\}}\alpha_i} x_3^{\sum_{i\in\varphi^{-1}\{3\}}\alpha_i} \cdots = \underbrace{\prod_{j=1}^{\infty}}_{=\prod_{j\in\{1,2,3,\ldots\}}} \underbrace{x_j^{\sum_{i\in\varphi^{-1}\{j\}}\alpha_i}}_{=\prod_{i\in\varphi^{-1}\{j\}} x_j^{\alpha_i}}$$

$$= \prod_{j\in\{1,2,3,\ldots\}} \underbrace{\prod_{i\in\varphi^{-1}\{j\}}}_{=\prod_{\substack{i\in[p];\\ \varphi(i)=j}}} x_j^{\alpha_i} = \prod_{j\in\{1,2,3,\ldots\}} \prod_{\substack{i\in[p];\\ \varphi(i)=j}} x_j^{\alpha_i}.$$

Comparing this with

$$x_{\varphi(1)}^{\alpha_1} x_{\varphi(2)}^{\alpha_2} \cdots x_{\varphi(p)}^{\alpha_p} = \underbrace{\prod_{i\in[p]}}_{\substack{=\prod_{j\in\{1,2,3,\ldots\}} \prod_{\substack{i\in[p];\\ \varphi(i)=j}} \\ (\text{since } \varphi(i)\in\{1,2,3,\ldots\} \text{ for each } i\in[p])}} x_{\varphi(i)}^{\alpha_i} = \prod_{j\in\{1,2,3,\ldots\}} \prod_{\substack{i\in[p];\\ \varphi(i)=j}} \underbrace{x_{\varphi(i)}^{\alpha_i}}_{\substack{=x_j^{\alpha_i} \\ (\text{since } \varphi(i)=j)}}$$

$$= \prod_{j\in\{1,2,3,\ldots\}} \prod_{\substack{i\in[p];\\ \varphi(i)=j}} x_j^{\alpha_i},$$

we obtain $\mathbf{x}^{\varphi_*\alpha} = x_{\varphi(1)}^{\alpha_1} x_{\varphi(2)}^{\alpha_2} \cdots x_{\varphi(p)}^{\alpha_p}$. This proves Proposition 13.46.44. □

Lemma 13.46.45. *Let $\ell \in \mathbb{N}$. Let X be a set. Then, the map*

$$X^{\{1,2,\ldots,\ell\}} \to X^\ell, \qquad \varphi \mapsto (\varphi(1), \varphi(2), \ldots, \varphi(\ell))$$

is a bijection.

Proof of Lemma 13.46.45. It is well-known that the ℓ-tuples of elements of X are in bijection with the maps $\{1, 2, \ldots, \ell\} \to X$ [563]. Lemma 13.46.45 is merely a way to precisely formulate this bijection. Thus, we omit its proof. □

Proposition 13.46.46. *Let $\ell \in \mathbb{N}$. Let $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_\ell) \in \mathbb{N}^\ell$ be such that ($\alpha_i > 0$ for each $i \in [\ell]$). Then,*

$$p_{\alpha_1} p_{\alpha_2} \cdots p_{\alpha_\ell} = \sum_{\varphi:[\ell]\to\{1,2,3,\ldots\}} \mathbf{x}^{\varphi_*\alpha}$$

in $\mathbf{k}[[\mathbf{x}]]$.

Proof of Proposition 13.46.46. We assumed that ($\alpha_i > 0$ for each $i \in [\ell]$). Thus, each $i \in [\ell]$ satisfies $\alpha_i > 0$ and therefore

$$p_{\alpha_i} = x_1^{\alpha_i} + x_2^{\alpha_i} + x_3^{\alpha_i} + \cdots \qquad (\text{by the definition of } p_{\alpha_i})$$

$$= \sum_{j\in\{1,2,3,\ldots\}} x_j^{\alpha_i}. \tag{13.46.108}$$

Lemma 13.46.45 (applied to $X = \{1, 2, 3, \ldots\}$) shows that the map

$$\{1,2,3,\ldots\}^{\{1,2,\ldots,\ell\}} \to \{1,2,3,\ldots\}^\ell, \qquad \varphi \mapsto (\varphi(1), \varphi(2), \ldots, \varphi(\ell))$$

is a bijection. In view of $[\ell] = \{1, 2, \ldots, \ell\}$, this rewrites as follows: The map

$$\{1,2,3,\ldots\}^{[\ell]} \to \{1,2,3,\ldots\}^\ell, \qquad \varphi \mapsto (\varphi(1), \varphi(2), \ldots, \varphi(\ell)) \tag{13.46.109}$$

is a bijection.

[563] In fact, depending on your definition of an "ℓ-tuple", you might even consider the ℓ-tuples of elements of X to be **exactly** the maps $\{1, 2, \ldots, \ell\} \to X$.

Now,

$$\begin{aligned}
p_{\alpha_1} p_{\alpha_2} \cdots p_{\alpha_\ell} &= \prod_{i=1}^{\ell} \underbrace{p_{\alpha_i}}_{\substack{=\sum_{j\in\{1,2,3,\ldots\}} x_j^{\alpha_i} \\ \text{(by (13.46.108))}}} = \prod_{i=1}^{\ell} \sum_{j\in\{1,2,3,\ldots\}} x_j^{\alpha_i} \\
&= \sum_{(j_1,j_2,\ldots,j_\ell)\in\{1,2,3,\ldots\}^\ell} \prod_{i=1}^{\ell} x_{j_i}^{\alpha_i} \qquad \text{(by the product rule)} \\
&= \underbrace{\sum_{\varphi\in\{1,2,3,\ldots\}^{[\ell]}}}_{=\sum_{\varphi:[\ell]\to\{1,2,3,\ldots\}}} \underbrace{\prod_{i=1}^{\ell} x_{\varphi(i)}^{\alpha_i}}_{=x_{\varphi(1)}^{\alpha_1} x_{\varphi(2)}^{\alpha_2} \cdots x_{\varphi(\ell)}^{\alpha_\ell}} \\
&\qquad \left(\begin{array}{c}\text{here, we have substituted } (\varphi(1), \varphi(2), \ldots, \varphi(\ell)) \text{ for } (j_1, j_2, \ldots, j_\ell) \\ \text{in the sum, since the map (13.46.109) is a bijection}\end{array}\right) \\
&= \sum_{\varphi:[\ell]\to\{1,2,3,\ldots\}} x_{\varphi(1)}^{\alpha_1} x_{\varphi(2)}^{\alpha_2} \cdots x_{\varphi(\ell)}^{\alpha_\ell}.
\end{aligned}$$

Comparing this with

$$\sum_{\varphi:[\ell]\to\{1,2,3,\ldots\}} \underbrace{\mathbf{x}^{\varphi_*\alpha}}_{\substack{=x_{\varphi(1)}^{\alpha_1} x_{\varphi(2)}^{\alpha_2} \cdots x_{\varphi(\ell)}^{\alpha_\ell} \\ \text{(by Proposition 13.46.44 (applied to } p=\ell))}} = \sum_{\varphi:[\ell]\to\{1,2,3,\ldots\}} x_{\varphi(1)}^{\alpha_1} x_{\varphi(2)}^{\alpha_2} \cdots x_{\varphi(\ell)}^{\alpha_\ell},$$

we obtain $p_{\alpha_1} p_{\alpha_2} \cdots p_{\alpha_\ell} = \sum_{\varphi:[\ell]\to\{1,2,3,\ldots\}} \mathbf{x}^{\varphi_*\alpha}$. This proves Proposition 13.46.46. □

Finally, let us state a proposition that follows immediately from the definition of a weak composition:

Proposition 13.46.47. *Let μ be a weak composition. Then, there exists a $q \in \mathbb{N}$ such that $\mu = (\mu_1, \mu_2, \ldots, \mu_q)$.* [564]

Now, let us resume the solution of Exercise 2.2.13.

(j) Let λ be a partition. Let μ be a weak composition. Let $\ell = \ell(\lambda)$. We must prove that the number $b_{\lambda,\mu}$ is well-defined. In other words, we must prove that there are only finitely many maps $\varphi : \{1, 2, \ldots, \ell\} \to \{1, 2, 3, \ldots\}$ satisfying $\left(\mu_j = \sum_{\substack{i\in\{1,2,\ldots,\ell\};\\ \varphi(i)=j}} \lambda_i \text{ for all } j \geq 1\right)$.

We have $\ell = \ell(\lambda)$; thus, $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_\ell)$. Hence, λ is an ℓ-tuple in $\mathbb{N}^\ell$.

[564] Keep in mind that we are identifying any k-tuple $(a_1, a_2, \ldots, a_k) \in \mathbb{N}^k$ with the weak composition $(a_1, a_2, \ldots, a_k, 0, 0, 0, \ldots)$. Thus, the q-tuple $(\mu_1, \mu_2, \ldots, \mu_q)$ is identified with the weak composition $(\mu_1, \mu_2, \ldots, \mu_q, 0, 0, 0, \ldots)$.

We have $\ell = \ell(\lambda)$. Hence, the definition of $\mathfrak{B}'_{\lambda,\mu}$ yields

$$\mathfrak{B}'_{\lambda,\mu} = \left\{ \varphi : [\ell] \to \{1,2,3,\ldots\} \mid \underbrace{\mu_j = \sum_{\substack{i \in [\ell];\\ \varphi(i)=j}} \lambda_i \text{ for all } j \in \{1,2,3,\ldots\}}_{\substack{\iff \left(\mu_j = \sum_{\substack{i\in[\ell];\\ \varphi(i)=j}} \lambda_i \text{ for all } j \geq 1\right) \\ \text{(since the } j \in \{1,2,3,\ldots\} \text{ are precisely the integers } j \geq 1\text{)}}} \right\}$$

$$= \left\{ \varphi : [\ell] \to \{1,2,3,\ldots\} \mid \mu_j = \sum_{\substack{i \in [\ell];\\ \varphi(i)=j}} \lambda_i \text{ for all } j \geq 1 \right\}$$

$$= \left\{ \varphi : \{1,2,\ldots,\ell\} \to \{1,2,3,\ldots\} \mid \mu_j = \sum_{\substack{i \in \{1,2,\ldots,\ell\};\\ \varphi(i)=j}} \lambda_i \text{ for all } j \geq 1 \right\} \tag{13.46.110}$$

(since $[\ell] = \{1,2,\ldots,\ell\}$).

Proposition 13.46.47 shows that there exists some $q \in \mathbb{N}$ such that $\mu = (\mu_1, \mu_2, \ldots, \mu_q)$. Consider this q. Thus, μ is a q-tuple in $\mathbb{N}^q$ (since $\mu = (\mu_1, \mu_2, \ldots, \mu_q)$). Hence, the set $\mathfrak{B}_{\lambda,\mu,\ell,q}$ is well-defined (since $\lambda \in \mathbb{N}^\ell$ and $\mu \in \mathbb{N}^q$). Also, $\mu \in \mathrm{WC}$ (since μ is a weak composition). Thus, Proposition 13.46.39 (applied to $p = \ell$) shows that $\mathfrak{B}'_{\lambda,\mu} \cong \mathfrak{B}_{\lambda,\mu,\ell,q}$ as sets. But the definition of $\mathfrak{B}_{\lambda,\mu,\ell,q}$ yields

$$\mathfrak{B}_{\lambda,\mu,\ell,q} = \{\varphi : [\ell] \to [q] \mid \mu = \varphi_* \lambda\} \subset \{\varphi : [\ell] \to [q]\} = [q]^{[\ell]}.$$

Hence, $\mathfrak{B}_{\lambda,\mu,\ell,q}$ is a finite set (since $[q]^{[\ell]}$ is a finite set). Therefore, the set $\mathfrak{B}'_{\lambda,\mu}$ is a finite set as well (since $\mathfrak{B}'_{\lambda,\mu} \cong \mathfrak{B}_{\lambda,\mu,\ell,q}$ as sets). In view of (13.46.110), this rewrites as follows: The set

$$\left\{ \varphi : \{1,2,\ldots,\ell\} \to \{1,2,3,\ldots\} \mid \mu_j = \sum_{\substack{i \in \{1,2,\ldots,\ell\};\\ \varphi(i)=j}} \lambda_i \text{ for all } j \geq 1 \right\}$$

is a finite set. In other words, there are only finitely many maps $\varphi : \{1,2,\ldots,\ell\} \to \{1,2,3,\ldots\}$ satisfying $\left(\mu_j = \sum_{\substack{i \in \{1,2,\ldots,\ell\};\\ \varphi(i)=j}} \lambda_i \text{ for all } j \geq 1 \right)$. This solves Exercise 2.2.13(j).

(k) Let $\lambda \in \mathrm{Par}_n$. Let $\ell = \ell(\lambda)$. Thus, $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_\ell)$ (by the definition of $\ell(\lambda)$). Thus, $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_\ell) \in \mathbb{N}^\ell$. Therefore, for every map $\varphi : [\ell] \to \{1,2,3,\ldots\}$, a weak composition $\varphi_* \lambda \in \mathrm{WC}$ is defined (according to Definition 13.46.43, applied to $p = \ell$, $\alpha = \lambda$ and $\alpha_i = \lambda_i$). The definition of this weak composition $\varphi_* \lambda$ yields

$$\varphi_* \lambda = \left(\sum_{i \in \varphi^{-1}\{1\}} \lambda_i, \sum_{i \in \varphi^{-1}\{2\}} \lambda_i, \sum_{i \in \varphi^{-1}\{3\}} \lambda_i, \ldots \right) \tag{13.46.111}$$

for any map $\varphi : [\ell] \to \{1, 2, 3, \ldots\}$. Hence, for each map $\varphi : [\ell] \to \{1, 2, 3, \ldots\}$, we have the following chain of logical equivalences:

$$\begin{aligned}
&\left(\underbrace{\mu}_{=(\mu_1,\mu_2,\mu_3,\ldots)} = \underbrace{\varphi_*\lambda}_{\substack{=\left(\sum_{i\in\varphi^{-1}\{1\}}\lambda_i,\sum_{i\in\varphi^{-1}\{2\}}\lambda_i,\sum_{i\in\varphi^{-1}\{3\}}\lambda_i,\ldots\right) \\ \text{(by (13.46.111))}}} \right) \\
\Longleftrightarrow\ &\left((\mu_1, \mu_2, \mu_3, \ldots) = \left(\sum_{i\in\varphi^{-1}\{1\}} \lambda_i, \sum_{i\in\varphi^{-1}\{2\}} \lambda_i, \sum_{i\in\varphi^{-1}\{3\}} \lambda_i, \ldots \right) \right) \\
\Longleftrightarrow\ &\left(\mu_j = \underbrace{\sum_{i\in\varphi^{-1}\{j\}}}_{=\sum_{\substack{i\in[\ell];\\ \varphi(i)=j}}} \lambda_i \text{ for all } j \geq 1 \right) \\
\Longleftrightarrow\ &\left(\mu_j = \sum_{\substack{i\in[\ell];\\ \varphi(i)=j}} \lambda_i \text{ for all } j \geq 1 \right).
\end{aligned}$$

Thus,

$$\begin{aligned}
&\{\varphi : [\ell] \to \{1, 2, 3, \ldots\} \mid \mu = \varphi_*\lambda\} \\
&= \left\{ \varphi : [\ell] \to \{1, 2, 3, \ldots\} \mid \mu_j = \sum_{\substack{i\in[\ell];\\ \varphi(i)=j}} \lambda_i \text{ for all } j \geq 1 \right\}. \qquad (13.46.112)
\end{aligned}$$

We have $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_\ell)$ with $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_\ell > 0$ (since $\ell = \ell(\lambda)$). Hence, $p_\lambda = p_{\lambda_1} p_{\lambda_2} \cdots p_{\lambda_\ell}$ (by the definition of p_λ).

From $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_\ell > 0$, we also obtain ($\lambda_i > 0$ for each $i \in [\ell]$). Hence, Proposition 13.46.46 (applied to $\alpha = \lambda$ and $\alpha_i = \lambda_i$) yields

$$p_{\lambda_1} p_{\lambda_2} \cdots p_{\lambda_\ell} = \sum_{\varphi:[\ell]\to\{1,2,3,\ldots\}} \mathbf{x}^{\varphi_*\lambda}$$

in $\mathbf{k}[[\mathbf{x}]]$.

Now, let $\mu \in \operatorname{Par}_n$. Thus, $\mu \in \operatorname{Par}_n \subset \operatorname{Par} \subset \operatorname{WC}$. From $p_\lambda = p_{\lambda_1} p_{\lambda_2} \cdots p_{\lambda_\ell} = \sum_{\varphi:[\ell]\to\{1,2,3,\ldots\}} \mathbf{x}^{\varphi_* \lambda}$, we obtain

$$\begin{aligned}
&\left[\mathbf{x}^{\mu}\right]\left(p_\lambda\right) \\
&= \left[\mathbf{x}^{\mu}\right]\left(\sum_{\varphi:[\ell]\to\{1,2,3,\ldots\}} \mathbf{x}^{\varphi_*\lambda}\right) = \sum_{\varphi:[\ell]\to\{1,2,3,\ldots\}} \underbrace{\left[\mathbf{x}^{\mu}\right]\left(\mathbf{x}^{\varphi_*\lambda}\right)}_{\substack{=\delta_{\mu,\varphi_*\lambda} \\ \text{(by (13.46.1) (applied to } \alpha=\varphi_*\lambda))}} \\
&= \sum_{\varphi:[\ell]\to\{1,2,3,\ldots\}} \delta_{\mu,\varphi_*\lambda} \\
&= \sum_{\substack{\varphi:[\ell]\to\{1,2,3,\ldots\}; \\ \mu=\varphi_*\lambda}} \underbrace{\delta_{\mu,\varphi_*\lambda}}_{\substack{=1 \\ \text{(since } \mu=\varphi_*\lambda)}} + \sum_{\substack{\varphi:[\ell]\to\{1,2,3,\ldots\}; \\ \mu\neq\varphi_*\lambda}} \underbrace{\delta_{\mu,\varphi_*\lambda}}_{\substack{=0 \\ \text{(since } \mu\neq\varphi_*\lambda)}} \\
&= \sum_{\substack{\varphi:[\ell]\to\{1,2,3,\ldots\}; \\ \mu=\varphi_*\lambda}} 1 + \underbrace{\sum_{\substack{\varphi:[\ell]\to\{1,2,3,\ldots\}; \\ \mu\neq\varphi_*\lambda}} 0}_{=0} = \sum_{\substack{\varphi:[\ell]\to\{1,2,3,\ldots\}; \\ \mu=\varphi_*\lambda}} 1 \\
&= \left|\left\{\varphi:[\ell]\to\{1,2,3,\ldots\} \mid \mu=\varphi_*\lambda\right\}\right| \cdot 1 \\
&= \left|\left\{\varphi:[\ell]\to\{1,2,3,\ldots\} \mid \mu=\varphi_*\lambda\right\}\right| \\
&= \left|\left\{\varphi:[\ell]\to\{1,2,3,\ldots\} \mid \mu_j = \sum_{\substack{i\in[\ell]; \\ \varphi(i)=j}} \lambda_i \text{ for all } j\geq 1\right\}\right| \qquad (13.46.113)
\end{aligned}$$

(by (13.46.112)).

On the other hand, the definition of $b_{\lambda,\mu}$ shows that $b_{\lambda,\mu}$ is the number of all maps $\varphi : \{1,2,\ldots,\ell\} \to \{1,2,3,\ldots\}$ satisfying $\left(\mu_j = \sum_{\substack{i\in\{1,2,\ldots,\ell\}; \\ \varphi(i)=j}} \lambda_i \text{ for all } j\geq 1\right)$ (since $\ell = \ell(\lambda)$). In light of $\{1,2,\ldots,\ell\} = [\ell]$, this rewrites as follows: $b_{\lambda,\mu}$ is the number of all maps $\varphi : [\ell] \to \{1,2,3,\ldots\}$ satisfying $\left(\mu_j = \sum_{\substack{i\in[\ell]; \\ \varphi(i)=j}} \lambda_i \text{ for all } j\geq 1\right)$. In other words,

$$b_{\lambda,\mu} = \left|\left\{\varphi:[\ell]\to\{1,2,3,\ldots\} \mid \mu_j = \sum_{\substack{i\in[\ell]; \\ \varphi(i)=j}} \lambda_i \text{ for all } j\geq 1\right\}\right|.$$

Comparing this with (13.46.113), we obtain

$$\left[\mathbf{x}^{\mu}\right]\left(p_\lambda\right) = b_{\lambda,\mu}. \qquad (13.46.114)$$

Now, forget that we fixed μ. We thus have proven the equality (13.46.114) for every $\mu \in \operatorname{Par}_n$.

But $\lambda \in \operatorname{Par}_n$, so that $|\lambda| = n$. Hence, $n = \left|\underbrace{\lambda}_{=(\lambda_1,\lambda_2,\ldots,\lambda_\ell)}\right| = |(\lambda_1,\lambda_2,\ldots,\lambda_\ell)| = \lambda_1+\lambda_2+\cdots+\lambda_\ell$.

Let $i \in \{1,2,\ldots,\ell\}$. Then, p_{λ_i} is a homogeneous element of Λ having degree λ_i (because for each positive integer m, the element p_m is a homogeneous element of Λ having degree m).

Now, forget that we fixed i. We thus have shown that for each $i \in \{1,2,\ldots,\ell\}$, the element p_{λ_i} is a homogeneous element of Λ having degree λ_i. In other words, $p_{\lambda_1}, p_{\lambda_2}, \ldots, p_{\lambda_\ell}$ are homogeneous elements of Λ having degrees $\lambda_1, \lambda_2, \ldots, \lambda_\ell$, respectively. Hence, the product $p_{\lambda_1} p_{\lambda_2} \cdots p_{\lambda_\ell}$ of these elements is a homogeneous element of Λ having degree $\lambda_1+\lambda_2+\cdots+\lambda_\ell$. In light of $p_\lambda = p_{\lambda_1} p_{\lambda_2} \cdots p_{\lambda_\ell}$ and $n =$

$\lambda_1 + \lambda_2 + \cdots + \lambda_\ell$, this rewrites as follows: The element p_λ is a homogeneous element of Λ having degree n. In other words, $p_\lambda \in \Lambda_n$. Thus, Exercise 2.2.13(a) (applied to $f = p_\lambda$) yields

$$p_\lambda = \sum_{\mu \in \operatorname{Par}_n} \underbrace{\left(\left[\mathbf{x}^\mu\right]\left(p_\lambda\right)\right)}_{\substack{= b_{\lambda,\mu} \\ \text{(by (13.46.114))}}} m_\mu = \sum_{\mu \in \operatorname{Par}_n} b_{\lambda,\mu} m_\mu .$$

This solves Exercise 2.2.13(k).

Before we come to the solution of Exercise 2.2.13(l), let us state a simple lemma:

Lemma 13.46.48. *Let λ be a partition. Let μ be a weak composition. Then:*

(a) *We have $b_{\lambda,\mu} = \left|\mathfrak{B}'_{\lambda,\mu}\right|$.*

(b) *Let $p = \ell(\lambda)$. Let $q \in \mathbb{N}$ be such that $\mu = (\mu_1, \mu_2, \ldots, \mu_q)$ (that is, $\mu_i = 0$ for all integers $i > q$). Then, the set $\mathfrak{B}_{\lambda,\mu,p,q}$ is well-defined and satisfies $b_{\lambda,\mu} = \left|\mathfrak{B}_{\lambda,\mu,p,q}\right|$.*

Proof of Lemma 13.46.48. We know that λ is a partition.

We have $p = \ell(\lambda)$. Hence, the definition of $\mathfrak{B}'_{\lambda,\mu}$ yields

$$\mathfrak{B}'_{\lambda,\mu} = \left\{ \varphi : [p] \to \{1,2,3,\ldots\} \ \mid \ \underbrace{\mu_j = \sum_{\substack{i \in [p]; \\ \varphi(i) = j}} \lambda_i \text{ for all } j \in \{1,2,3,\ldots\}}_{\substack{\Longleftrightarrow \left(\mu_j = \sum_{\substack{i \in [p]; \\ \varphi(i)=j}} \lambda_i \text{ for all } j \geq 1\right) \\ \text{(since the } j \in \{1,2,3,\ldots\} \text{ are precisely the integers } j \geq 1\text{)}}} \right\}$$

$$= \left\{ \varphi : [p] \to \{1,2,3,\ldots\} \ \mid \ \mu_j = \sum_{\substack{i \in [p]; \\ \varphi(i) = j}} \lambda_i \text{ for all } j \geq 1 \right\}. \tag{13.46.115}$$

On the other hand, the definition of $b_{\lambda,\mu}$ shows that $b_{\lambda,\mu}$ is the number of all maps $\varphi : \{1,2,\ldots,p\} \to \{1,2,3,\ldots\}$ satisfying $\left(\mu_j = \sum_{\substack{i \in \{1,2,\ldots,p\}; \\ \varphi(i)=j}} \lambda_i \text{ for all } j \geq 1\right)$ (since $p = p(\lambda)$). In view of $\{1,2,\ldots,p\} = [p]$, this rewrites as follows: $b_{\lambda,\mu}$ is the number of all maps $\varphi : [p] \to \{1,2,3,\ldots\}$ satisfying $\left(\mu_j = \sum_{\substack{i \in [p]; \\ \varphi(i)=j}} \lambda_i \text{ for all } j \geq 1\right)$.

In other words,

$$b_{\lambda,\mu} = \left| \underbrace{\left\{ \varphi : [p] \to \{1,2,3,\ldots\} \ \mid \ \mu_j = \sum_{\substack{i \in [p]; \\ \varphi(i) = j}} \lambda_i \text{ for all } j \geq 1 \right\}}_{\substack{= \mathfrak{B}'_{\lambda,\mu} \\ \text{(by (13.46.115))}}} \right| = \left|\mathfrak{B}'_{\lambda,\mu}\right| .$$

This proves Lemma 13.46.48(a).

(b) From $p = \ell(\lambda)$, we obtain $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_p)$; thus, λ is a p-tuple in $\mathbb{N}^p$. Also, $\mu = (\mu_1, \mu_2, \ldots, \mu_q) \in \mathbb{N}^q$. Thus, μ is a q-tuple in $\mathbb{N}^q$. Hence, the set $\mathfrak{B}_{\lambda,\mu,p,q}$ is well-defined (since $\lambda \in \mathbb{N}^p$ and $\mu \in \mathbb{N}^q$). Proposition 13.46.39 yields that $\mathfrak{B}'_{\lambda,\mu} \cong \mathfrak{B}_{\lambda,\mu,p,q}$ as sets. But Lemma 13.46.48(a) shows that $b_{\lambda,\mu} = \left|\mathfrak{B}'_{\lambda,\mu}\right| = |\mathfrak{B}_{\lambda,\mu,p,q}|$ (since $\mathfrak{B}'_{\lambda,\mu} \cong \mathfrak{B}_{\lambda,\mu,p,q}$ as sets). Thus, Lemma 13.46.48(b) is proven. □

We now resume the solution of Exercise 2.2.13.

(l) Let $\lambda \in \operatorname{Par}_n$ and $\mu \in \operatorname{Par}_n$ be any partitions that don't satisfy $\mu \triangleright \lambda$. We must prove that $b_{\lambda,\mu} = 0$.

Indeed, assume the contrary. Thus, $b_{\lambda,\mu} \neq 0$. But Lemma 13.46.48(a) yields $b_{\lambda,\mu} = \left|\mathfrak{B}'_{\lambda,\mu}\right|$ (since μ is a weak composition (since μ is a partition)). Hence, $\left|\mathfrak{B}'_{\lambda,\mu}\right| = b_{\lambda,\mu} \neq 0$. In other words, $\mathfrak{B}'_{\lambda,\mu} \neq \varnothing$. Thus, Proposition 13.46.40 shows that $\mu \triangleright \lambda$. This contradicts the fact that we don't have $\mu \triangleright \lambda$.

This contradiction proves that our assumption was wrong. Hence, $b_{\lambda,\mu} = 0$ is proven. This solves Exercise 2.2.13(l).

(m) Let $\lambda \in \operatorname{Par}_n$. We must prove that $b_{\lambda,\lambda}$ is a positive integer.

Let $k = \ell(\lambda)$. Thus, $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_k)$ (by the definition of $\ell(\lambda)$). Clearly, λ is a weak composition (since λ is a partition). Hence, Lemma 13.46.48(b) (applied to $\mu = \lambda$, $p = k$ and $q = k$) shows that the set $\mathfrak{B}_{\lambda,\lambda,k,k}$ is well-defined and satisfies $b_{\lambda,\lambda} = |\mathfrak{B}_{\lambda,\lambda,k,k}|$. But $\lambda \in \operatorname{Par}_n \subset \operatorname{Par}$. Hence, Proposition 13.46.41(b) (applied to $\mu = \lambda$) shows that $\mathfrak{B}_{\lambda,\lambda,k,k}$ is a subgroup of $\mathfrak{S}_k$. Thus, the set $\mathfrak{B}_{\lambda,\lambda,k,k}$ contains the neutral element of $\mathfrak{S}_k$. Thus, the set $\mathfrak{B}_{\lambda,\lambda,k,k}$ contains at least one element, and therefore is nonempty. Thus, $|\mathfrak{B}_{\lambda,\lambda,k,k}| > 0$.

Hence, $b_{\lambda,\lambda} = |\mathfrak{B}_{\lambda,\lambda,k,k}| > 0$. Since $b_{\lambda,\lambda}$ is an integer, we can therefore conclude that $b_{\lambda,\lambda}$ is a positive integer. This solves Exercise 2.2.13(m).

(n) Let $\mu = (\mu_1, \mu_2, \ldots, \mu_k) \in \operatorname{Par}_n$ be a partition. Let $k = \ell(\mu)$. We must show that the integer $b_{\mu,\mu}$ is the size of the subgroup of $\mathfrak{S}_k$ consisting of all permutations $\sigma \in \mathfrak{S}_k$ having each i satisfy $\mu_{\sigma(i)} = \mu_i$. In particular, we must show that this subgroup is indeed a subgroup.

Clearly, μ is a weak composition (since μ is a partition). Hence, Lemma 13.46.48(b) (applied to $\lambda = \mu$, $p = k$ and $q = k$) shows that the set $\mathfrak{B}_{\mu,\mu,k,k}$ is well-defined and satisfies $b_{\mu,\mu} = |\mathfrak{B}_{\mu,\mu,k,k}|$.

We have $\mu \in \operatorname{Par}_n \subset \operatorname{Par}$. Hence, Proposition 13.46.41(a) shows that

$$\mathfrak{B}_{\mu,\mu,k,k} = \left\{\sigma \in \mathfrak{S}_k \ \mid \ \mu_{\sigma(i)} = \mu_i \text{ for each } i \in [k]\right\}. \tag{13.46.116}$$

Furthermore, Proposition 13.46.41(b) shows that $\mathfrak{B}_{\mu,\mu,k,k}$ is a subgroup of $\mathfrak{S}_k$.

Now,

$$\begin{aligned}
&\left(\text{the set of all permutations } \sigma \in \mathfrak{S}_k \text{ having each } i \text{ satisfy } \mu_{\sigma(i)} = \mu_i\right) \\
&= \left\{\sigma \in \mathfrak{S}_k \ \mid \ \text{each } i \text{ satisfies } \mu_{\sigma(i)} = \mu_i\right\} \\
&= \left\{\sigma \in \mathfrak{S}_k \ \mid \ \text{each } i \in \underbrace{\{1, 2, \ldots, k\}}_{=[k]} \text{ satisfies } \mu_{\sigma(i)} = \mu_i\right\} \\
&= \left\{\sigma \in \mathfrak{S}_k \ \mid \ \text{each } i \in [k] \text{ satisfies } \mu_{\sigma(i)} = \mu_i\right\} \\
&= \left\{\sigma \in \mathfrak{S}_k \ \mid \ \mu_{\sigma(i)} = \mu_i \text{ for each } i \in [k]\right\} \\
&= \mathfrak{B}_{\mu,\mu,k,k} \qquad (\text{by } (13.46.116)).
\end{aligned} \tag{13.46.117}$$

But recall that $\mathfrak{B}_{\mu,\mu,k,k}$ is a subgroup of $\mathfrak{S}_k$. In view of (13.46.117), this rewrites as follows: The set of all permutations $\sigma \in \mathfrak{S}_k$ having each i satisfy $\mu_{\sigma(i)} = \mu_i$ is a subgroup of $\mathfrak{S}_k$. This subgroup thus is the subgroup of $\mathfrak{S}_k$ consisting of all permutations $\sigma \in \mathfrak{S}_k$ having each i satisfy $\mu_{\sigma(i)} = \mu_i$. The size of this subgroup is clearly

$$\left|\underbrace{\left\{\sigma \in \mathfrak{S}_k \ \mid \ \text{each } i \text{ satisfies } \mu_{\sigma(i)} = \mu_i\right\}}_{=\mathfrak{B}_{\mu,\mu,k,k}}\right| = |\mathfrak{B}_{\mu,\mu,k,k}| = b_{\mu,\mu} \qquad (\text{since } b_{\mu,\mu} = |\mathfrak{B}_{\mu,\mu,k,k}|).$$

Thus, $b_{\mu,\mu}$ is the size of the subgroup of $\mathfrak{S}_k$ consisting of all permutations $\sigma \in \mathfrak{S}_k$ having each i satisfy $\mu_{\sigma(i)} = \mu_i$. (In particular, we have shown that this subgroup is indeed a subgroup.) This solves Exercise 2.2.13(n).

Before we come to the solution of Exercise 2.2.13(o), let us prove some further auxiliary results.

Proposition 13.46.49. *Let $\mu \in \operatorname{Par}$. Let $k = \ell(\mu)$. Proposition 13.46.41(b) shows that the set $\mathfrak{B}_{\mu,\mu,k,k}$ is a subgroup of $\mathfrak{S}_k$. Thus, $\mathfrak{B}_{\mu,\mu,k,k}$ is a group.*

Let $p \in \mathbb{N}$. Let $\lambda \in \mathbb{N}^p$. Then, the set $\mathfrak{B}_{\lambda,\mu,p,k}$ can be made into a left $\mathfrak{B}_{\mu,\mu,k,k}$-set (i.e., it can be equipped with an action of the group $\mathfrak{B}_{\mu,\mu,k,k}$ from the left) in such a way that the group $\mathfrak{B}_{\mu,\mu,k,k}$ acts freely on $\mathfrak{B}_{\lambda,\mu,p,k}$.

Proof of Proposition 13.46.49. For any $\alpha \in \mathfrak{B}_{\mu,\mu,k,k}$ and $\beta \in \mathfrak{B}_{\lambda,\mu,p,k}$, we have $\alpha \circ \beta \in \mathfrak{B}_{\lambda,\mu,p,k}$ [565]. Thus, we can try to define an action of the group $\mathfrak{B}_{\mu,\mu,k,k}$ on the set $\mathfrak{B}_{\lambda,\mu,p,k}$ from the left by setting

$$(\alpha\beta = \alpha \circ \beta \qquad \text{for all } \alpha \in \mathfrak{B}_{\mu,\mu,k,k} \text{ and } \beta \in \mathfrak{B}_{\lambda,\mu,p,k}). \tag{13.46.118}$$

In order to show that this definition actually defines an action of the group $\mathfrak{B}_{\mu,\mu,k,k}$ on the set $\mathfrak{B}_{\lambda,\mu,p,k}$, we need to prove the following two observations:

Observation 1: We have $\operatorname{id}_{\{1,2,\ldots,k\}} \circ \gamma = \gamma$ for all $\gamma \in \mathfrak{B}_{\lambda,\mu,p,k}$.

Observation 2: We have $(\alpha \circ \beta) \circ \gamma = \alpha \circ (\beta \circ \gamma)$ for all $\alpha \in \mathfrak{B}_{\mu,\mu,k,k}$, $\beta \in \mathfrak{B}_{\mu,\mu,k,k}$ and $\gamma \in \mathfrak{B}_{\lambda,\mu,p,k}$.

However, both Observation 1 and Observation 2 are obvious. Thus, we have shown that (13.46.118) actually defines an action of the group $\mathfrak{B}_{\mu,\mu,k,k}$ on the set $\mathfrak{B}_{\lambda,\mu,p,k}$. Consider this action. Thus, the set $\mathfrak{B}_{\lambda,\mu,p,k}$ has been made into a left $\mathfrak{B}_{\mu,\mu,k,k}$-set.

We shall now prove that the group $\mathfrak{B}_{\mu,\mu,k,k}$ acts freely on $\mathfrak{B}_{\lambda,\mu,p,k}$. In order to do so, we must prove the following observation:

Observation 3: If $\alpha \in \mathfrak{B}_{\mu,\mu,k,k}$, $\beta \in \mathfrak{B}_{\mu,\mu,k,k}$ and $\gamma \in \mathfrak{B}_{\lambda,\mu,p,k}$ are such that $\alpha \circ \gamma = \beta \circ \gamma$, then $\alpha = \beta$.

[*Proof of Observation 3:* Let $\alpha \in \mathfrak{B}_{\mu,\mu,k,k}$, $\beta \in \mathfrak{B}_{\mu,\mu,k,k}$ and $\gamma \in \mathfrak{B}_{\lambda,\mu,p,k}$ be such that $\alpha \circ \gamma = \beta \circ \gamma$. We must prove that $\alpha = \beta$.

We have $\alpha \in \mathfrak{B}_{\mu,\mu,k,k} = \{\varphi : [k] \to [k] \mid \mu = \varphi_* \mu\}$ (by the definition of $\mathfrak{B}_{\mu,\mu,k,k}$). In other words, α is a map $[k] \to [k]$ and satisfies $\mu = \alpha_* \mu$. The same argument (applied to β instead of α) shows that β is a map $[k] \to [k]$ and satisfies $\mu = \beta_* \mu$.

We have $\gamma \in \mathfrak{B}_{\lambda,\mu,p,k} = \{\varphi : [p] \to [k] \mid \mu = \varphi_* \lambda\}$ (by the definition of $\mathfrak{B}_{\lambda,\mu,p,k}$). In other words, γ is a map $[p] \to [k]$ and satisfies $\mu = \gamma_* \lambda$.

Write the p-tuple λ in the form $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_p)$.

Also, $\ell(\mu) = k$. Hence, $\mu = (\mu_1, \mu_2, \ldots, \mu_k) \in \mathbb{N}^k$ and $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_k > 0$. From $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_k > 0$, we obtain ($\mu_i > 0$ for each $i \in [k]$). Hence, Lemma 13.46.31 (applied to $q = k$, $\varphi = \gamma$, $\alpha = \lambda$, $\alpha_i = \lambda_i$, $\beta = \mu$ and $\beta_i = \mu_i$) shows that the map γ is surjective. Hence, we can cancel γ from the equality $\alpha \circ \gamma = \beta \circ \gamma$. Thus, we obtain $\alpha = \beta$. This proves Observation 3.]

Observation 3 shows that the group $\mathfrak{B}_{\mu,\mu,k,k}$ acts freely on $\mathfrak{B}_{\lambda,\mu,p,k}$ (by the definition of "acting freely"). Thus, the set $\mathfrak{B}_{\lambda,\mu,p,k}$ can be made into a left $\mathfrak{B}_{\mu,\mu,k,k}$-set in such a way that the group $\mathfrak{B}_{\mu,\mu,k,k}$ acts freely on $\mathfrak{B}_{\lambda,\mu,p,k}$ (namely, by using the above-defined left action of $\mathfrak{B}_{\mu,\mu,k,k}$ on $\mathfrak{B}_{\lambda,\mu,p,k}$). This proves Proposition 13.46.49. □

Let us now recall a basic fact from abstract algebra:

Proposition 13.46.50. *Let G be a finite group. Let X be a finite left G-set. Assume that G acts freely on X. Then, $|G| \mid |X|$.*

[565] *Proof.* Let $\alpha \in \mathfrak{B}_{\mu,\mu,k,k}$ and $\beta \in \mathfrak{B}_{\lambda,\mu,p,k}$. We must show that $\alpha \circ \beta \in \mathfrak{B}_{\lambda,\mu,p,k}$.

We have $\alpha \in \mathfrak{B}_{\mu,\mu,k,k} = \{\varphi : [k] \to [k] \mid \mu = \varphi_* \mu\}$ (by the definition of $\mathfrak{B}_{\mu,\mu,k,k}$). In other words, α is a map $[k] \to [k]$ and satisfies $\mu = \alpha_* \mu$.

We have $\beta \in \mathfrak{B}_{\lambda,\mu,p,k} = \{\varphi : [p] \to [k] \mid \mu = \varphi_* \lambda\}$ (by the definition of $\mathfrak{B}_{\lambda,\mu,p,k}$). In other words, β is a map $[p] \to [k]$ and satisfies $\mu = \beta_* \lambda$.

Now, Proposition 13.46.28 (applied to k, k, β, α and λ instead of q, r, φ, ψ and α) yields $(\alpha \circ \beta)_* \lambda = \alpha_* \left(\underbrace{\beta_* \lambda}_{=\mu} \right) = \alpha_* \mu = \mu$. In other words, $\mu = (\alpha \circ \beta)_* \lambda$. Hence, $\alpha \circ \beta$ is a map $[p] \to [k]$ (since α is a map $[k] \to [k]$, and since β is a map $[p] \to [k]$), and satisfies $\mu = (\alpha \circ \beta)_* \lambda$. In other words, $\alpha \circ \beta \in \{\varphi : [p] \to [k] \mid \mu = \varphi_* \lambda\}$. This rewrites as $\alpha \circ \beta \in \mathfrak{B}_{\lambda,\mu,p,k}$ (since $\mathfrak{B}_{\lambda,\mu,p,k} = \{\varphi : [p] \to [k] \mid \mu = \varphi_* \lambda\}$). Qed.

Proof of Proposition 13.46.50. This fact is well-known, so let us merely sketch the proof: The G-set X is a disjoint union of orbits (since every G-set is a disjoint union of orbits). Thus, the G-set X is a disjoint union of finitely many orbits (since X is finite). In other words, we have $X = O_1 \cup O_2 \cup \cdots \cup O_k$ for some list $(O_1, O_2, \ldots, O_k)$ of disjoint orbits of G on X. Consider this list $(O_1, O_2, \ldots, O_k)$.

Every $i \in \{1, 2, \ldots, k\}$ satisfies $|O_i| = |G|$ [566]. Hence, $\sum_{i=1}^k \underbrace{|O_i|}_{=|G|} = \sum_{i=1}^k |G| = k\,|G|$.

From $X = O_1 \cup O_2 \cup \cdots \cup O_k$, we obtain

$$|X| = |O_1 \cup O_2 \cup \cdots \cup O_k| = |O_1| + |O_2| + \cdots + |O_k| \qquad \text{(since the orbits } O_1, O_2, \ldots, O_k \text{ are disjoint)}$$
$$= \sum_{i=1}^{k} |O_i| = k\,|G|\,.$$

Thus, $|G| \mid k\,|G| = |X|$. This proves Proposition 13.46.50. □

We now resume the solution of Exercise 2.2.13.

(o) Let $\lambda \in \operatorname{Par}_n$ and $\mu \in \operatorname{Par}_n$. We must prove that $b_{\mu,\mu} \mid b_{\lambda,\mu}$.

Let $p = \ell(\lambda)$. Thus, $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_p)$ (by the definition of $\ell(\lambda)$). Hence, $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_p) \in \mathbb{N}^p$.

We have $\mu \in \operatorname{Par}_n \subset \operatorname{Par}$. Let $k = \ell(\mu)$. Proposition 13.46.41(b) shows that the set $\mathfrak{B}_{\mu,\mu,k,k}$ is a subgroup of $\mathfrak{S}_k$. Thus, $\mathfrak{B}_{\mu,\mu,k,k}$ is a finite group (since $\mathfrak{S}_k$ is a finite group).

But λ is a partition (since $\lambda \in \operatorname{Par}_n \subset \operatorname{Par}$). Also, μ is a partition (since $\mu \in \operatorname{Par}$), thus a weak composition. Furthermore, $k = \ell(\mu)$; thus, $\mu = (\mu_1, \mu_2, \ldots, \mu_k)$ (by the definition of $\ell(\mu)$). Hence, Lemma 13.46.48(b) (applied to $q = k$) shows that the set $\mathfrak{B}_{\lambda,\mu,p,k}$ is well-defined and satisfies $b_{\lambda,\mu} = |\mathfrak{B}_{\lambda,\mu,p,k}|$.

Furthermore, Lemma 13.46.48(b) (applied to k, k and μ instead of p, q and λ) shows that the set $\mathfrak{B}_{\mu,\mu,k,k}$ is well-defined and satisfies $b_{\mu,\mu} = |\mathfrak{B}_{\mu,\mu,k,k}|$.

The definition of $\mathfrak{B}_{\lambda,\mu,p,k}$ yields $\mathfrak{B}_{\lambda,\mu,p,k} = \{\varphi : [p] \to [k] \mid \mu = \varphi_* \lambda\} \subset \{\varphi : [p] \to [k]\} = [k]^{[p]}$. Thus, $\mathfrak{B}_{\lambda,\mu,p,k}$ is a finite set (since $[k]^{[p]}$ is a finite set).

Proposition 13.46.49 shows that the set $\mathfrak{B}_{\lambda,\mu,p,k}$ can be made into a left $\mathfrak{B}_{\mu,\mu,k,k}$-set (i.e., it can be equipped with an action of the group $\mathfrak{B}_{\mu,\mu,k,k}$ from the left) in such a way that the group $\mathfrak{B}_{\mu,\mu,k,k}$ acts freely on $\mathfrak{B}_{\lambda,\mu,p,k}$. Consider this action. Proposition 13.46.50 (applied to $G = \mathfrak{B}_{\mu,\mu,k,k}$ and $X = \mathfrak{B}_{\lambda,\mu,p,k}$) now yields $|\mathfrak{B}_{\mu,\mu,k,k}| \mid |\mathfrak{B}_{\lambda,\mu,p,k}|$. Thus, $b_{\mu,\mu} = |\mathfrak{B}_{\mu,\mu,k,k}| \mid |\mathfrak{B}_{\lambda,\mu,p,k}| = b_{\lambda,\mu}$ (since $b_{\lambda,\mu} = |\mathfrak{B}_{\lambda,\mu,p,k}|$). This solves Exercise 2.2.13(o).

13.47. **Solution to Exercise 2.2.14.**

Solution to Exercise 2.2.14. Let us first consider the polynomial ring $\mathbf{k}[x_1, x_2, x_3, \ldots]$ in countably many indeterminates $x_1, x_2, x_3, \ldots$.

Let $f : \mathbf{k}[x_1, x_2, x_3, \ldots] \to A$ be the $\mathbf{k}$-algebra homomorphism that sends $x_1, x_2, x_3, \ldots$ to $v_1, v_2, v_3, \ldots$, respectively. This is well-defined by the universal property of the polynomial ring $\mathbf{k}[x_1, x_2, x_3, \ldots]$ (since A is commutative). The $\mathbf{k}$-algebra homomorphism f is an instance of an evaluation homomorphism; it sends each polynomial $P \in \mathbf{k}[x_1, x_2, x_3, \ldots]$ to $P(v_1, v_2, v_3, \ldots) \in A$. In other words,

$$f(P) = P(v_1, v_2, v_3, \ldots) \qquad \text{for each } P \in \mathbf{k}[x_1, x_2, x_3, \ldots]\,. \tag{13.47.1}$$

The map f is a $\mathbf{k}$-algebra homomorphism, and thus is $\mathbf{k}$-linear.

For each partition $\lambda \in \operatorname{Par}$, we define a monomial $\mathbf{x}_\lambda \in \mathbf{k}[x_1, x_2, x_3, \ldots]$ by

$$\mathbf{x}_\lambda = x_{\lambda_1} x_{\lambda_2} \cdots x_{\lambda_{\ell(\lambda)}}.$$

(This is well-defined, since $\lambda_1, \lambda_2, \ldots, \lambda_{\ell(\lambda)}$ are positive integers whenever $\lambda \in \operatorname{Par}$.)

If λ is a partition and i is a positive integer, then $m_i(\lambda)$ shall denote the multiplicity of i in λ (that is, the number of parts of λ equal to i).

Let WC denote the set of all weak compositions.

[566] *Proof.* Let $i \in \{1, 2, \ldots, k\}$. Thus, O_i is an orbit of G on X. In other words, $O_i = Gy$ for some $y \in X$. Consider this y.

The elements gy for $g \in G$ are all distinct (because if $gy = hy$ for two elements $g, h \in G$, then we must have $g = h$ (because G acts freely on X)). Thus, the number of these elements is precisely $|G|$. In other words, $|\{gy \mid g \in G\}| = |G|$. In view of $\{gy \mid g \in G\} = Gy = O_i$, this rewrites as $|O_i| = |G|$. Qed.

Every partition λ can be uniquely written in the form $\lambda = (1^{m_1} 2^{m_2} 3^{m_3} \cdots)$ for some weak composition $(m_1, m_2, m_3, \ldots)$. Thus, the map

$$\begin{aligned} \mathbf{part} : \mathrm{WC} &\to \mathrm{Par}, \\ (m_1, m_2, m_3, \ldots) &\mapsto (1^{m_1} 2^{m_2} 3^{m_3} \cdots) \end{aligned}$$

is a bijection. The inverse of this bijection is the map

$$\begin{aligned} \mathbf{mults} : \mathrm{Par} &\to \mathrm{WC}, \\ \lambda &\mapsto (m_1(\lambda), m_2(\lambda), m_3(\lambda), \ldots). \end{aligned}$$

Thus, this map **mults** is a bijection, too (since it is the inverse of a bijection).

It is easy to see that

$$\mathbf{x}_\lambda = \mathbf{x}^{\mathbf{mults}(\lambda)} \qquad \text{for any } \lambda \in \mathrm{Par} \tag{13.47.2}$$

[567]. In other words, $(\mathbf{x}_\lambda)_{\lambda \in \mathrm{Par}} = \left(\mathbf{x}^{\mathbf{mults}(\lambda)}\right)_{\lambda \in \mathrm{Par}}$.

The monomials in $\mathbf{k}[x_1, x_2, x_3, \ldots]$ have the form $\mathbf{x}^\alpha$ for $\alpha \in \mathrm{WC}$. It is well-known that these monomials form a basis of the $\mathbf{k}$-module $\mathbf{k}[x_1, x_2, x_3, \ldots]$ (since any polynomial ring has a basis consisting of the monomials). In other words, the family $(\mathbf{x}^\alpha)_{\alpha \in \mathrm{WC}}$ is a basis of the $\mathbf{k}$-module $\mathbf{k}[x_1, x_2, x_3, \ldots]$. The family $\left(\mathbf{x}^{\mathbf{mults}(\lambda)}\right)_{\lambda \in \mathrm{Par}}$ is a reindexing of this basis $(\mathbf{x}^\alpha)_{\alpha \in \mathrm{WC}}$ (since $\mathbf{mults} : \mathrm{Par} \to \mathrm{WC}$ is a bijection), and thus must also be a basis of the $\mathbf{k}$-module $\mathbf{k}[x_1, x_2, x_3, \ldots]$ (since a reindexing of a basis of a $\mathbf{k}$-module is always a basis itself). In view of $(\mathbf{x}_\lambda)_{\lambda \in \mathrm{Par}} = \left(\mathbf{x}^{\mathbf{mults}(\lambda)}\right)_{\lambda \in \mathrm{Par}}$, we can rewrite this as follows: The family $(\mathbf{x}_\lambda)_{\lambda \in \mathrm{Par}}$ is a basis of the $\mathbf{k}$-module $\mathbf{k}[x_1, x_2, x_3, \ldots]$. Hence, this family $(\mathbf{x}_\lambda)_{\lambda \in \mathrm{Par}}$ is $\mathbf{k}$-linearly independent and spans the $\mathbf{k}$-module $\mathbf{k}[x_1, x_2, x_3, \ldots]$.

Let us now step to the solution of Exercise 2.2.14:

(a) Let us use the following notation: If U is any $\mathbf{k}$-module, and if $(u_i)_{i \in I}$ is any family of elements of U, then $\langle u_i \mid i \in I \rangle_{\mathbf{k}}$ shall denote the $\mathbf{k}$-submodule of U spanned by this family $(u_i)_{i \in I}$.

We know that the family $(\mathbf{x}_\lambda)_{\lambda \in \mathrm{Par}}$ spans the $\mathbf{k}$-module $\mathbf{k}[x_1, x_2, x_3, \ldots]$. In other words,

$$\mathbf{k}[x_1, x_2, x_3, \ldots] = \langle \mathbf{x}_\lambda \mid \lambda \in \mathrm{Par} \rangle_{\mathbf{k}}.$$

[567] *Proof of (13.47.2):* Let $\lambda \in \mathrm{Par}$. Then, λ_p is a positive integer for each $p \in \{1, 2, \ldots, \ell(\lambda)\}$.

We have $\lambda = \left(\lambda_1, \lambda_2, \ldots, \lambda_{\ell(\lambda)}\right)$, and the entries $\lambda_1, \lambda_2, \ldots, \lambda_{\ell(\lambda)}$ of λ are positive integers, whereas each integer $p > \ell(\lambda)$ satisfies $\lambda_p = 0$. Thus, the parts of λ are $\lambda_1, \lambda_2, \ldots, \lambda_{\ell(\lambda)}$ (since a *part* of a partition means a nonzero entry of the partition). For each $i \in \{1, 2, 3, \ldots\}$, we have

$$\begin{aligned} m_i(\lambda) &= (\text{the multiplicity of } i \text{ in } \lambda) \qquad (\text{by the definition of } m_i(\lambda)) \\ &= (\text{the number of parts of } \lambda \text{ equal to } i) \\ &= (\text{the number of } p \in \{1, 2, \ldots, \ell(\lambda)\} \text{ satisfying } \lambda_p = i) \end{aligned}$$

(since the parts of λ are $\lambda_1, \lambda_2, \ldots, \lambda_{\ell(\lambda)}$). Hence, for each $i \in \{1, 2, 3, \ldots\}$, we have

$$(\text{the number of } p \in \{1, 2, \ldots, \ell(\lambda)\} \text{ satisfying } \lambda_p = i) = m_i(\lambda). \tag{13.47.3}$$

The definition of **mults** yields $\mathbf{mults}(\lambda) = (m_1(\lambda), m_2(\lambda), m_3(\lambda), \ldots)$. Thus, the definition of $\mathbf{x}^{\mathbf{mults}(\lambda)}$ yields

$$\mathbf{x}^{\mathbf{mults}(\lambda)} = x_1^{m_1(\lambda)} x_2^{m_2(\lambda)} x_3^{m_3(\lambda)} \cdots = \prod_{i=1}^{\infty} x_i^{m_i(\lambda)}. \tag{13.47.4}$$

But the definition of $\mathbf{x}_\lambda$ yields

$$\mathbf{x}_\lambda = x_{\lambda_1} x_{\lambda_2} \cdots x_{\lambda_{\ell(\lambda)}} = \prod_{p \in \{1, 2, \ldots, \ell(\lambda)\}} x_{\lambda_p} = \prod_{i=1}^{\infty} \prod_{\substack{p \in \{1, 2, \ldots, \ell(\lambda)\}; \\ \lambda_p = i}} x_{\lambda_p}$$

(here, we have split the product according to the value of λ_p, since λ_p is a positive integer for each $p \in \{1, 2, \ldots, \ell(\lambda)\}$). Hence,

$$\mathbf{x}_\lambda = \prod_{i=1}^{\infty} \prod_{\substack{p \in \{1, 2, \ldots, \ell(\lambda)\}; \\ \lambda_p = i}} \underbrace{x_{\lambda_p}}_{\substack{= x_i \\ (\text{since } \lambda_p = i)}} = \prod_{i=1}^{\infty} \underbrace{\prod_{\substack{p \in \{1, 2, \ldots, \ell(\lambda)\}; \\ \lambda_p = i}} x_i}_{\substack{= x_i^{(\text{the number of } p \in \{1, 2, \ldots, \ell(\lambda)\} \text{ satisfying } \lambda_p = i)} = x_i^{m_i(\lambda)} \\ (\text{by (13.47.3)})}} = \prod_{i=1}^{\infty} x_i^{m_i(\lambda)}.$$

Comparing this with (13.47.4), we obtain $\mathbf{x}_\lambda = \mathbf{x}^{\mathbf{mults}(\lambda)}$. This proves (13.47.2).

We have

$$f\left(\mathbf{x}_\lambda\right) = v_\lambda \qquad \text{for any } \lambda \in \operatorname{Par} \tag{13.47.5}$$

[568].

But A is commutative; hence,

$$\begin{aligned}
&\left(\text{the } \mathbf{k}\text{-subalgebra of } A \text{ generated by } v_1, v_2, v_3, \ldots\right) \\
&= \left\{ \underbrace{P\left(v_1, v_2, v_3, \ldots\right)}_{\substack{=f(P) \\ \text{(by (13.47.1))}}} \ \mid \ P \in \mathbf{k}\left[x_1, x_2, x_3, \ldots\right] \right\} \\
&= \left\{ f\left(P\right) \ \mid \ P \in \mathbf{k}\left[x_1, x_2, x_3, \ldots\right] \right\} = f\left(\underbrace{\mathbf{k}\left[x_1, x_2, x_3, \ldots\right]}_{=\left\langle \mathbf{x}_\lambda \ \mid \ \lambda \in \operatorname{Par} \right\rangle_{\mathbf{k}}} \right) \\
&= f\left(\left\langle \mathbf{x}_\lambda \ \mid \ \lambda \in \operatorname{Par} \right\rangle_{\mathbf{k}}\right) = \left\langle \underbrace{f\left(\mathbf{x}_\lambda\right)}_{\substack{=v_\lambda \\ \text{(by (13.47.5))}}} \ \mid \ \lambda \in \operatorname{Par} \right\rangle_{\mathbf{k}} \qquad \left(\text{since the map } f \text{ is } \mathbf{k}\text{-linear}\right) \\
&= \left\langle v_\lambda \ \mid \ \lambda \in \operatorname{Par} \right\rangle_{\mathbf{k}} = \left(\text{the } \mathbf{k}\text{-submodule of } A \text{ spanned by the family } \left(v_\lambda\right)_{\lambda \in \operatorname{Par}}\right).
\end{aligned}$$

This solves Exercise 2.2.14(a).

(b) We have the following chain of logical equivalences:

$$\begin{aligned}
&\left(\text{the elements } v_1, v_2, v_3, \ldots \text{ generate the } \mathbf{k}\text{-algebra } A\right) \\
\Longleftrightarrow\ &\left(\text{the } \mathbf{k}\text{-subalgebra of } A \text{ generated by } v_1, v_2, v_3, \ldots \text{ is } A\right) \\
\Longleftrightarrow\ &\left(\text{the } \mathbf{k}\text{-submodule of } A \text{ spanned by the family } \left(v_\lambda\right)_{\lambda \in \operatorname{Par}} \text{ is } A\right) \\
&\left(\begin{array}{c} \text{since Exercise 2.2.14(a) shows that} \\ \text{the } \mathbf{k}\text{-subalgebra of } A \text{ generated by } v_1, v_2, v_3, \ldots \\ \text{is the } \mathbf{k}\text{-submodule of } A \text{ spanned by the family } \left(v_\lambda\right)_{\lambda \in \operatorname{Par}} \end{array}\right) \\
\Longleftrightarrow\ &\left(\text{the family } \left(v_\lambda\right)_{\lambda \in \operatorname{Par}} \text{ spans the } \mathbf{k}\text{-module } A\right).
\end{aligned}$$

This solves Exercise 2.2.14(b).

(c) We shall prove the "$\Longrightarrow$" and the "$\Longleftarrow$" directions of Exercise 2.2.14(c) separately:

$\Longrightarrow$: Assume that the elements $v_1, v_2, v_3, \ldots$ are algebraically independent over $\mathbf{k}$. We must show that the family $\left(v_\lambda\right)_{\lambda \in \operatorname{Par}}$ is $\mathbf{k}$-linearly independent.

Let $\left(a_\lambda\right)_{\lambda \in \operatorname{Par}} \in \mathbf{k}^{\operatorname{Par}}$ be a family of scalars such that (all but finitely many $\lambda \in \operatorname{Par}$ satisfy $a_\lambda = 0$) and $\sum_{\lambda \in \operatorname{Par}} a_\lambda v_\lambda = 0$. We shall show that $\left(a_\lambda\right)_{\lambda \in \operatorname{Par}} = \left(0\right)_{\lambda \in \operatorname{Par}}$.

[568] *Proof of (13.47.5):* Let $\lambda \in \operatorname{Par}$. Then, the definition of $\mathbf{x}_\lambda$ yields $\mathbf{x}_\lambda = x_{\lambda_1} x_{\lambda_2} \cdots x_{\lambda_{\ell(\lambda)}}$. Applying the map f to both sides of this equality, we obtain

$$\begin{aligned}
f\left(\mathbf{x}_\lambda\right) &= f\left(x_{\lambda_1} x_{\lambda_2} \cdots x_{\lambda_{\ell(\lambda)}}\right) = f\left(x_{\lambda_1}\right) f\left(x_{\lambda_2}\right) \cdots f\left(x_{\lambda_{\ell(\lambda)}}\right) \qquad \left(\text{since } f \text{ is a } \mathbf{k}\text{-algebra homomorphism}\right) \\
&= \prod_{i=1}^{\ell(\lambda)} \underbrace{f\left(x_{\lambda_i}\right)}_{\substack{=v_{\lambda_i} \\ \text{(by the definition of } f\text{)}}} = \prod_{i=1}^{\ell(\lambda)} v_{\lambda_i} = v_{\lambda_1} v_{\lambda_2} \cdots v_{\lambda_{\ell(\lambda)}}.
\end{aligned}$$

Comparing this with

$$v_\lambda = v_{\lambda_1} v_{\lambda_2} \cdots v_{\lambda_{\ell(\lambda)}} \qquad \left(\text{by the definition of } v_\lambda\right),$$

we obtain $f\left(\mathbf{x}_\lambda\right) = v_\lambda$. This proves (13.47.5).

Define a polynomial $P \in \mathbf{k}[x_1, x_2, x_3, \ldots]$ by $P = \sum_{\lambda \in \operatorname{Par}} a_\lambda \mathbf{x}_\lambda$. Then,

$$f\left(\underbrace{P}_{=\sum_{\lambda\in\operatorname{Par}} a_\lambda \mathbf{x}_\lambda}\right) = f\left(\sum_{\lambda\in\operatorname{Par}} a_\lambda \mathbf{x}_\lambda\right) = \sum_{\lambda\in\operatorname{Par}} a_\lambda \underbrace{f(\mathbf{x}_\lambda)}_{\substack{=v_\lambda \\ \text{(by (13.47.5))}}} \qquad \text{(since the map } f \text{ is } \mathbf{k}\text{-linear)}$$
$$= \sum_{\lambda\in\operatorname{Par}} a_\lambda v_\lambda = 0.$$

Comparing this with (13.47.1), we obtain $P(v_1, v_2, v_3, \ldots) = 0$. This entails $P = 0$, because the elements $v_1, v_2, v_3, \ldots$ are algebraically independent. Comparing this with $P = \sum_{\lambda\in\operatorname{Par}} a_\lambda \mathbf{x}_\lambda$, we obtain $\sum_{\lambda\in\operatorname{Par}} a_\lambda \mathbf{x}_\lambda = 0$. But since the family $(\mathbf{x}_\lambda)_{\lambda\in\operatorname{Par}}$ is $\mathbf{k}$-linearly independent, this entails that $(a_\lambda)_{\lambda\in\operatorname{Par}} = (0)_{\lambda\in\operatorname{Par}}$.

Forget that we fixed $(a_\lambda)_{\lambda\in\operatorname{Par}}$. We thus have proved that if $(a_\lambda)_{\lambda\in\operatorname{Par}} \in \mathbf{k}^{\operatorname{Par}}$ is a family of scalars such that (all but finitely many $\lambda \in \operatorname{Par}$ satisfy $a_\lambda = 0$) and $\sum_{\lambda\in\operatorname{Par}} a_\lambda v_\lambda = 0$, then $(a_\lambda)_{\lambda\in\operatorname{Par}} = (0)_{\lambda\in\operatorname{Par}}$. In other words, the family $(v_\lambda)_{\lambda\in\operatorname{Par}}$ is $\mathbf{k}$-linearly independent. This proves the "$\Longrightarrow$" direction of Exercise 2.2.14(c).

$\Longleftarrow$: Assume that the family $(v_\lambda)_{\lambda\in\operatorname{Par}}$ is $\mathbf{k}$-linearly independent. We shall show that the elements $v_1, v_2, v_3, \ldots$ are algebraically independent over $\mathbf{k}$.

Indeed, let $P \in \mathbf{k}[x_1, x_2, x_3, \ldots]$ be a polynomial that satisfies $P(v_1, v_2, v_3, \ldots) = 0$. We shall show that $P = 0$.

Recall that the family $(\mathbf{x}_\lambda)_{\lambda\in\operatorname{Par}}$ spans the $\mathbf{k}$-module $\mathbf{k}[x_1, x_2, x_3, \ldots]$. Hence, we can write the polynomial $P \in \mathbf{k}[x_1, x_2, x_3, \ldots]$ as a $\mathbf{k}$-linear combination of this family. In other words, there exists a family $(a_\lambda)_{\lambda\in\operatorname{Par}} \in \mathbf{k}^{\operatorname{Par}}$ of scalars such that (all but finitely many $\lambda \in \operatorname{Par}$ satisfy $a_\lambda = 0$) and $P = \sum_{\lambda\in\operatorname{Par}} a_\lambda \mathbf{x}_\lambda$. Consider this $(a_\lambda)_{\lambda\in\operatorname{Par}}$.

Applying the map f to both sides of the equality $P = \sum_{\lambda\in\operatorname{Par}} a_\lambda \mathbf{x}_\lambda$, we obtain

$$f(P) = f\left(\sum_{\lambda\in\operatorname{Par}} a_\lambda \mathbf{x}_\lambda\right) = \sum_{\lambda\in\operatorname{Par}} a_\lambda \underbrace{f(\mathbf{x}_\lambda)}_{\substack{=v_\lambda \\ \text{(by (13.47.5))}}} \qquad \text{(since the map } f \text{ is } \mathbf{k}\text{-linear)}$$
$$= \sum_{\lambda\in\operatorname{Par}} a_\lambda v_\lambda.$$

But (13.47.1) yields

$$f(P) = P(v_1, v_2, v_3, \ldots) = 0.$$

Comparing these two equalities, we obtain $\sum_{\lambda\in\operatorname{Par}} a_\lambda v_\lambda = 0$. This entails that $(a_\lambda)_{\lambda\in\operatorname{Par}} = (0)_{\lambda\in\operatorname{Par}}$ (since the family $(v_\lambda)_{\lambda\in\operatorname{Par}}$ is $\mathbf{k}$-linearly independent). In other words, $a_\lambda = 0$ for each $\lambda \in \operatorname{Par}$. Hence, $P = \sum_{\lambda\in\operatorname{Par}} \underbrace{a_\lambda}_{=0} \mathbf{x}_\lambda = \sum_{\lambda\in\operatorname{Par}} 0\mathbf{x}_\lambda = 0$.

Now, forget that we fixed P. We thus have showed that every polynomial $P \in \mathbf{k}[x_1, x_2, x_3, \ldots]$ satisfying $P(v_1, v_2, v_3, \ldots) = 0$ must satisfy $P = 0$. In other words, the elements $v_1, v_2, v_3, \ldots$ are algebraically independent over $\mathbf{k}$. This proves the "$\Longleftarrow$" direction of Exercise 2.2.14(c).

Thus, both "$\Longrightarrow$" and "$\Longleftarrow$" directions of Exercise 2.2.14(c) are solved.

13.48. **Solution to Exercise 2.2.15.** *Solution to Exercise 2.2.15.* Before we start solving the exercise, let us recall our notion of a "monomial". For us, a *pure monomial* is (formally speaking) just a symbol $\mathbf{x}^\alpha$ indexed by a weak composition α. (Thus, pure monomials are combinatorial objects; in particular, they have nothing to do with the ground ring $\mathbf{k}$ [569], and do not "come with coefficients".)

We let Mon be the set of all pure monomials. We define the product of two pure monomials $\mathbf{x}^\alpha$ and $\mathbf{x}^\beta$ by $\mathbf{x}^\alpha \mathbf{x}^\beta = \mathbf{x}^{\alpha+\beta}$, where $\alpha+\beta$ denotes the entrywise sum of the two weak compositions α and β (in other words, if $\alpha = (\alpha_1, \alpha_2, \alpha_3, \ldots)$ and $\beta = (\beta_1, \beta_2, \beta_3, \ldots)$, then $\alpha + \beta = (\alpha_1 + \beta_1, \alpha_2 + \beta_2, \alpha_3 + \beta_3, \ldots)$). This notion

[569] In particular, the pure monomials $\mathbf{x}^\alpha$ for different weak compositions α are distinct, although their images in the polynomial ring $\mathbf{k}[\mathbf{x}]$ are equal when $\mathbf{k} = 0$.

of product makes the set Mon of all pure monomials into a monoid. The neutral element of this monoid Mon is the pure monomial $\mathbf{x}^{(0,0,0,\ldots)}$, which we denote by 1. Whenever i is a positive integer, we write x_i for the pure monomial $\mathbf{x}^{t_i}$, where t_i is the weak composition $(\delta_{1,i}, \delta_{2,i}, \delta_{3,i}, \ldots) = \left(\underbrace{0, 0, \ldots, 0}_{i-1 \text{ zeroes}}, 1, 0, 0, 0, \ldots\right)$. This notation allows us to rewrite any pure monomial $\mathbf{x}^\alpha$ in the form $x_1^{\alpha_1} x_2^{\alpha_2} x_3^{\alpha_3} \cdots = \prod_{i \geq 1} x_i^{\alpha_i}$, where α is written in the form $\alpha = (\alpha_1, \alpha_2, \alpha_3, \ldots)$.

So far we have only discussed pure monomials, but not the actual monomials we encounter in polynomial rings or rings of power series. In practice, however, most authors (including us) identify pure monomials with actual monomials. Let us now introduce this identification.

For each weak composition α, we identify the pure monomial $\mathbf{x}^\alpha \in \operatorname{Mon}$ with the monomial $\mathbf{x}^\alpha$ in the polynomial ring $\mathbf{k}[\mathbf{x}]$. This identification is mostly harmless, because the map

$$\operatorname{Mon} \to \mathbf{k}[\mathbf{x}], \qquad \mathbf{x}^\alpha \mapsto \mathbf{x}^\alpha$$

is a monoid homomorphism (from Mon to the multiplicative monoid of $\mathbf{k}[\mathbf{x}]$). [570] Using this injectivity, we thus can consider every monomial $\mathbf{x}^\alpha \in \operatorname{Mon}$ as an element of $\mathbf{k}[\mathbf{x}]$, and thus also as an element of $\mathbf{k}[[\mathbf{x}]]$. Whenever we write sums of monomials (such as $\mathbf{x}^\alpha + \mathbf{x}^\beta$ or $\sum_{\mathfrak{m} \in \operatorname{Mon}} \mathfrak{m}$), we always mean these sums to be computed in $\mathbf{k}[\mathbf{x}]$ or $\mathbf{k}[[\mathbf{x}]]$.

For any pure monomial $\mathbf{x}^\alpha$, we let $\deg(\mathbf{x}^\alpha)$ denote the nonnegative integer $\alpha_1 + \alpha_2 + \alpha_3 + \cdots$. This integer is called the *degree* of the monomial $\mathbf{x}^\alpha$.

Now, we notice that every pure monomial $\mathfrak{m} \in \operatorname{Mon}$ can be uniquely represented in the form $\mathbf{x}^\alpha$ for a weak composition α. Hence, we can substitute $\mathbf{x}^\alpha$ for $\mathfrak{m}$ in the sum $\sum_{\mathfrak{m} \in \operatorname{Mon}} \mathfrak{m} t^{\deg \mathfrak{m}}$. We thus obtain

$$\sum_{\mathfrak{m} \in \operatorname{Mon}} \mathfrak{m} t^{\deg \mathfrak{m}} = \sum_{\substack{\alpha \text{ is a weak} \\ \text{composition}}} \mathbf{x}^\alpha t^{\deg(\mathbf{x}^\alpha)} = \sum_{\substack{(\alpha_1,\alpha_2,\alpha_3,\ldots) \text{ is a weak} \\ \text{composition}}} \underbrace{\mathbf{x}^{(\alpha_1,\alpha_2,\alpha_3,\ldots)}}_{=\prod_{i\geq 1} x_i^{\alpha_i}} \underbrace{t^{\deg\left(\mathbf{x}^{(\alpha_1,\alpha_2,\alpha_3,\ldots)}\right)}}_{\substack{=t^{\alpha_1+\alpha_2+\alpha_3+\cdots} \\ (\text{since } \deg\left(\mathbf{x}^{(\alpha_1,\alpha_2,\alpha_3,\ldots)}\right)=\alpha_1+\alpha_2+\alpha_3+\cdots)}}$$

$$\left(\text{here, we renamed the summation index } \alpha \text{ as } (\alpha_1, \alpha_2, \alpha_3, \ldots)\right)$$

$$= \sum_{\substack{(\alpha_1,\alpha_2,\alpha_3,\ldots) \text{ is a weak} \\ \text{composition}}} \left(\prod_{i \geq 1} x_i^{\alpha_i}\right) \underbrace{t^{\alpha_1+\alpha_2+\alpha_3+\cdots}}_{=t^{\sum_{i\geq 1}\alpha_i}=\prod_{i\geq 1} t^{\alpha_i}} = \sum_{\substack{(\alpha_1,\alpha_2,\alpha_3,\ldots) \text{ is a weak} \\ \text{composition}}} \underbrace{\left(\prod_{i \geq 1} x_i^{\alpha_i}\right)\left(\prod_{i\geq 1} t^{\alpha_i}\right)}_{=\prod_{i\geq 1}\left(x_i^{\alpha_i} t^{\alpha_i}\right)}$$

$$= \sum_{\substack{(\alpha_1,\alpha_2,\alpha_3,\ldots) \text{ is a weak} \\ \text{composition}}} \prod_{i \geq 1} \underbrace{\left(x_i^{\alpha_i} t^{\alpha_i}\right)}_{=(x_i t)^{\alpha_i}} = \sum_{\substack{(\alpha_1,\alpha_2,\alpha_3,\ldots) \text{ is a weak} \\ \text{composition}}} \prod_{i \geq 1} (x_i t)^{\alpha_i}$$

$$= \sum_{\substack{(k_1,k_2,k_3,\ldots) \text{ is a weak} \\ \text{composition}}} \prod_{i \geq 1} (x_i t)^{k_i}$$

(here, we renamed the summation index $(\alpha_1, \alpha_2, \alpha_3, \ldots)$ as $(k_1, k_2, k_3, \ldots)$). Compared with

$$\underbrace{\prod_{i=1}^{\infty}}_{=\prod_{i\geq 1}} \underbrace{(1 - x_i t)^{-1}}_{\substack{=\sum_{k\in\mathbb{N}} (x_i t)^k \\ (\text{by the formula for the} \\ \text{geometric series})}} = \prod_{i \geq 1} \sum_{k \in \mathbb{N}} (x_i t)^k = \sum_{\substack{(k_1,k_2,k_3,\ldots) \text{ is a weak} \\ \text{composition}}} \prod_{i \geq 1} (x_i t)^{k_i}$$

$$(\text{by the product rule}),$$

this yields

$$\sum_{\mathfrak{m} \in \operatorname{Mon}} \mathfrak{m} t^{\deg \mathfrak{m}} = \prod_{i=1}^{\infty} (1 - x_i t)^{-1}. \qquad (13.48.1)$$

[570] Also, this map is injective unless $\mathbf{k} = 0$. However, we do not need the injectivity of this map, because we will not derive any equalities between pure monomials from equalities between monomials in $\mathbf{k}[\mathbf{x}]$.

On the other hand, every pure monomial $\mathfrak{m} \in \operatorname{Mon}$ can be uniquely represented in the form $x_{i_1} x_{i_2} \cdots x_{i_n}$ for some $n \in \mathbb{N}$ and some weakly increasing n-tuple $(i_1, i_2, \ldots, i_n)$ of positive integers.[571] Hence, we can substitute $x_{i_1} x_{i_2} \cdots x_{i_n}$ for $\mathfrak{m}$ in the sum $\sum_{\mathfrak{m} \in \operatorname{Mon}} \mathfrak{m} t^{\deg \mathfrak{m}}$. We thus obtain

$$\begin{aligned}
\sum_{\mathfrak{m} \in \operatorname{Mon}} \mathfrak{m} t^{\deg \mathfrak{m}} &= \sum_{n \in \mathbb{N}} \underbrace{\sum_{\substack{(i_1, i_2, \ldots, i_n) \text{ is a} \\ \text{weakly increasing} \\ n\text{-tuple of positive integers}}}}_{= \sum_{i_1 \leq i_2 \leq \cdots \leq i_n}} (x_{i_1} x_{i_2} \cdots x_{i_n}) \underbrace{t^{\deg\left(x_{i_1} x_{i_2} \cdots x_{i_n}\right)}}_{\substack{= t^n \\ (\text{since } \deg\left(x_{i_1} x_{i_2} \cdots x_{i_n}\right) = n)}} \\
&= \sum_{n \in \mathbb{N}} \sum_{i_1 \leq i_2 \leq \cdots \leq i_n} (x_{i_1} x_{i_2} \cdots x_{i_n}) t^n = \sum_{n \in \mathbb{N}} \underbrace{\left(\sum_{i_1 \leq i_2 \leq \cdots \leq i_n} x_{i_1} x_{i_2} \cdots x_{i_n} \right)}_{\substack{= h_n \\ (\text{since (2.2.3) yields} \\ h_n = \sum_{i_1 \leq i_2 \leq \cdots \leq i_n} x_{i_1} x_{i_2} \cdots x_{i_n})}} t^n \\
&= \sum_{n \in \mathbb{N}} \underbrace{h_n}_{= h_n(\mathbf{x})} t^n = \sum_{n \geq 0} h_n(\mathbf{x}) t^n = \underbrace{h_0(\mathbf{x})}_{= h_0 = 1} + h_1(\mathbf{x}) t + h_2(\mathbf{x}) t^2 + \cdots \\
&= 1 + h_1(\mathbf{x}) t + h_2(\mathbf{x}) t^2 + \cdots .
\end{aligned}$$

Compared with (13.48.1), this yields

$$\prod_{i=1}^{\infty} (1 - x_i t)^{-1} = 1 + h_1(\mathbf{x}) t + h_2(\mathbf{x}) t^2 + \cdots = \sum_{n \geq 0} h_n(\mathbf{x}) t^n.$$

Thus, the first of the two identities that we need to prove is proven.

Next, let us define the notion of a *squarefree monomial*. Indeed, let us say that a pure monomial $\mathbf{x}^\alpha \in \operatorname{Mon}$ is *squarefree* if and only if every entry of the weak composition α belongs to $\{0, 1\}$. Thus, of course, every squarefree monomial $\mathfrak{m} \in \operatorname{Mon}$ can be uniquely represented in the form $\mathbf{x}^\alpha$ for a weak composition α whose every entry belongs to $\{0, 1\}$. Hence, we can substitute $\mathbf{x}^\alpha$ for $\mathfrak{m}$ in the sum $\sum_{\substack{\mathfrak{m} \in \operatorname{Mon}; \\ \mathfrak{m} \text{ is squarefree}}} \mathfrak{m} t^{\deg \mathfrak{m}}$. We thus

[571] Indeed, this n-tuple $(i_1, i_2, \ldots, i_n)$ can be found by writing $\mathfrak{m}$ as a product of x_j's, and sorting those x_j's in weakly increasing order.

obtain

$$\begin{aligned}
&\sum_{\substack{\mathfrak{m}\in\operatorname{Mon};\\ \mathfrak{m}\text{ is squarefree}}} \mathfrak{m}t^{\deg\mathfrak{m}} \\
&= \underbrace{\sum_{\substack{\alpha\text{ is a weak}\\ \text{composition;}\\ \text{every entry of }\alpha\\ \text{belongs to }\{0,1\}}}}_{\substack{=\sum_{\substack{\alpha\in\{0,1\}^\infty\text{ is a weak}\\ \text{composition}}}\\ \text{(because if }\alpha\text{ is a weak}\\ \text{composition, then the statement}\\ \text{(every entry of }\alpha\text{ belongs to }\{0,1\})\\ \text{is equivalent to }(\alpha\in\{0,1\}^\infty))}} \mathbf{x}^\alpha t^{\deg(\mathbf{x}^\alpha)} = \sum_{\substack{\alpha\in\{0,1\}^\infty\text{ is a weak}\\ \text{composition}}} \mathbf{x}^\alpha t^{\deg(\mathbf{x}^\alpha)} \\
&= \sum_{\substack{(\alpha_1,\alpha_2,\alpha_3,\ldots)\in\{0,1\}^\infty\text{ is a}\\ \text{weak composition}}} \underbrace{\mathbf{x}^{(\alpha_1,\alpha_2,\alpha_3,\ldots)}}_{=\prod_{i\geq 1}x_i^{\alpha_i}} \underbrace{t^{\deg\left(\mathbf{x}^{(\alpha_1,\alpha_2,\alpha_3,\ldots)}\right)}}_{\substack{=t^{\alpha_1+\alpha_2+\alpha_3+\cdots}\\ \text{(since }\deg\left(\mathbf{x}^{(\alpha_1,\alpha_2,\alpha_3,\ldots)}\right)=\alpha_1+\alpha_2+\alpha_3+\cdots)}} \\
&\qquad \left(\text{here, we renamed the summation index }\alpha\text{ as }(\alpha_1,\alpha_2,\alpha_3,\ldots)\right) \\
&= \sum_{\substack{(\alpha_1,\alpha_2,\alpha_3,\ldots)\in\{0,1\}^\infty\text{ is a}\\ \text{weak composition}}} \left(\prod_{i\geq 1}x_i^{\alpha_i}\right) \underbrace{t^{\alpha_1+\alpha_2+\alpha_3+\cdots}}_{=t^{\sum_{i\geq 1}\alpha_i}=\prod_{i\geq 1}t^{\alpha_i}} = \sum_{\substack{(\alpha_1,\alpha_2,\alpha_3,\ldots)\in\{0,1\}^\infty\text{ is a}\\ \text{weak composition}}} \underbrace{\left(\prod_{i\geq 1}x_i^{\alpha_i}\right)\left(\prod_{i\geq 1}t^{\alpha_i}\right)}_{=\prod_{i\geq 1}\left(x_i^{\alpha_i}t^{\alpha_i}\right)} \\
&= \sum_{\substack{(\alpha_1,\alpha_2,\alpha_3,\ldots)\in\{0,1\}^\infty\text{ is a}\\ \text{weak composition}}} \prod_{i\geq 1} \underbrace{\left(x_i^{\alpha_i}t^{\alpha_i}\right)}_{=(x_it)^{\alpha_i}} = \sum_{\substack{(\alpha_1,\alpha_2,\alpha_3,\ldots)\in\{0,1\}^\infty\text{ is a}\\ \text{weak composition}}} \prod_{i\geq 1}(x_it)^{\alpha_i} \\
&= \sum_{\substack{(k_1,k_2,k_3,\ldots)\in\{0,1\}^\infty\text{ is a}\\ \text{weak composition}}} \prod_{i\geq 1}(x_it)^{k_i}
\end{aligned}$$

(here, we renamed the summation index $(\alpha_1,\alpha_2,\alpha_3,\ldots)$ as $(k_1,k_2,k_3,\ldots)$). Compared with

$$\underbrace{\prod_{i=1}^{\infty}}_{=\prod_{i\geq 1}} \underbrace{(1+x_it)}_{\substack{=\sum_{k\in\{0,1\}}(x_it)^k\\ \text{(since }\sum_{k\in\{0,1\}}(x_it)^k\\ =(x_it)^0+(x_it)^1=1+x_it\\ \text{(since }(x_it)^0=1\text{ and }(x_it)^1=x_it))}} = \prod_{i\geq 1}\sum_{k\in\{0,1\}}(x_it)^k = \sum_{\substack{(k_1,k_2,k_3,\ldots)\in\{0,1\}^\infty\text{ is a}\\ \text{weak composition}}} \prod_{i\geq 1}(x_it)^{k_i}$$
$$(\text{by the product rule}),$$

this yields

$$\sum_{\substack{\mathfrak{m}\in\operatorname{Mon};\\ \mathfrak{m}\text{ is squarefree}}} \mathfrak{m}t^{\deg\mathfrak{m}} = \prod_{i=1}^{\infty}(1+x_it). \tag{13.48.2}$$

On the other hand, every squarefree monomial $\mathfrak{m}\in\operatorname{Mon}$ can be uniquely represented in the form $x_{i_1}x_{i_2}\cdots x_{i_n}$ for some $n\in\mathbb{N}$ and some strictly increasing n-tuple $(i_1,i_2,\ldots,i_n)$ of positive integers.[572]

[572] *Proof.* This follows from the following observations:

- If $\mathfrak{m}\in\operatorname{Mon}$ is a squarefree monomial, then $\mathfrak{m}$ can be represented in the form $x_{i_1}x_{i_2}\cdots x_{i_n}$ for some $n\in\mathbb{N}$ and some strictly increasing n-tuple $(i_1,i_2,\ldots,i_n)$ of positive integers. (Indeed, we know already that every pure monomial $\mathfrak{m}\in\operatorname{Mon}$ can be uniquely represented in the form $x_{i_1}x_{i_2}\cdots x_{i_n}$ for some $n\in\mathbb{N}$ and some weakly increasing n-tuple $(i_1,i_2,\ldots,i_n)$ of positive integers. If $\mathfrak{m}$ is squarefree, then this n-tuple $(i_1,i_2,\ldots,i_n)$ must consist of n **distinct** integers (because otherwise, the product $x_{i_1}x_{i_2}\cdots x_{i_n}$ would fail to be squarefree, which would contradict the fact that $x_{i_1}x_{i_2}\cdots x_{i_n}=\mathfrak{m}$ is squarefree), and therefore is strictly increasing (because it is weakly increasing). Hence, every squarefree monomial $\mathfrak{m}$ can be be represented in the form $x_{i_1}x_{i_2}\cdots x_{i_n}$ for some $n\in\mathbb{N}$ and some strictly increasing n-tuple $(i_1,i_2,\ldots,i_n)$ of positive integers.)

Hence, we can substitute $x_{i_1}x_{i_2}\cdots x_{i_n}$ for $\mathfrak{m}$ in the sum $\sum\limits_{\substack{\mathfrak{m}\in\operatorname{Mon};\\ \mathfrak{m}\text{ is squarefree}}} \mathfrak{m}t^{\deg\mathfrak{m}}$. We thus obtain

$$\begin{aligned}
\sum_{\substack{\mathfrak{m}\in\operatorname{Mon};\\ \mathfrak{m}\text{ is squarefree}}} \mathfrak{m}t^{\deg\mathfrak{m}} &= \sum_{n\in\mathbb{N}} \underbrace{\sum_{\substack{(i_1,i_2,\ldots,i_n)\text{ is a}\\ \text{strictly increasing}\\ n\text{-tuple of positive integers}}}}_{=\sum_{i_1<i_2<\cdots<i_n}} \left(x_{i_1}x_{i_2}\cdots x_{i_n}\right) \underbrace{t^{\deg\left(x_{i_1}x_{i_2}\cdots x_{i_n}\right)}}_{\substack{=t^n\\ (\text{since }\deg\left(x_{i_1}x_{i_2}\cdots x_{i_n}\right)=n)}} \\
&= \sum_{n\in\mathbb{N}} \sum_{i_1<i_2<\cdots<i_n} \left(x_{i_1}x_{i_2}\cdots x_{i_n}\right) t^n = \sum_{n\in\mathbb{N}} \underbrace{\left(\sum_{i_1<i_2<\cdots<i_n} x_{i_1}x_{i_2}\cdots x_{i_n}\right)}_{\substack{=e_n\\ (\text{since (2.2.2) yields}\\ e_n=\sum_{i_1<i_2<\cdots<i_n} x_{i_1}x_{i_2}\cdots x_{i_n})}} t^n \\
&= \sum_{n\in\mathbb{N}} \underbrace{e_n}_{=e_n(\mathbf{x})} t^n = \sum_{n\geq 0} e_n\left(\mathbf{x}\right) t^n = \underbrace{e_0\left(\mathbf{x}\right)}_{=e_0=1} + e_1\left(\mathbf{x}\right) t + e_2\left(\mathbf{x}\right) t^2 + \cdots \\
&= 1 + e_1\left(\mathbf{x}\right) t + e_2\left(\mathbf{x}\right) t^2 + \cdots .
\end{aligned}$$

Compared with (13.48.2), this yields

$$\prod_{i=1}^{\infty}\left(1+x_it\right) = 1 + e_1\left(\mathbf{x}\right) t + e_2\left(\mathbf{x}\right) t^2 + \cdots = \sum_{n\geq 0} e_n\left(\mathbf{x}\right) t^n.$$

This completes the solution to Exercise 2.2.15.

13.49. **Solution to Exercise 2.3.4.** *Solution to Exercise 2.3.4.* (a) This is straightforward. If λ and μ are two partitions such that $\mu\subseteq\lambda$, and if $\mathcal{L}$ is any total order on the positive integers, then we say that an assignment T of entries in $\{1,2,3,\ldots\}$ to the cells of the Ferrers diagram of λ/μ is an *$\mathcal{L}$-column-strict tableau* if it is weakly $\mathcal{L}$-increasing left-to-right in rows, and strictly $\mathcal{L}$-increasing top-to-bottom in columns. (This definition of $\mathcal{L}$-column-strict tableaux clearly extends the definition given in Remark 2.2.5 for tableaux of shape λ.) Now, the analogue of Proposition 2.2.6 is the following statement:

Proposition 13.49.1. *Let λ and μ be two partitions such that $\mu\subseteq\lambda$. Then, for any total order $\mathcal{L}$ on the positive integers,*

$$s_{\lambda/\mu} = \sum_T \mathbf{x}^{\operatorname{cont}(T)}$$

as T runs through all $\mathcal{L}$-column-strict tableaux of shape λ/μ.

The proof of Proposition 13.49.1 is completely analogous to the proof of Proposition 2.2.6.

(b) If α and β are two partitions such that $\beta\subseteq\alpha$, then let $Y\left(\alpha/\beta\right)$ denote the skew Ferrers diagram α/β. This is a finite set of cells in $\{1,2,3,...\}^2$.

We know that the skew Ferrers diagram λ'/μ' can be obtained from the skew Ferrers diagram λ/μ by a 180° rotation. In other words, there exists a 180° rotation r such that $r\left(Y\left(\lambda/\mu\right)\right) = Y\left(\lambda'/\mu'\right)$. Consider this r.

Let $\mathcal{L}$ be the total order on the set of all positive integers which is defined by $\cdots <_{\mathcal{L}} 3 <_{\mathcal{L}} 2 <_{\mathcal{L}} 1$ (in other words, let $\mathcal{L}$ be the reverse of the usual total order on the set of all positive integers). Recall the definition

- If $\mathfrak{m}\in\operatorname{Mon}$ is a squarefree monomial, then the representation of $\mathfrak{m}$ in the form $x_{i_1}x_{i_2}\cdots x_{i_n}$ for some $n\in\mathbb{N}$ and some strictly increasing n-tuple $\left(i_1,i_2,\ldots,i_n\right)$ of positive integers is unique. (Indeed, even if we only require $\left(i_1,i_2,\ldots,i_n\right)$ to be weakly increasing rather than strictly increasing, then the representation is unique (because we know already that every pure monomial $\mathfrak{m}\in\operatorname{Mon}$ can be uniquely represented in the form $x_{i_1}x_{i_2}\cdots x_{i_n}$ for some $n\in\mathbb{N}$ and some weakly increasing n-tuple $\left(i_1,i_2,\ldots,i_n\right)$ of positive integers).)
- Every product of the form $x_{i_1}x_{i_2}\cdots x_{i_n}$ for some $n\in\mathbb{N}$ and some strictly increasing n-tuple $\left(i_1,i_2,\ldots,i_n\right)$ of positive integers is a squarefree monomial. (This is obvious, because the elements $i_1,i_2,\ldots,i_n$ of a strictly increasing n-tuple $\left(i_1,i_2,\ldots,i_n\right)$ are distinct.)

of $\mathcal{L}$-column-strict tableaux that we gave in the solution of part (a) of this exercise. According to this definition, an $\mathcal{L}$-column-strict tableau of shape λ'/μ' is an assignment of entries in $\{1,2,3,\ldots\}$ to the cells of the Ferrers diagram of λ'/μ' which is weakly $\mathcal{L}$-increasing left-to-right in rows, and strictly $\mathcal{L}$-increasing top-to-bottom in columns. Since "weakly $\mathcal{L}$-increasing" is the same as "weakly decreasing" (because $\mathcal{L}$ is the reverse of the usual total order on the set of all positive integers), and since "strictly $\mathcal{L}$-increasing" is the same as "strictly decreasing" (for the same reason), this rewrites as follows: An $\mathcal{L}$-column-strict tableau of shape λ'/μ' is an assignment of entries in $\{1,2,3,\ldots\}$ to the cells of the Ferrers diagram of λ'/μ' which is weakly decreasing left-to-right in rows, and strictly decreasing top-to-bottom in columns. Hence, if T is an $\mathcal{L}$-column-strict tableau of shape λ'/μ', then $T\circ r$ (this composition is well-defined[573]) is an assignment of entries in $\{1,2,3,\ldots\}$ to the cells of the Ferrers diagram of λ/μ which is weakly decreasing right-to-left in rows, and strictly decreasing bottom-to-top in columns[574]. In other words, if T is an $\mathcal{L}$-column-strict tableau of shape λ'/μ', then $T\circ r$ is an assignment of entries in $\{1,2,3,\ldots\}$ to the cells of the Ferrers diagram of λ/μ which is weakly increasing left-to-right in rows, and strictly increasing top-to-bottom in columns. In other words, if T is an $\mathcal{L}$-column-strict tableau of shape λ'/μ', then $T\circ r$ is a column-strict tableau (in the usual sense) of shape λ/μ. Hence, we have constructed a map

$$\{\mathcal{L}\text{-column-strict tableaux of shape } \lambda'/\mu'\} \to \{\text{column-strict tableaux of shape } \lambda/\mu\},$$

which sends every T to $T\circ r$. This map is easily seen to be a bijection. Therefore, we can substitute $T\circ r$ for T in the sum $\sum_{\substack{T \text{ is a column-strict} \\ \text{tableau of shape } \lambda/\mu}} \mathbf{x}^{\operatorname{cont}(T)}$. We thus obtain

$$\sum_{\substack{T \text{ is a column-strict} \\ \text{tableau of shape } \lambda/\mu}} \mathbf{x}^{\operatorname{cont}(T)} = \sum_{\substack{T \text{ is an } \mathcal{L}\text{-column-strict} \\ \text{tableau of shape } \lambda'/\mu'}} \underbrace{\mathbf{x}^{\operatorname{cont}(T\circ r)}}_{\substack{=\mathbf{x}^{\operatorname{cont}(T)} \\ \text{(since the multiset of entries} \\ \text{of } T\circ r \text{ is the multiset} \\ \text{of entries of } T)}}$$

$$= \sum_{\substack{T \text{ is an } \mathcal{L}\text{-column-strict} \\ \text{tableau of shape } \lambda'/\mu'}} \mathbf{x}^{\operatorname{cont}(T)}. \qquad (13.49.1)$$

But Proposition 13.49.1 (applied to λ' and μ' instead of λ and μ) yields

$$s_{\lambda'/\mu'} = \sum_T \mathbf{x}^{\operatorname{cont}(T)}$$

as T runs through all $\mathcal{L}$-column-strict tableaux of shape λ'/μ'. In other words,

$$s_{\lambda'/\mu'} = \sum_{\substack{T \text{ is an } \mathcal{L}\text{-column-strict} \\ \text{tableau of shape } \lambda'/\mu'}} \mathbf{x}^{\operatorname{cont}(T)} = \sum_{\substack{T \text{ is a column-strict} \\ \text{tableau of shape } \lambda/\mu}} \mathbf{x}^{\operatorname{cont}(T)} \qquad (13.49.2)$$

(by (13.49.1)).

But the definition of $s_{\lambda/\mu}$ yields $s_{\lambda/\mu} = \sum_T \mathbf{x}^{\operatorname{cont}(T)}$, where the sum ranges over all column-strict tableaux T of shape λ/μ. In other words,

$$s_{\lambda/\mu} = \sum_{\substack{T \text{ is a column-strict} \\ \text{tableau of shape } \lambda/\mu}} \mathbf{x}^{\operatorname{cont}(T)}.$$

Compared with (13.49.2), this yields $s_{\lambda/\mu} = s_{\lambda'/\mu'}$. This solves part (b) of the exercise.

---

[573] because T is an assignment of entries in $\{1,2,3,\ldots\}$ to the cells of the Ferrers diagram of λ'/μ', that is, a map $Y(\lambda'/\mu') \to \{1,2,3,\ldots\}$, whereas r is a map sending $Y(\lambda/\mu)$ to $Y(\lambda'/\mu')$

[574] because r is a 180° rotation, and thus interchanges "left-to-right" with "right-to-left" and interchanges "top-to-bottom" with "bottom-to-top"

13.50. **Solution to Exercise 2.3.5.** *Solution to Exercise 2.3.5.* If φ and ψ are two partitions such that $\psi \subseteq \varphi$, then let $Y(\varphi/\psi)$ denote the skew Ferrers diagram φ/ψ (this is a subset of $\{1,2,3,...\}^2$).

Whenever Z is a subset of $\mathbb{Z}^2$, we define a *column-strict Z-tableau* to be an assignment of entries in $\{1,2,3,...\}$ to the elements of Z which is weakly increasing left-to-right in rows and strictly increasing top-to-bottom in columns. It is clear that if φ and ψ are two partitions such that $\psi \subseteq \varphi$, then a column-strict tableau of shape φ/ψ is the same as a column-strict $Y(\varphi/\psi)$-tableau. We define the notation $\operatorname{cont}(T)$ (and therefore, $\mathbf{x}^{\operatorname{cont}(T)}$) for a column-strict Z-tableau T in the same way as it is defined for a column-strict tableau of shape λ/μ (for some partitions λ and μ).

The following is now more or less obvious:

Lemma 13.50.1. *Let φ and ψ be two partitions such that $\psi \subseteq \varphi$. Let Z be a subset of $\mathbb{Z}^2$. Assume that the skew Ferrers diagram φ/ψ can be obtained from Z by parallel translation. Then,*

$$s_{\varphi/\psi} = \sum_{\substack{T \text{ is a column-strict} \\ Z\text{-tableau}}} \mathbf{x}^{\operatorname{cont}(T)}.$$

Proof of Lemma 13.50.1. Let R be the parallel translation which sends the set Z to $Y(\varphi/\psi)$.

The definition of $s_{\varphi/\psi}$ yields $s_{\varphi/\psi} = \sum_{\substack{T \text{ is a column-strict} \\ \text{tableau of shape } \varphi/\psi}} \mathbf{x}^{\operatorname{cont}(T)}$. It remains to prove that the right hand side of this equality equals $\sum_{\substack{T \text{ is a column-strict} \\ Z\text{-tableau}}} \mathbf{x}^{\operatorname{cont}(T)}$. To achieve this, it is clearly enough to find a bijection

$$\Gamma : (\text{the set of all column-strict tableaux of shape } \varphi/\psi) \to (\text{the set of all column-strict } Z\text{-tableaux})$$

which satisfies

$$\left(\mathbf{x}^{\operatorname{cont}(\Gamma(T))} = \mathbf{x}^{\operatorname{cont}(T)} \qquad \text{for every column-strict tableau } T \text{ of shape } \varphi/\psi\right).$$

But this is very easy: The bijection Γ sends every column-strict tableau T of shape φ/ψ to the column-strict Z-tableau $T \circ R$. (The notation $T \circ R$ makes sense because T, being a column-strict tableau of shape φ/ψ, is an assignment of entries in $\{1,2,3,...\}$ to the cells of the skew Ferrers diagram φ/ψ, that is, a map $Y(\varphi/\psi) \to \{1,2,3,...\}$. Visually speaking, $T \circ R$ is the result of moving the tableau T so that it takes up the cells of Z rather than the cells of φ/ψ.) Lemma 13.50.1 is proven. □

Now, let us return to the solution of the exercise. Clearly, the subsets $F_{\text{rows}\leq k}$ and $F_{\text{rows}>k}$ of F are disjoint, and their union is F.

Lemma 13.50.1 (applied to α, β and $F_{\text{rows}\leq k}$ instead of φ, ψ and Z) yields

$$(13.50.1) \qquad s_{\alpha/\beta} = \sum_{\substack{T \text{ is a column-strict} \\ F_{\text{rows}\leq k}\text{-tableau}}} \mathbf{x}^{\operatorname{cont}(T)} = \sum_{\substack{P \text{ is a column-strict} \\ F_{\text{rows}\leq k}\text{-tableau}}} \mathbf{x}^{\operatorname{cont}(P)}$$

(here, we renamed the summation index T as P). Also, Lemma 13.50.1 (applied to γ, δ and $F_{\text{rows}>k}$ instead of φ, ψ and Z) yields

$$(13.50.2) \qquad s_{\gamma/\delta} = \sum_{\substack{T \text{ is a column-strict} \\ F_{\text{rows}>k}\text{-tableau}}} \mathbf{x}^{\operatorname{cont}(T)} = \sum_{\substack{Q \text{ is a column-strict} \\ F_{\text{rows}>k}\text{-tableau}}} \mathbf{x}^{\operatorname{cont}(Q)}$$

(here, we renamed the summation index T as Q). Multiplying the identities (13.50.1) and (13.50.2), we obtain

$$\begin{aligned} s_{\alpha/\beta}s_{\gamma/\delta} &= \left(\sum_{\substack{P \text{ is a column-strict} \\ F_{\text{rows}\leq k}\text{-tableau}}} \mathbf{x}^{\operatorname{cont}(P)} \right) \left(\sum_{\substack{Q \text{ is a column-strict} \\ F_{\text{rows}>k}\text{-tableau}}} \mathbf{x}^{\operatorname{cont}(Q)} \right) \\ &= \sum_{\substack{P \text{ is a column-strict} \\ F_{\text{rows}\leq k}\text{-tableau}}} \sum_{\substack{Q \text{ is a column-strict} \\ F_{\text{rows}>k}\text{-tableau}}} \mathbf{x}^{\operatorname{cont}(P)}\mathbf{x}^{\operatorname{cont}(Q)} \\ &= \sum_{\substack{(P,Q)\in \\ \left(\text{the set of all column-strict } F_{\text{rows}\leq k}\text{-tableaux}\right) \\ \times\left(\text{the set of all column-strict } F_{\text{rows}>k}\text{-tableaux}\right)}} \mathbf{x}^{\operatorname{cont}(P)}\mathbf{x}^{\operatorname{cont}(Q)}. \end{aligned} \tag{13.50.3}$$

On the other hand, the definition of $s_{\lambda/\mu}$ yields

$$s_{\lambda/\mu} = \sum_{\substack{T \text{ is a column-strict} \\ \text{tableau of shape } \lambda/\mu}} \mathbf{x}^{\operatorname{cont}(T)}. \tag{13.50.4}$$

Our goal is to prove that the left-hand side of (13.50.4) equals the left-hand side of (13.50.3). For this, it is clearly enough to show that the right-hand side of (13.50.4) equals the right-hand side of (13.50.3). But to achieve this, it clearly suffices to exhibit a bijection

$$\begin{aligned} \Phi &: (\text{the set of all column-strict tableaux of shape } \lambda/\mu) \\ &\to (\text{the set of all column-strict } F_{\text{rows}\leq k}\text{-tableaux}) \times (\text{the set of all column-strict } F_{\text{rows}>k}\text{-tableaux}) \end{aligned}$$

which has the property that
(13.50.5)

$$\text{if } (P,Q) = \Phi(T) \text{ for some column-strict tableau } T \text{ of shape } \lambda/\mu, \text{ then } \mathbf{x}^{\operatorname{cont}(P)}\mathbf{x}^{\operatorname{cont}(Q)} = \mathbf{x}^{\operatorname{cont}(T)}.$$

We claim that such a bijection Φ can be defined by

$$\left(\Phi(T) = \left(T\mid_{F_{\text{rows}\leq k}}, T\mid_{F_{\text{rows}>k}}\right) \qquad \text{for every column-strict tableau } T \text{ of shape } \lambda/\mu\right). \tag{13.50.6}$$

Indeed, it is clear that we can define a map

$$\begin{aligned} \Phi &: (\text{the set of all column-strict tableaux of shape } \lambda/\mu) \\ &\to (\text{the set of all column-strict } F_{\text{rows}\leq k}\text{-tableaux}) \times (\text{the set of all column-strict } F_{\text{rows}>k}\text{-tableaux}) \end{aligned}$$

by (13.50.6), and that this map Φ satisfies (13.50.5). All that remains to be proven is that this map Φ is a bijection. It is clear that Φ is injective, so we only need to prove that Φ is surjective.

Let $(P,Q) \in$ (the set of all column-strict $F_{\text{rows}\leq k}$-tableaux)×(the set of all column-strict $F_{\text{rows}>k}$-tableaux) be arbitrary. We are going to prove that (P,Q) lies in the image of Φ.

Define a map $T : F \to \{1,2,3,...\}$ by setting

$$\left(T(p) = \begin{cases} P(p), & \text{if } p \in F_{\text{rows}\leq k}; \\ Q(p), & \text{if } p \in F_{\text{rows}>k} \end{cases} \qquad \text{for all } p \in F\right).$$

This map T is clearly well-defined (since the subsets $F_{\text{rows}\leq k}$ and $F_{\text{rows}>k}$ of F are disjoint, and their union is F), and thus is an assignment of entries in $\{1,2,3,...\}$ to the cells of the skew Ferrers diagram λ/μ (since F is the set of those cells). It furthermore satisfies $T\mid_{F_{\text{rows}\leq k}}= P$ and $T\mid_{F_{\text{rows}>k}}= Q$. We shall now show that this assignment T is a column-strict tableau of shape λ/μ.

This rests on the following observation:

> *Assertion A:* Let c and d be two cells lying in F. Assume that the cells c and d either lie in one and the same row, or lie in one and the same column. Then, either both c and d belong to $F_{\text{rows}\leq k}$, or both c and d belong to $F_{\text{rows}>k}$.

Assertion A is an easy consequence of our assumption that $\mu_k \geq \lambda_{k+1}$. [575] Now, we want to prove that T is a column-strict tableau of shape λ/μ. To do so, we need to check that T is weakly increasing left-to-right in rows, and strictly increasing top-to-bottom in columns. We will only prove the latter part of this statement, as the former part is proven analogously. So we are going to show that T is strictly increasing top-to-bottom in columns. In other words, we are going to show that if c and d are two cells of λ/μ lying in one and the same column, with d lying strictly further south than c, then $T(c) < T(d)$. Indeed, consider two such cells c and d. Assertion A shows that either both c and d belong to $F_{\text{rows}\leq k}$, or both c and d belong to $F_{\text{rows}>k}$. Let us WLOG assume that we are in the first of these two cases (the other case is exactly analogous). Then, both c and d belong to $F_{\text{rows}\leq k}$, so that we have $T(c) = \left(T\mid_{F_{\text{rows}\leq k}}\right)(c)$ and $T(d) = \left(T\mid_{F_{\text{rows}\leq k}}\right)(d)$. Since $T\mid_{F_{\text{rows}\leq k}} = P$, these two equalities rewrite as $T(c) = P(c)$ and $T(d) = P(d)$. But since P is strictly increasing top-to-bottom in columns (because P is a column-strict tableau), we have $P(c) < P(d)$, and thus $T(c) = P(c) < P(d) = T(d)$. Thus, we have proven that $T(c) < T(d)$. This completes the proof that T is a column-strict tableau of shape λ/μ. Hence, $\Phi(T)$ is well-defined, and the definition of $\Phi(T)$ shows that $\Phi(T) = \left(\underbrace{T\mid_{F_{\text{rows}\leq k}}}_{=P}, \underbrace{T\mid_{F_{\text{rows}>k}}}_{=Q}\right) = (P,Q)$. Thus, (P,Q) lies in the image of Φ.

Now, let us forget that we fixed (P,Q). We thus have shown that every $(P,Q) \in (\text{the set of all column-strict } F_{\text{rows}\leq k}\text{-tableaux}) \times (\text{the set of all column-strict } F_{\text{rows}>k}\text{-tableaux})$ lies in the image of Φ. In other words, the map Φ is surjective, which (as we know that Φ is injective) yields that Φ is a bijection. As explained above, this completes the solution of Exercise 2.3.5.

13.51. **Solution to Exercise 2.3.7.** *Solution to Exercise 2.3.7.* As usual, let T denote the twist map $\Lambda \otimes \Lambda \to \Lambda \otimes \Lambda$ (that is, the $\mathbf{k}$-linear map sending every $c \otimes d \in \Lambda \otimes \Lambda$ to $d \otimes c$). By the definition of "cocommutative", we know that the Hopf algebra Λ is cocommutative if and only if the diagram

(13.51.1)
$$\begin{tikzcd}
\Lambda \otimes \Lambda \arrow[rr, "T"] & & \Lambda \otimes \Lambda \\
& \Lambda \arrow[ul, "\Delta"] \arrow[ur, "\Delta"'] &
\end{tikzcd}$$

commutes. Hence, in order to solve Exercise 2.3.7(a), it is enough to check that the diagram (13.51.1) commutes.

The set $\{h_n\}_{n=1,2,\ldots}$ generates the $\mathbf{k}$-algebra Λ (due to Proposition 2.4.1). In other words, the set $\{h_1, h_2, h_3, \ldots\}$ is a generating set of the $\mathbf{k}$-algebra Λ.

By the axioms of a bialgebra, the comultiplication Δ of Λ is a $\mathbf{k}$-algebra homomorphism (since Λ is a bialgebra). Hence, $T \circ \Delta$ also is a $\mathbf{k}$-algebra homomorphism (since T and Δ are $\mathbf{k}$-algebra homomorphisms).

[575] *Proof of Assertion A:* Assume the contrary. Then, one of the cells c and d belongs to $F_{\text{rows}\leq k}$, whereas the other belongs to $F_{\text{rows}>k}$. We WLOG assume that c belongs to $F_{\text{rows}\leq k}$, whereas d belongs to $F_{\text{rows}>k}$ (since otherwise, we can simply switch c with d).

Write the cell c in the form (a,b), so that c lies in row a and column b. Write the cell d in the form (a',b'), so that d lies in row a' and column b'.

We have $(a,b) = c \in F_{\text{rows}\leq k}$, so that $a \leq k$ (by the definition of $F_{\text{rows}\leq k}$). We have $(a',b') = d \in F_{\text{rows}>k}$, so that $a' > k$ (by the definition of $F_{\text{rows}>k}$).

We have $a \leq k < a'$ (since $a' > k$), hence $a \neq a'$. Thus, the cells c and d lie in different rows (since the cell c lies in row a, whereas the cell d lies in row a'). As a consequence, the cells c and d must lie in one and the same column (since we assumed that the cells c and d either lie in one and the same row, or lie in one and the same column). In other words, $b = b'$ (since the cell c lies in column b, while the cell d lies in column b').

Now, $(a,b) = c \in F = Y(\lambda/\mu)$. By the definition of $Y(\lambda/\mu)$, this shows that $\mu_a < b \leq \lambda_a$. Similarly, $\mu_{a'} < b' \leq \lambda_{a'}$.

We have $a' > k$, hence $a' \geq k+1$ (since a' and k are integers), thus $k+1 \leq a'$.

Since μ is a partition, we have $\mu_1 \geq \mu_2 \geq \mu_3 \geq \ldots$. Hence, $\mu_a \geq \mu_k$ (since $a \leq k$). Thus, $\mu_a \geq \mu_k \geq \lambda_{k+1}$.

Since λ is a partition, we have $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \ldots$. Thus, $\lambda_{k+1} \geq \lambda_{a'}$ (since $k+1 \leq a'$). Thus, $\mu_a \geq \lambda_{k+1} \geq \lambda_{a'}$, so that $\lambda_{a'} \leq \mu_a$ and thus $b' \leq \lambda_{a'} \leq \mu_a < b = b'$, which is absurd. This contradiction proves that our assumption was wrong, and thus Assertion A is proven.

Every positive integer n satisfies

$$\begin{aligned}
(T \circ \Delta)(h_n) &= T\left(\underbrace{\Delta h_n}_{\substack{=\sum_{i+j=n} h_i \otimes h_j \\ \text{(by Proposition 2.3.6(iii))}}}\right) \\
&\qquad \left(\begin{array}{c}\text{here, we are using the notation } \sum_{i+j=n} h_i \otimes h_j \\ \text{in the same way as explained in Proposition 2.3.6}\end{array}\right) \\
&= T\left(\sum_{i+j=n} h_i \otimes h_j\right) = \sum_{i+j=n} h_j \otimes h_i \qquad \text{(by the definition of the twist map } T\text{)} \\
&= \underbrace{\sum_{j+i=n}}_{=\sum_{i+j=n}} h_i \otimes h_j \qquad \text{(here, we renamed the summation index } (i,j) \text{ as } (j,i)\text{)} \\
&= \sum_{i+j=n} h_i \otimes h_j = \Delta h_n \qquad \text{(by Proposition 2.3.6(iii))} \\
&= \Delta(h_n).
\end{aligned}$$

In other words, for every positive integer n, the two maps $T \circ \Delta$ and Δ are equal to each other on the element h_n. In other words, the two maps $T \circ \Delta$ and Δ are equal to each other on the set $\{h_1, h_2, h_3, ...\}$. Hence, the two maps $T \circ \Delta$ and Δ are equal to each other on a generating set of the $\mathbf{k}$-algebra Λ (since the set $\{h_1, h_2, h_3, ...\}$ is a generating set of the $\mathbf{k}$-algebra Λ). Since these two maps $T \circ \Delta$ and Δ are $\mathbf{k}$-algebra homomorphisms, this shows that the two maps $T \circ \Delta$ and Δ must be identical (because if two $\mathbf{k}$-algebra homomorphisms with the same domain and the same target are equal to each other on a generating set of their domain, then these two homomorphisms must be identical). In other words, $T \circ \Delta = \Delta$. Hence, the diagram (13.51.1) commutes. As we know, this shows that the Hopf algebra Λ is cocommutative. This solves Exercise 2.3.7(a).

(b) Let λ and ν be two partitions. We have shown above that $T \circ \Delta = \Delta$. Hence, $\Delta = T \circ \Delta$. Applying both sides of this equality to $s_{\lambda/\nu}$, we obtain

$$\begin{aligned}
\Delta s_{\lambda/\nu} = (T \circ \Delta)(s_{\lambda/\nu}) &= T\left(\underbrace{\Delta s_{\lambda/\nu}}_{\substack{=\sum_{\substack{\mu \in \operatorname{Par}: \\ \nu \subseteq \mu \subseteq \lambda}} s_{\mu/\nu} \otimes s_{\lambda/\mu} \\ \text{(by Proposition 2.3.6(v))}}}\right) = T\left(\sum_{\substack{\mu \in \operatorname{Par}: \\ \nu \subseteq \mu \subseteq \lambda}} s_{\mu/\nu} \otimes s_{\lambda/\mu}\right) \\
&= \sum_{\substack{\mu \in \operatorname{Par}: \\ \nu \subseteq \mu \subseteq \lambda}} s_{\lambda/\mu} \otimes s_{\mu/\nu} \qquad \text{(by the definition of the twist map } T\text{)}.
\end{aligned}$$

This solves Exercise 2.3.7(b).

13.52. **Solution to Exercise 2.3.8.** *Solution to Exercise 2.3.8.* (a) Whenever $\alpha = (\alpha_1, \alpha_2, \alpha_3, ...)$ is a weak composition satisfying ($\alpha_i = 0$ for every $i > n$), the monomial $\mathbf{x}^\alpha$ is a monomial in $\mathbf{k}[x_1, x_2, ..., x_n]$. This will be often used in the following.

Recall that $s_{\lambda/\mu}$ is defined as $\sum_{\substack{T \text{ is a column-strict} \\ \text{tableau of shape } \lambda/\mu}} \mathbf{x}^{\operatorname{cont}(T)}$. Now, $s_{\lambda/\mu}(x_1, x_2, \ldots, x_n)$ is the result of substituting $x_1, x_2, ..., x_n, 0, 0, 0, ...$ for $x_1, x_2, x_3, ...$ in $s_{\lambda/\mu}$. This substitution has the following effect on any given monomial $\mathbf{x}^\alpha$:

- if none of the indeterminates $x_{n+1}, x_{n+2}, x_{n+3}, ...$ occur in this monomial $\mathbf{x}^{\alpha}$, then the monomial $\mathbf{x}^{\alpha}$ stays fixed;
- otherwise, the monomial $\mathbf{x}^{\alpha}$ goes to 0.

Hence, the effect of this substitution on the power series $s_{\lambda/\mu} = \sum_{\substack{T \text{ is a column-strict} \\ \text{tableau of shape } \lambda/\mu}} \mathbf{x}^{\operatorname{cont}(T)}$ is that:

- every addend in the sum $\sum_{\substack{T \text{ is a column-strict} \\ \text{tableau of shape } \lambda/\mu}} \mathbf{x}^{\operatorname{cont}(T)}$ for which none of the indeterminates $x_{n+1}, x_{n+2}, x_{n+3}, ...$ occur in the monomial $\mathbf{x}^{\operatorname{cont}(T)}$ stays fixed;
- all other addends go to 0.

The result of the substitution is therefore $\sum_{\substack{T \text{ is a column-strict} \\ \text{tableau of shape } \lambda/\mu; \\ \text{none of the indeterminates } x_{n+1}, x_{n+2}, x_{n+3}, ... \\ \text{occur in the monomial } \mathbf{x}^{\operatorname{cont}(T)}}} \mathbf{x}^{\operatorname{cont}(T)}$. We thus have

$$s_{\lambda/\mu}(x_1, x_2, \ldots, x_n) = \sum_{\substack{T \text{ is a column-strict} \\ \text{tableau of shape } \lambda/\mu; \\ \text{none of the indeterminates } x_{n+1}, x_{n+2}, x_{n+3}, ... \\ \text{occur in the monomial } \mathbf{x}^{\operatorname{cont}(T)}}} \mathbf{x}^{\operatorname{cont}(T)}.$$

But this rewrites as

$$s_{\lambda/\mu}(x_1, x_2, \ldots, x_n) = \sum_{\substack{T \text{ is a column-strict} \\ \text{tableau of shape } \lambda/\mu; \\ \text{all entries of } T \text{ belong} \\ \text{to } \{1,2,...,n\}}} \mathbf{x}^{\operatorname{cont}(T)}$$

(because for a column-strict tableau T, saying that none of the indeterminates $x_{n+1}, x_{n+2}, x_{n+3}, ...$ occur in the monomial $\mathbf{x}^{\operatorname{cont}(T)}$ is equivalent to saying that all entries of T belong to $\{1, 2, ..., n\}$). This solves Exercise 2.3.8(a).

(b) Let λ be a partition having more than n parts. We have to prove that $s_{\lambda}(x_1, x_2, \ldots, x_n) = 0$. Since

$$\underbrace{s_{\lambda}}_{= s_{\lambda/\varnothing}}(x_1, x_2, \ldots, x_n) = s_{\lambda/\varnothing}(x_1, x_2, \ldots, x_n) = \sum_{\substack{T \text{ is a column-strict} \\ \text{tableau of shape } \lambda/\varnothing; \\ \text{all entries of } T \text{ belong} \\ \text{to } \{1,2,...,n\}}} \mathbf{x}^{\operatorname{cont}(T)}$$

(by Exercise 2.3.8(a), applied to $\mu = \varnothing$), this goal will clearly be achieved if we can show that the sum $\sum_{\substack{T \text{ is a column-strict} \\ \text{tableau of shape } \lambda/\varnothing; \\ \text{all entries of } T \text{ belong} \\ \text{to } \{1,2,...,n\}}} \mathbf{x}^{\operatorname{cont}(T)}$ is empty, i.e., that there exists no column-strict tableau T of shape $\lambda/\varnothing$ such that all entries of T belong to $\{1, 2, ..., n\}$.

Assume the contrary. Thus, there exists a column-strict tableau T of shape $\lambda/\varnothing$ such that all entries of T belong to $\{1, 2, ..., n\}$. This tableau has more than n rows (since the partition λ has more than n parts), and thus the first column of this tableau must have more than n entries. These entries must be strictly increasing top-to-bottom (since the entries of a column-strict tableau strictly increase top-to-bottom along columns) and hence be distinct, but at the same time (like all entries of T) they must belong to $\{1, 2, ..., n\}$. So we have found more than n entries which are distinct and belong to $\{1, 2, ..., n\}$. This contradicts the fact that the set $\{1, 2, ..., n\}$ does not have more than n distinct elements. This contradiction concludes our proof, and Exercise 2.3.8(b) is solved.

13.53. **Solution to Exercise 2.4.4.** *Solution to Exercise 2.4.4.*

Proof of Proposition 2.4.3. The map ω is precisely the $\mathbf{k}$-algebra homomorphism ω defined in the proof of Proposition 2.4.1. Thus, in particular, ω is a $\mathbf{k}$-algebra homomorphism; hence, $\omega(1) = 1$ and $\omega(0) = 0$.

(a) Let $n \in \mathbb{Z}$. We must prove that $\omega(e_n) = h_n$.

If n is a positive integer, then this follows immediately from the definition of ω. Thus, for the rest of the proof of $\omega\left(e_n\right)=h_n$, we can WLOG assume that n is not a positive integer. Assume this.

From $e_0=1$, we obtain $\omega\left(e_0\right)=\omega\left(1\right)=1=h_0$ (since $h_0=1$). Hence, $\omega\left(e_n\right)=h_n$ holds for $n=0$. Thus, for the rest of the proof of $\omega\left(e_n\right)=h_n$, we can WLOG assume that we don't have $n=0$. Assume this.

Now, n is a negative integer (since n is not a positive integer, and since we don't have $n=0$). This yields $h_n=0$ and $e_n=0$. Now, $\omega\left(\underbrace{e_n}_{=0}\right)=\omega\left(0\right)=0=h_n$. This proves $\omega\left(e_n\right)=h_n$. This proves Proposition 2.4.3(a).

(b) The proof of Proposition 2.4.3(b) is completely analogous to the proof of Proposition 2.4.3(a) given above, except that e_n and h_n trade places (and that we need to use (2.4.9) instead of the definition of ω).

(d) In the proof of Proposition 2.4.1, we have shown that ω is an involution and a $\mathbf{k}$-algebra automorphism of Λ. This proves Proposition 2.4.3(d).

(e) Proposition 1.4.10 (applied to $A=\Lambda$) shows that the antipode S of Λ is a $\mathbf{k}$-algebra anti-endomorphism. In other words, S is a $\mathbf{k}$-algebra anti-homomorphism from Λ to Λ. But Exercise 1.5.8(a) (applied to $A=\Lambda$ and $B=\Lambda$) shows that the $\mathbf{k}$-algebra anti-homomorphisms from Λ to Λ are the same as the $\mathbf{k}$-algebra homomorphisms from Λ to Λ (since the $\mathbf{k}$-algebra Λ is commutative). Hence, S is a $\mathbf{k}$-algebra homomorphism from Λ to Λ (since S is a $\mathbf{k}$-algebra anti-homomorphism from Λ to Λ).

Recall that Λ is a graded $\mathbf{k}$-bialgebra. Let $q=-1\in\mathbf{k}$. Consider the $\mathbf{k}$-linear map $D_q:\Lambda\to\Lambda$ constructed in Exercise 1.3.24 (applied to $A=\Lambda$). From Exercise 1.3.24 (applied to $A=\Lambda$), we know that this map D_q is a $\mathbf{k}$-bialgebra homomorphism; thus, in particular, D_q is a $\mathbf{k}$-algebra homomorphism. The definition of D_q shows that

$$D_q\left(a\right)=q^n a \qquad \text{for each } n\in\mathbb{N} \text{ and each } a\in\Lambda_n. \tag{13.53.1}$$

Now, the map $\omega\circ D_q:\Lambda\to\Lambda$ is a $\mathbf{k}$-algebra homomorphism (since it is the composition of the two $\mathbf{k}$-algebra homomorphisms ω and D_q).

Let n be a positive integer. Then, $\omega\left(e_n\right)=h_n$ (by the definition of ω). But $e_n\in\Lambda_n$; hence, (13.53.1) (applied to $a=e_n$) yields $D_q\left(e_n\right)=q^n e_n=\left(-1\right)^n e_n$ (since $q=-1$). Now,

$$\begin{aligned}\left(\omega\circ D_q\right)\left(e_n\right)&=\omega\left(\underbrace{D_q\left(e_n\right)}_{=\left(-1\right)^n e_n}\right)=\omega\left(\left(-1\right)^n e_n\right)=\left(-1\right)^n\underbrace{\omega\left(e_n\right)}_{=h_n} \qquad \text{(since the map } \omega \text{ is } \mathbf{k}\text{-linear)}\\&=\left(-1\right)^n h_n.\end{aligned}$$

Comparing this with

$$S\left(e_n\right)=\left(-1\right)^n h_n \qquad \text{(by Proposition 2.4.1(ii))},$$

we obtain $\left(\omega\circ D_q\right)\left(e_n\right)=S\left(e_n\right)$.

Now, forget that we fixed n. We thus have shown that $\left(\omega\circ D_q\right)\left(e_n\right)=S\left(e_n\right)$ for each positive integer n. In other words, the two $\mathbf{k}$-algebra homomorphisms $\omega\circ D_q$ and S (from Λ to Λ) agree on each element of the family $\left\{e_n\right\}_{n\geq 1}$ of elements of Λ.

But Proposition 2.4.1 tells us that the family $\left\{e_n\right\}_{n\geq 1}$ generates the $\mathbf{k}$-algebra Λ. Thus, if two $\mathbf{k}$-algebra homomorphisms from Λ agree on each element of this family $\left\{e_n\right\}_{n\geq 1}$, then these two homomorphisms must be equal. Hence, the two $\mathbf{k}$-algebra homomorphisms $\omega\circ D_q$ and S must be equal (since they agree on each element of this family $\left\{e_n\right\}_{n\geq 1}$). In other words, we have $\omega\circ D_q=S$.

Now, let $n \in \mathbb{N}$ and let $f \in \Lambda_n$. We must show that $S(f) = (-1)^n \omega(f)$. And indeed, we have $\underbrace{(\omega \circ D_q)}_{=S}(f) = S(f)$, so that

$$\begin{aligned} S(f) = (\omega \circ D_q)(f) = \omega\left(\underbrace{D_q(f)}_{\substack{=q^n f \\ \text{(by (13.53.1), applied to } a=f)}}\right) = \omega(q^n f) \\ = q^n \omega(f) \qquad \text{(since the map } \omega \text{ is } \mathbf{k}\text{-linear)} \\ = (-1)^n \omega(f) \qquad \text{(since } q = -1). \end{aligned}$$

This proves Proposition 2.4.3(e).

(c) Let n be a positive integer. Proposition 2.4.1(i) yields $S(p_n) = -p_n$. But (2.4.11) (applied to $f = p_n$) yields $S(p_n) = (-1)^n \omega(p_n)$ (since $p_n \in \Lambda_n$). Comparing these two equalities yields $(-1)^n \omega(p_n) = -p_n$, and this quickly rewrites as $\omega(p_n) = (-1)^{n-1} p_n$. This proves Proposition 2.4.3(c).

(f) To show that ω is a coalgebra homomorphism, it suffices to check that $(\omega \otimes \omega) \circ \Delta = \Delta \circ \omega$ and $\epsilon = \epsilon \circ \omega$.

Let us first prove that $(\omega \otimes \omega) \circ \Delta = \Delta \circ \omega$. Indeed, both sides of this equality being $\mathbf{k}$-algebra homomorphisms, it only needs to be checked on the algebra generators h_n of Λ (indeed, we know from Proposition 2.4.1 that the family $\{h_n\}_{n \geq 1}$ generates the $\mathbf{k}$-algebra Λ). On these generators this is easy to check: Comparing

$$\begin{aligned} ((\omega \otimes \omega) \circ \Delta)(h_n) = (\omega \otimes \omega)(\Delta(h_n)) = (\omega \otimes \omega)\left(\sum_{i+j=n} h_i \otimes h_j\right) \qquad \text{(by Proposition 2.3.6(iii))} \\ = \sum_{i+j=n} \underbrace{\omega(h_i)}_{\substack{=e_i \\ \text{(by Proposition 2.4.3(b))}}} \otimes \underbrace{\omega(h_j)}_{\substack{=e_j \\ \text{(by Proposition 2.4.3(b))}}} = \sum_{i+j=n} e_i \otimes e_j \end{aligned}$$

and

$$(\Delta \circ \omega)(h_n) = \Delta\left(\underbrace{\omega(h_n)}_{\substack{=e_n \\ \text{(by Proposition 2.4.3(b))}}}\right) = \Delta(e_n) = \sum_{i+j=n} e_i \otimes e_j \qquad \text{(by Proposition 2.3.6(ii))}$$

shows that $((\omega \otimes \omega) \circ \Delta)(h_n) = (\Delta \circ \omega)(h_n)$ for all $n \geq 1$. So we obtain $(\omega \otimes \omega) \circ \Delta = \Delta \circ \omega$. The equality $\epsilon = \epsilon \circ \omega$ is even easier to check. Thus, ω is an algebra and coalgebra morphism, thus a bialgebra morphism, thus a Hopf morphism by Corollary 1.4.27. Since ω is invertible (by part (d)), we conclude that ω is a Hopf algebra automorphism. This proves Proposition 2.4.3(f).

(g) Part (g) of Proposition 2.4.3 can be proved in the same way as we proved part (f) above, with the following three differences:

- We need to know that S is a $\mathbf{k}$-algebra homomorphism. (This was shown during the proof of Proposition 2.4.3(e) above.)
- We need to know that S is invertible. (This can be concluded from Corollary 1.4.12, or from the fact that ω is invertible using Proposition 2.4.3(e).)
- Instead of the formula $\omega(h_n) = e_n$ we now need to use the formula $S(h_n) = (-1)^n e_n$ (which is Proposition 2.4.1(iii)), and thus we incur some signs in the computation.

[*Remark:* Let us sketch alternative approaches to proving parts (f) and (g) of Proposition 2.4.3:

Alternative proof of Proposition 2.4.3(g) (sketched): We have already seen (during the proof of Proposition 2.4.3(e) above) that S is a $\mathbf{k}$-algebra homomorphism. Corollary 1.4.12 (applied to $S = \Lambda$) yields that S is an involution; thus, S is invertible. Moreover, the Hopf algebra Λ is cocommutative (by Exercise 2.3.7(a)). But Exercise 1.4.28 (applied to $A = \Lambda$) yields that the antipode S of Λ is a $\mathbf{k}$-coalgebra anti-endomorphism of Λ. In other words, S is a $\mathbf{k}$-coalgebra anti-homomorphism from Λ to Λ. But Exercise 1.5.8(b) (applied to $A = \Lambda$ and $B = \Lambda$) shows that the $\mathbf{k}$-coalgebra anti-homomorphisms from Λ to Λ are the same as the

k-coalgebra homomorphisms from Λ to Λ (since the **k**-coalgebra Λ is cocommutative). Hence, S is a **k**-coalgebra homomorphism from Λ to Λ (since S is a **k**-coalgebra anti-homomorphism from Λ to Λ). We now know that S is a **k**-algebra homomorphism and a **k**-coalgebra homomorphism at the same time. Hence, S is a **k**-bialgebra homomorphism, therefore a Hopf morphism (by Corollary 1.4.27), and thus a Hopf algebra automorphism (since S is invertible). This proves Proposition 2.4.3(g) again.

Alternative proof of Proposition 2.4.3(f) (sketched): Proposition 2.4.3(g) shows that S is a Hopf algebra automorphism. Define q and D_q as in the proof of Proposition 2.4.3(e) above. It is easy to see that D_q is an involution; thus, D_q is invertible. But D_q is a **k**-bialgebra homomorphism (as we have seen in the proof of Proposition 2.4.3(e) above), therefore a Hopf morphism (by Corollary 1.4.27), and thus a Hopf algebra automorphism (since D_q is invertible). Now, recall that $\omega \circ D_q = S$ (as we have seen in the proof of Proposition 2.4.3(e) above). Hence, $\omega = S \circ (D_q)^{-1}$ (since D_q is invertible). This shows that ω is a Hopf algebra automorphism (since both S and D_q are Hopf algebra automorphisms). This proves Proposition 2.4.3(f) again.]

(h) Let λ be a partition.

We can easily obtain (2.4.14) by multiplicativity:

[*Proof of* (2.4.14)*:* Let ℓ be the length $\ell(\lambda)$ of the partition λ. Then, p_λ is defined as $p_{\lambda_1} p_{\lambda_2} \cdots p_{\lambda_\ell}$, and thus we have

$$\begin{aligned}
\omega(p_\lambda) &= \omega(p_{\lambda_1} p_{\lambda_2} \cdots p_{\lambda_\ell}) = \omega(p_{\lambda_1})\,\omega(p_{\lambda_2}) \cdots \omega(p_{\lambda_\ell}) && \text{(since } \omega \text{ is a } \mathbf{k}\text{-algebra homomorphism)} \\
&= (-1)^{\lambda_1 - 1} p_{\lambda_1} \cdot (-1)^{\lambda_2 - 1} p_{\lambda_2} \cdots (-1)^{\lambda_\ell - 1} p_{\lambda_\ell} && \text{(by Proposition 2.4.3(c))} \\
&= \underbrace{(-1)^{(\lambda_1 + \lambda_2 + \cdots + \lambda_\ell) - \ell}}_{\substack{= (-1)^{|\lambda| - \ell(\lambda)} \\ \text{(since } \lambda_1 + \lambda_2 + \cdots + \lambda_\ell = |\lambda| \\ \text{and } \ell = \ell(\lambda))}} \underbrace{p_{\lambda_1} p_{\lambda_2} \cdots p_{\lambda_\ell}}_{= p_\lambda} = (-1)^{|\lambda| - \ell(\lambda)} p_\lambda.
\end{aligned}$$

This proves (2.4.14).]

A similar argument (but using Proposition 2.4.3(b) instead of Proposition 2.4.3(c)) proves (2.4.12). Finally, a similar argument (but using Proposition 2.4.3(a) instead of Proposition 2.4.3(c)) proves (2.4.13).

We have now proved all three equalities (2.4.12), (2.4.13) and (2.4.14). Thus, Proposition 2.4.3(h) is proved.

(i) Proposition 2.4.3(d) yields that the map ω is a **k**-algebra automorphism of Λ and an involution. Thus, ω is an involution. In other words, ω is invertible, and the inverse of ω is ω itself.

We know that Λ is a connected graded bialgebra. Thus, Proposition 1.4.16 (applied to $A = \Lambda$) yields that Λ has a unique antipode S, which is a graded map $\Lambda \xrightarrow{S} \Lambda$, endowing it with a Hopf structure. Thus, in particular, the antipode S of Λ is a graded map. In other words,

$$S(\Lambda_n) \subset \Lambda_n \qquad \text{for each } n \in \mathbb{N}. \tag{13.53.2}$$

Now, let $n \in \mathbb{N}$. Let $f \in \Lambda_n$. Then, Proposition 2.4.3(e) yields $S(f) = (-1)^n \omega(f)$. Hence,

$$(-1)^n \underbrace{S(f)}_{= (-1)^n \omega(f)} = \underbrace{(-1)^n (-1)^n}_{\substack{= (-1)^{n+n} = (-1)^{2n} = 1 \\ \text{(since } 2n \text{ is even)}}} \omega(f) = \omega(f),$$

so that $\omega(f) = (-1)^n S\left(\underbrace{f}_{\in \Lambda_n}\right) \in (-1)^n \underbrace{S(\Lambda_n)}_{\substack{\subset \Lambda_n \\ \text{(by (13.53.2))}}} \subset (-1)^n \Lambda_n \subset \Lambda_n$ (since Λ_n is a **k**-module).

Forget that we fixed f. We thus have proved that $\omega(f) \in \Lambda_n$ for each $f \in \Lambda_n$. In other words, $\omega(\Lambda_n) \subset \Lambda_n$.

Forget that we fixed n. We thus have shown that $\omega(\Lambda_n) \subset \Lambda_n$ for each $n \in \mathbb{N}$. Hence, the **k**-linear map $\omega : \Lambda \to \Lambda$ is graded. In other words, the inverse of ω is graded (since the inverse of ω is ω itself).

Now, we know that the **k**-linear map ω is graded, and its inverse is also graded. In other words, the **k**-linear map ω is an isomorphism of graded **k**-modules. This proves Proposition 2.4.3(i).

(j) From (2.4.13), we know that $\omega(e_\lambda) = h_\lambda$ for each partition λ. In other words, $\omega(e_\lambda) = h_\lambda$ for each $\lambda \in \operatorname{Par}$. In other words, $(\omega(e_\lambda))_{\lambda \in \operatorname{Par}} = (h_\lambda)_{\lambda \in \operatorname{Par}}$.

Proposition 2.2.10 says (among other things) that the family $\{e_\lambda\}$, as λ runs through all partitions, is a graded basis of the graded $\mathbf{k}$-module Λ. In other words, the family $(e_\lambda)_{\lambda\in\operatorname{Par}}$ is a graded basis of the graded $\mathbf{k}$-module Λ. On the other hand, Proposition 2.4.3(i) shows that the map ω is an isomorphism of graded $\mathbf{k}$-modules.

Now, recall the following well-known fact: If V and W are two graded $\mathbf{k}$-modules, and if $f : V \to W$ is an isomorphism of graded $\mathbf{k}$-modules, then f sends any graded basis of V to a graded basis of W. In other words, if V and W are two graded $\mathbf{k}$-modules, and if $f : V \to W$ is an isomorphism of graded $\mathbf{k}$-modules, and if $(v_i)_{i\in I}$ is a graded basis of V, then $(f(v_i))_{i\in I}$ is a graded basis of W.

We can apply this fact to $V = \Lambda$, $W = \Lambda$, $f = \omega$, $I = \operatorname{Par}$ and $(v_i)_{i\in I} = (e_\lambda)_{\lambda\in\operatorname{Par}}$ (since we know that $\omega : \Lambda \to \Lambda$ is an isomorphism of graded $\mathbf{k}$-modules, and since we know that the family $(e_\lambda)_{\lambda\in\operatorname{Par}}$ is a graded basis of Λ). We thus conclude that $(\omega(e_\lambda))_{\lambda\in\operatorname{Par}}$ is a graded basis of Λ. In other words, the family $(h_\lambda)_{\lambda\in\operatorname{Par}}$ is a graded basis of Λ (since $(\omega(e_\lambda))_{\lambda\in\operatorname{Par}} = (h_\lambda)_{\lambda\in\operatorname{Par}}$). This proves Proposition 2.4.3(j). □

We have now proved Proposition 2.4.3; thus, Exercise 2.4.4 is solved.

13.54. **Solution to Exercise 2.5.5.** *Solution to Exercise 2.5.5.* Recall that for each partition λ,

$$\text{the element } q_\lambda \in \Lambda \text{ is homogeneous of degree } |\lambda| . \tag{13.54.1}$$

(a) Let $(a_\lambda)_{\lambda\in\operatorname{Par}} \in \mathbf{k}^{\operatorname{Par}}$ and $(b_\lambda)_{\lambda\in\operatorname{Par}} \in \mathbf{k}^{\operatorname{Par}}$ be two families satisfying (2.5.3) in $\mathbf{k}[[\mathbf{x}]]$. We must prove that $(a_\lambda)_{\lambda\in\operatorname{Par}} = (b_\lambda)_{\lambda\in\operatorname{Par}}$.

Fix $n \in \mathbb{N}$. Consider the $\mathbf{k}$-linear map $\pi_n : \mathbf{k}[[\mathbf{x}]] \to \mathbf{k}[[\mathbf{x}]]$ that sends each power series $f \in \mathbf{k}[[\mathbf{x}]]$ to its homogeneous component of degree n. Thus, π_n has the following properties:

- If $f \in \mathbf{k}[[\mathbf{x}]]$ is a power series that is homogeneous of degree n, then
$$\pi_n(f) = f. \tag{13.54.2}$$
- If $f \in \mathbf{k}[[\mathbf{x}]]$ is a power series that is homogeneous of degree $\neq n$, then
$$\pi_n(f) = 0. \tag{13.54.3}$$
- The map π_n is $\mathbf{k}$-linear and continuous (with respect to the topology on $\mathbf{k}[[\mathbf{x}]]$).

Hence, each $\lambda \in \operatorname{Par}$ satisfying $|\lambda| = n$ satisfies

$$\pi_n(q_\lambda(\mathbf{x})) = q_\lambda(\mathbf{x}) \tag{13.54.4}$$

[576]. Furthermore, each $\lambda \in \operatorname{Par}$ satisfying $|\lambda| \neq n$ satisfies

$$\pi_n(q_\lambda(\mathbf{x})) = 0 \tag{13.54.5}$$

[577].

But the map π_n is $\mathbf{k}$-linear and continuous. Thus, it respects infinite sums. Hence,

$$\pi_n\left(\sum_{\lambda\in\operatorname{Par}} a_\lambda q_\lambda(\mathbf{x})\right) = \sum_{\lambda\in\operatorname{Par}} a_\lambda \pi_n(q_\lambda(\mathbf{x})) = \sum_{\substack{\lambda\in\operatorname{Par};\\ |\lambda|=n}} a_\lambda \underbrace{\pi_\lambda(q_\lambda(\mathbf{x}))}_{\substack{=q_\lambda(\mathbf{x})\\ \text{(by (13.54.4))}}} + \sum_{\substack{\lambda\in\operatorname{Par};\\ |\lambda|\neq n}} a_\lambda \underbrace{\pi_\lambda(q_\lambda(\mathbf{x}))}_{\substack{=0\\ \text{(by (13.54.5))}}}$$
$$= \sum_{\substack{\lambda\in\operatorname{Par};\\ |\lambda|=n}} a_\lambda q_\lambda(\mathbf{x}) + \underbrace{\sum_{\substack{\lambda\in\operatorname{Par};\\ |\lambda|\neq n}} a_\lambda 0}_{=0} = \sum_{\substack{\lambda\in\operatorname{Par};\\ |\lambda|=n}} a_\lambda q_\lambda(\mathbf{x}) . \tag{13.54.6}$$

[576] *Proof of (13.54.4):* Let $\lambda \in \operatorname{Par}$ be such that $|\lambda| = n$. From (13.54.1), we know that the element $q_\lambda \in \Lambda$ is homogeneous of degree $|\lambda|$. Thus, the power series $q_\lambda(\mathbf{x})$ is homogeneous of degree $|\lambda| = n$. Hence, (13.54.2) (applied to $f = q_\lambda(\mathbf{x})$) yields $\pi_n(q_\lambda(\mathbf{x})) = q_\lambda(\mathbf{x})$. This proves (13.54.4).

[577] *Proof of (13.54.5):* Let $\lambda \in \operatorname{Par}$ be such that $|\lambda| \neq n$. From (13.54.1), we know that the element $q_\lambda \in \Lambda$ is homogeneous of degree $|\lambda|$. Thus, the power series $q_\lambda(\mathbf{x})$ is homogeneous of degree $|\lambda| \neq n$. Hence, (13.54.3) (applied to $f = q_\lambda(\mathbf{x})$) yields $\pi_n(q_\lambda(\mathbf{x})) = 0$. This proves (13.54.5).

The same argument (applied to the family $(b_\lambda)_{\lambda\in\text{Par}}$ instead of $(a_\lambda)_{\lambda\in\text{Par}}$) yields

$$\pi_n\left(\sum_{\lambda\in\text{Par}} b_\lambda q_\lambda(\mathbf{x})\right) = \sum_{\substack{\lambda\in\text{Par};\\ |\lambda|=n}} b_\lambda q_\lambda(\mathbf{x}). \tag{13.54.7}$$

Now, applying the map π_n to both sides of the equality (2.5.3), we obtain

$$\pi_n\left(\sum_{\lambda\in\text{Par}} a_\lambda q_\lambda(\mathbf{x})\right) = \pi_n\left(\sum_{\lambda\in\text{Par}} b_\lambda q_\lambda(\mathbf{x})\right) = \sum_{\substack{\lambda\in\text{Par};\\ |\lambda|=n}} b_\lambda q_\lambda(\mathbf{x}).$$

Comparing this with (13.54.6), we obtain

$$\sum_{\substack{\lambda\in\text{Par};\\ |\lambda|=n}} a_\lambda q_\lambda(\mathbf{x}) = \sum_{\substack{\lambda\in\text{Par};\\ |\lambda|=n}} b_\lambda q_\lambda(\mathbf{x}). \tag{13.54.8}$$

Notice that both sums appearing in this equality are finite (since there are only finitely many $\lambda \in \text{Par}$ satisfying $|\lambda| = n$). Since the family $(q_\lambda(\mathbf{x}))_{\lambda\in\text{Par};\ |\lambda|=n}$ is $\mathbf{k}$-linearly independent[578], we can thus conclude from (13.54.8) that

$$a_\lambda = b_\lambda \qquad \text{for each } \lambda\in\text{Par} \text{ satisfying } |\lambda| = n. \tag{13.54.9}$$

Now, forget that we fixed n. We thus have proven (13.54.9) for each $n \in \mathbb{N}$.

Now, let $\lambda \in \text{Par}$ be arbitrary. Then, $|\lambda| \in \mathbb{N}$. Hence, (13.54.9) (applied to $n = |\lambda|$) yields $a_\lambda = b_\lambda$.

Now, forget that we fixed λ. We thus have proven that $a_\lambda = b_\lambda$ for each $\lambda \in \text{Par}$. In other words, $(a_\lambda)_{\lambda\in\text{Par}} = (b_\lambda)_{\lambda\in\text{Par}}$. This solves Exercise 2.5.5 (a).

(c) We first go afield. Recall that $(m_\lambda)_{\lambda\in\text{Par}}$ is a basis of the $\mathbf{k}$-module Λ. Hence, the family $(m_\mu \otimes m_\nu \otimes m_\lambda)_{(\mu,\nu,\lambda)\in\text{Par}^3}$ is a basis of the $\mathbf{k}$-module $\Lambda\otimes\Lambda\otimes\Lambda$.

On the other hand, $(q_\lambda)_{\lambda\in\text{Par}}$ is a basis of the $\mathbf{k}$-module Λ. Hence, the family $(q_\mu \otimes q_\nu \otimes q_\lambda)_{(\mu,\nu,\lambda)\in\text{Par}^3}$ is a basis of the $\mathbf{k}$-module $\Lambda\otimes\Lambda\otimes\Lambda$.

We shall now show that the family $(m_\mu(\mathbf{x})\, m_\nu(\mathbf{y})\, m_\lambda(\mathbf{z}))_{(\mu,\nu,\lambda)\in\text{Par}^3}$ of elements of $\mathbf{k}[[\mathbf{x},\mathbf{y},\mathbf{z}]]$ is $\mathbf{k}$-linearly independent.

If $\mathfrak{m}$ is a monomial, and if f is a power series, then we let $[\mathfrak{m}]\, f$ denote the coefficient of $\mathfrak{m}$ in f.

Any $\tau \in \text{Par}$ and $\lambda \in \text{Par}$ satisfy

$$[\mathbf{x}^\tau]\,(m_\lambda(\mathbf{x})) = \delta_{\tau,\lambda} \tag{13.54.10}$$

[579].

For any three partitions $\alpha, \beta, \gamma \in \text{Par}$ and any three partitions $\mu, \nu, \lambda \in \text{Par}$, we have

$$\left[\mathbf{x}^\alpha \mathbf{y}^\beta \mathbf{z}^\gamma\right](m_\mu(\mathbf{x})\, m_\nu(\mathbf{y})\, m_\lambda(\mathbf{z})) = \delta_{(\alpha,\beta,\gamma),(\mu,\nu,\lambda)} \tag{13.54.12}$$

[578] *Proof.* The family $\left(\underbrace{q_\lambda(\mathbf{x})}_{=q_\lambda}\right)_{\lambda\in\text{Par}} = (q_\lambda)_{\lambda\in\text{Par}}$ is a basis of the $\mathbf{k}$-module Λ, and thus is $\mathbf{k}$-linearly independent. Hence, the family $(q_\lambda(\mathbf{x}))_{\lambda\in\text{Par};\ |\lambda|=n}$ is $\mathbf{k}$-linearly independent as well (since it is a subfamily of this family $(q_\lambda(\mathbf{x}))_{\lambda\in\text{Par}}$).

[579] *Proof of (13.54.10):* Let $\tau \in \text{Par}$ and $\lambda \in \text{Par}$. The set $\mathfrak{S}_{(\infty)}\lambda$ clearly contains λ, since $\lambda = \underbrace{\text{id}}_{\in\mathfrak{S}_{(\infty)}} \cdot\lambda \in \mathfrak{S}_{(\infty)}\lambda$. Moreover, each $\alpha \in \mathfrak{S}_{(\infty)}\lambda$ satisfying $\alpha \neq \lambda$ must satisfy

$$\delta_{\tau,\alpha} = 0. \tag{13.54.11}$$

[*Proof of (13.54.11):* Let $\alpha \in \mathfrak{S}_{(\infty)}\lambda$ be such that $\alpha \neq \lambda$.

Assume (for the sake of contradiction) that $\tau = \alpha$. Then, $\alpha = \tau \in \text{Par}$. Hence, the sequence α is nonincreasing. But the sequence λ is also nonincreasing (since $\lambda \in \text{Par}$).

From $\alpha \in \mathfrak{S}_{(\infty)}\lambda$, we conclude that α is a rearrangement of the partition λ. But λ is also a rearrangement of λ. However, there is only one nonincreasing rearrangement of λ. In other words, if μ and ν are two nonincreasing rearrangements of λ, then $\mu = \nu$. Applying this to $\mu = \alpha$ and $\nu = \lambda$, we conclude that $\alpha = \lambda$ (since both α and λ are nonincreasing rearrangements of λ). This contradicts $\alpha \neq \lambda$.

This contradiction shows that our assumption (that $\tau = \alpha$) was false. Hence, we have $\tau \neq \alpha$. Thus, $\delta_{\tau,\alpha} = 0$. This proves (13.54.11).]

[580].

Hence, if $(c_{\mu,\nu,\lambda})_{(\mu,\nu,\lambda)\in\operatorname{Par}^3} \in \mathbf{k}^{\operatorname{Par}^3}$ is a family of elements of $\mathbf{k}$ satisfying

$$\sum_{(\mu,\nu,\lambda)\in\operatorname{Par}^3} c_{\mu,\nu,\lambda} m_\mu(\mathbf{x}) m_\nu(\mathbf{y}) m_\lambda(\mathbf{z}) = 0, \tag{13.54.13}$$

then $(c_{\mu,\nu,\lambda})_{(\mu,\nu,\lambda)\in\operatorname{Par}^3} = (0)_{(\mu,\nu,\lambda)\in\operatorname{Par}^3}$ [581]. Thus, the family $(m_\mu(\mathbf{x}) m_\nu(\mathbf{y}) m_\lambda(\mathbf{z}))_{(\mu,\nu,\lambda)\in\operatorname{Par}^3}$ of elements of $\mathbf{k}[[\mathbf{x},\mathbf{y},\mathbf{z}]]$ is $\mathbf{k}$-linearly independent.

Hence, the $\mathbf{k}$-linear map

$$\Lambda\otimes\Lambda\otimes\Lambda \to \mathbf{k}[[\mathbf{x},\mathbf{y},\mathbf{z}]], \qquad f\otimes g\otimes h \mapsto f(\mathbf{x})\,g(\mathbf{y})\,h(\mathbf{z})$$

We have $m_\lambda(\mathbf{x}) = m_\lambda = \sum_{\alpha\in\mathfrak{S}_{(\infty)}\lambda} \mathbf{x}^\alpha$ (by (2.1.1)). Thus,

$$\begin{aligned}
[\mathbf{x}^\tau]\left(\underbrace{m_\lambda(\mathbf{x})}_{=\sum_{\alpha\in\mathfrak{S}_{(\infty)}\lambda}\mathbf{x}^\alpha}\right) &= [\mathbf{x}^\tau]\left(\sum_{\alpha\in\mathfrak{S}_{(\infty)}\lambda}\mathbf{x}^\alpha\right) = \sum_{\alpha\in\mathfrak{S}_{(\infty)}\lambda}\underbrace{[\mathbf{x}^\tau](\mathbf{x}^\alpha)}_{=\delta_{\tau,\alpha}} = \sum_{\alpha\in\mathfrak{S}_{(\infty)}\lambda}\delta_{\tau,\alpha}\\
&= \delta_{\tau,\lambda} + \sum_{\substack{\alpha\in\mathfrak{S}_{(\infty)}\lambda;\\ \alpha\neq\lambda}} \underbrace{\delta_{\tau,\alpha}}_{\substack{=0\\ \text{(by (13.54.11))}}}\\
&\qquad\left(\begin{array}{c}\text{here, we have split off the addend for } \alpha=\lambda \text{ from}\\ \text{the sum (since } \lambda\in\mathfrak{S}_{(\infty)}\lambda)\end{array}\right)\\
&= \delta_{\tau,\lambda} + \underbrace{\sum_{\substack{\alpha\in\mathfrak{S}_{(\infty)}\lambda;\\ \alpha\neq\lambda}} 0}_{=0} = \delta_{\tau,\lambda}.
\end{aligned}$$

This proves (13.54.10).

[580] *Proof of (13.54.12):* Let $\alpha,\beta,\gamma\in\operatorname{Par}$ be three partitions. Let $\mu,\nu,\lambda\in\operatorname{Par}$ be three partitions. Then, (13.54.10) (applied to α and μ instead of τ and λ) yields $[\mathbf{x}^\alpha](m_\mu(\mathbf{x})) = \delta_{\alpha,\mu}$. Also, (13.54.10) (applied to β and ν instead of τ and λ) yields $[\mathbf{x}^\beta](m_\nu(\mathbf{x})) = \delta_{\beta,\nu}$. Renaming the indeterminates $\mathbf{x}$ as $\mathbf{y}$ in this fact, we obtain $[\mathbf{y}^\beta](m_\nu(\mathbf{y})) = \delta_{\beta,\nu}$. Finally, (13.54.10) (applied to γ and λ instead of τ and λ) yields $[\mathbf{x}^\gamma](m_\lambda(\mathbf{x})) = \delta_{\gamma,\lambda}$. Renaming the indeterminates $\mathbf{x}$ as $\mathbf{z}$ in this fact, we obtain $[\mathbf{z}^\gamma](m_\lambda(\mathbf{z})) = \delta_{\gamma,\lambda}$. Now,

$$\begin{aligned}
\left[\mathbf{x}^\alpha\mathbf{y}^\beta\mathbf{z}^\gamma\right](m_\mu(\mathbf{x})m_\nu(\mathbf{y})m_\lambda(\mathbf{z})) &= \underbrace{([\mathbf{x}^\alpha](m_\mu(\mathbf{x})))}_{=\delta_{\alpha,\mu}}\cdot\underbrace{\left(\left[\mathbf{y}^\beta\right](m_\nu(\mathbf{y}))\right)}_{=\delta_{\beta,\nu}}\cdot\underbrace{([\mathbf{z}^\gamma](m_\lambda(\mathbf{z})))}_{=\delta_{\gamma,\lambda}}\\
&= \delta_{\alpha,\mu}\cdot\delta_{\beta,\nu}\cdot\delta_{\gamma,\lambda} = \delta_{(\alpha,\beta,\gamma),(\mu,\nu,\lambda)}.
\end{aligned}$$

This proves (13.54.12).

[581] *Proof.* Let $(c_{\mu,\nu,\lambda})_{(\mu,\nu,\lambda)\in\operatorname{Par}^3} \in \mathbf{k}^{\operatorname{Par}^3}$ be a family of elements of $\mathbf{k}$ satisfying (13.54.13). We must show that $(c_{\mu,\nu,\lambda})_{(\mu,\nu,\lambda)\in\operatorname{Par}^3} = (0)_{(\mu,\nu,\lambda)\in\operatorname{Par}^3}$.

Fix any $(\alpha,\beta,\gamma)\in\operatorname{Par}^3$. Then,

$$\begin{aligned}
&\left[\mathbf{x}^\alpha\mathbf{y}^\beta\mathbf{z}^\gamma\right]\left(\sum_{(\mu,\nu,\lambda)\in\operatorname{Par}^3} c_{\mu,\nu,\lambda} m_\mu(\mathbf{x}) m_\nu(\mathbf{y}) m_\lambda(\mathbf{z})\right)\\
&= \sum_{(\mu,\nu,\lambda)\in\operatorname{Par}^3} c_{\mu,\nu,\lambda}\underbrace{\left[\mathbf{x}^\alpha\mathbf{y}^\beta\mathbf{z}^\gamma\right](m_\mu(\mathbf{x})m_\nu(\mathbf{y})m_\lambda(\mathbf{z}))}_{\substack{=\delta_{(\alpha,\beta,\gamma),(\mu,\nu,\lambda)}\\ \text{(by (13.54.12))}}}\\
&= \sum_{(\mu,\nu,\lambda)\in\operatorname{Par}^3} c_{\mu,\nu,\lambda}\delta_{(\alpha,\beta,\gamma),(\mu,\nu,\lambda)}\\
&= c_{\alpha,\beta,\gamma}\underbrace{\delta_{(\alpha,\beta,\gamma),(\alpha,\beta,\gamma)}}_{=1} + \sum_{\substack{(\mu,\nu,\lambda)\in\operatorname{Par}^3;\\ (\mu,\nu,\lambda)\neq(\alpha,\beta,\gamma)}} c_{\mu,\nu,\lambda}\underbrace{\delta_{(\alpha,\beta,\gamma),(\mu,\nu,\lambda)}}_{\substack{=0\\ \text{(since } (\alpha,\beta,\gamma)\neq(\mu,\nu,\lambda))}}\\
&\qquad\text{(here, we have split off the addend for } (\mu,\nu,\lambda)=(\alpha,\beta,\gamma) \text{ from the sum)}\\
&= c_{\alpha,\beta,\gamma} + \underbrace{\sum_{\substack{(\mu,\nu,\lambda)\in\operatorname{Par}^3;\\ (\mu,\nu,\lambda)\neq(\alpha,\beta,\gamma)}} c_{\mu,\nu,\lambda}0}_{=0} = c_{\alpha,\beta,\gamma},
\end{aligned}$$

maps the basis $(m_\mu \otimes m_\nu \otimes m_\lambda)_{(\mu,\nu,\lambda)\in \operatorname{Par}^3}$ of $\Lambda \otimes \Lambda \otimes \Lambda$ to the linearly independent family $(m_\mu(\mathbf{x}) m_\nu(\mathbf{y}) m_\lambda(\mathbf{z}))_{(\mu,\nu,\lambda)\in \operatorname{Par}^3}$. Consequently, this map is injective[582]. Therefore, the family $(q_\mu(\mathbf{x}) q_\nu(\mathbf{y}) q_\lambda(\mathbf{z}))_{(\mu,\nu,\lambda)\in \operatorname{Par}^3}$ of elements of $\mathbf{k}[[\mathbf{x},\mathbf{y},\mathbf{z}]]$ is also linearly independent (because it is the image of the basis $(q_\mu \otimes q_\nu \otimes q_\lambda)_{(\mu,\nu,\lambda)\in \operatorname{Par}^3}$ of $\Lambda \otimes \Lambda \otimes \Lambda$ under this injective map[583]). If the sums appearing in (2.5.5) were finite, then this observation would already yield Exercise 2.5.5 (c). However, these sums are infinite (and linear independence makes no claims about infinite sums being 0), so the solution of Exercise 2.5.5 (c) takes some more work:

Let $(a_{\lambda,\mu,\nu})_{(\mu,\nu,\lambda)\in \operatorname{Par}^3} \in \mathbf{k}^{\operatorname{Par}^3}$ and $(b_{\lambda,\mu,\nu})_{(\mu,\nu,\lambda)\in \operatorname{Par}^3} \in \mathbf{k}^{\operatorname{Par}^3}$ be two families satisfying (2.5.5) in $\mathbf{k}[[\mathbf{x},\mathbf{y},\mathbf{z}]]$. We must prove that $(a_{\lambda,\mu,\nu})_{(\mu,\nu,\lambda)\in \operatorname{Par}^3} = (b_{\lambda,\mu,\nu})_{(\mu,\nu,\lambda)\in \operatorname{Par}^3}$.

For each $(\mu,\nu,\lambda) \in \operatorname{Par}^3$,

$$\text{the power series } q_\mu(\mathbf{x}) q_\nu(\mathbf{y}) q_\lambda(\mathbf{z}) \text{ is homogeneous of degree } |\mu| + |\nu| + |\lambda| \tag{13.54.14}$$

[584].

Fix $n \in \mathbb{N}$. Note that there are only finitely many $(\mu,\nu,\lambda) \in \operatorname{Par}^3$ satisfying $|\mu| + |\nu| + |\lambda| = n$ [585].

Consider the $\mathbf{k}$-linear map $\pi_n : \mathbf{k}[[\mathbf{x},\mathbf{y},\mathbf{z}]] \to \mathbf{k}[[\mathbf{x},\mathbf{y},\mathbf{z}]]$ that sends each power series $f \in \mathbf{k}[[\mathbf{x},\mathbf{y},\mathbf{z}]]$ to its homogeneous component of degree n. Thus, π_n has the following properties:

- If $f \in \mathbf{k}[[\mathbf{x},\mathbf{y},\mathbf{z}]]$ is a power series that is homogeneous of degree n, then

$$\pi_n(f) = f. \tag{13.54.15}$$

- If $f \in \mathbf{k}[[\mathbf{x},\mathbf{y},\mathbf{z}]]$ is a power series that is homogeneous of degree $\neq n$, then

$$\pi_n(f) = 0. \tag{13.54.16}$$

- The map π_n is $\mathbf{k}$-linear and continuous (with respect to the topology on $\mathbf{k}[[\mathbf{x},\mathbf{y},\mathbf{z}]]$).

Hence, each $(\mu,\nu,\lambda) \in \operatorname{Par}^3$ satisfying $|\mu| + |\nu| + |\lambda| = n$ satisfies

$$\pi_n(q_\mu(\mathbf{x}) q_\nu(\mathbf{y}) q_\lambda(\mathbf{z})) = q_\mu(\mathbf{x}) q_\nu(\mathbf{y}) q_\lambda(\mathbf{z}) \tag{13.54.17}$$

[586]. Furthermore, each $(\mu,\nu,\lambda) \in \operatorname{Par}^3$ satisfying $|\mu| + |\nu| + |\lambda| \neq n$ satisfies

$$\pi_n(q_\mu(\mathbf{x}) q_\nu(\mathbf{y}) q_\lambda(\mathbf{z})) = 0 \tag{13.54.18}$$

[587].

so that

$$c_{\alpha,\beta,\gamma} = \left[\mathbf{x}^\alpha \mathbf{y}^\beta \mathbf{z}^\gamma\right] \left(\underbrace{\sum_{(\mu,\nu,\lambda)\in \operatorname{Par}^3} c_{\mu,\nu,\lambda} m_\mu(\mathbf{x}) m_\nu(\mathbf{y}) m_\lambda(\mathbf{z})}_{=0} \right) = \left[\mathbf{x}^\alpha \mathbf{y}^\beta \mathbf{z}^\gamma\right] 0 = 0.$$

Now, forget that we fixed (α,β,γ). We thus have shown that $c_{\alpha,\beta,\gamma} = 0$ for each $(\alpha,\beta,\gamma) \in \operatorname{Par}^3$. In other words, $(c_{\alpha,\beta,\gamma})_{(\alpha,\beta,\gamma)\in \operatorname{Par}^3} = (0)_{(\alpha,\beta,\gamma)\in \operatorname{Par}^3}$. Renaming the index (α,β,γ) as (μ,ν,λ) in this equality, we obtain $(c_{\mu,\nu,\lambda})_{(\mu,\nu,\lambda)\in \operatorname{Par}^3} = (0)_{(\mu,\nu,\lambda)\in \operatorname{Par}^3}$. Qed.

[582] because any $\mathbf{k}$-linear map that maps a basis of its domain to a linearly independent family must be injective

[583] Here, we are using the fact that the image of a basis under an injective $\mathbf{k}$-linear map is always linearly independent.

[584] *Proof of (13.54.14):* Let $(\mu,\nu,\lambda) \in \operatorname{Par}^3$.

The power series $q_\lambda(\mathbf{x}) = q_\lambda$ is homogeneous of degree $|\lambda|$ (by (13.54.1)). Renaming the indeterminates $\mathbf{x}$ as $\mathbf{z}$ in this fact, we conclude that the power series $q_\lambda(\mathbf{z})$ is homogeneous of degree $|\lambda|$. Similarly, the power series $q_\mu(\mathbf{x})$ is homogeneous of degree $|\mu|$. Similarly, the power series $q_\nu(\mathbf{y})$ is homogeneous of degree $|\nu|$.

Now, we know that the three power series $q_\mu(\mathbf{x}), q_\nu(\mathbf{y}), q_\lambda(\mathbf{z})$ are homogeneous of degrees $|\mu|, |\nu|, |\lambda|$, respectively. Hence, their product $q_\mu(\mathbf{x}) q_\nu(\mathbf{y}) q_\lambda(\mathbf{z})$ is homogeneous of degree $|\mu| + |\nu| + |\lambda|$. This proves (13.54.14).

[585] *Proof.* If $(\mu,\nu,\lambda) \in \operatorname{Par}^3$ satisfies $|\mu| + |\nu| + |\lambda| = n$, then each of the three partitions μ,ν,λ has to have size $\leq n$ (because $|\mu| \leq |\mu| + |\nu| + |\lambda| = n$, and similarly $|\nu| \leq n$ and $|\lambda| \leq n$). This leaves only finitely many possibilities for each of these partitions μ,ν,λ. Thus, there are only finitely many $(\mu,\nu,\lambda) \in \operatorname{Par}^3$ satisfying $|\mu| + |\nu| + |\lambda| = n$.

[586] *Proof of (13.54.17):* Let $(\mu,\nu,\lambda) \in \operatorname{Par}^3$ be such that $|\mu| + |\nu| + |\lambda| = n$. From (13.54.14), we know that the power series $q_\mu(\mathbf{x}) q_\nu(\mathbf{y}) q_\lambda(\mathbf{z})$ is homogeneous of degree $|\mu| + |\nu| + |\lambda| = n$. Hence, (13.54.15) (applied to $f = q_\mu(\mathbf{x}) q_\nu(\mathbf{y}) q_\lambda(\mathbf{z})$) yields $\pi_n(q_\mu(\mathbf{x}) q_\nu(\mathbf{y}) q_\lambda(\mathbf{z})) = q_\mu(\mathbf{x}) q_\nu(\mathbf{y}) q_\lambda(\mathbf{z})$. This proves (13.54.17).

[587] *Proof of (13.54.18):* Let $(\mu,\nu,\lambda) \in \operatorname{Par}^3$ be such that $|\mu| + |\nu| + |\lambda| \neq n$. From (13.54.14), we know that the power series $q_\mu(\mathbf{x}) q_\nu(\mathbf{y}) q_\lambda(\mathbf{z})$ is homogeneous of degree $|\mu| + |\nu| + |\lambda| \neq n$. Hence, (13.54.16) (applied to $f = q_\mu(\mathbf{x}) q_\nu(\mathbf{y}) q_\lambda(\mathbf{z})$) yields $\pi_n(q_\mu(\mathbf{x}) q_\nu(\mathbf{y}) q_\lambda(\mathbf{z})) = 0$. This proves (13.54.18).

But the map π_n is $\mathbf{k}$-linear and continuous. Thus, it respects infinite sums. Hence,

$$\begin{aligned}
&\pi_n \left(\sum_{(\mu,\nu,\lambda)\in \operatorname{Par}^3} a_{\lambda,\mu,\nu} q_\mu (\mathbf{x}) q_\nu (\mathbf{y}) q_\lambda (\mathbf{z}) \right) \\
&= \sum_{(\mu,\nu,\lambda)\in \operatorname{Par}^3} a_{\lambda,\mu,\nu} \pi_n (q_\mu (\mathbf{x}) q_\nu (\mathbf{y}) q_\lambda (\mathbf{z})) \\
&= \sum_{\substack{(\mu,\nu,\lambda)\in \operatorname{Par}^3; \\ |\mu|+|\nu|+|\lambda|=n}} a_{\lambda,\mu,\nu} \underbrace{\pi_n (q_\mu (\mathbf{x}) q_\nu (\mathbf{y}) q_\lambda (\mathbf{z}))}_{\substack{=q_\mu (\mathbf{x}) q_\nu (\mathbf{y}) q_\lambda (\mathbf{z}) \\ \text{(by (13.54.17))}}} + \sum_{\substack{(\mu,\nu,\lambda)\in \operatorname{Par}^3; \\ |\mu|+|\nu|+|\lambda|\neq n}} a_{\lambda,\mu,\nu} \underbrace{\pi_n (q_\mu (\mathbf{x}) q_\nu (\mathbf{y}) q_\lambda (\mathbf{z}))}_{\substack{=0 \\ \text{(by (13.54.18))}}} \\
&= \sum_{\substack{(\mu,\nu,\lambda)\in \operatorname{Par}^3; \\ |\mu|+|\nu|+|\lambda|=n}} a_{\lambda,\mu,\nu} q_\mu (\mathbf{x}) q_\nu (\mathbf{y}) q_\lambda (\mathbf{z}) + \underbrace{\sum_{\substack{(\mu,\nu,\lambda)\in \operatorname{Par}^3; \\ |\mu|+|\nu|+|\lambda|\neq n}} a_{\lambda,\mu,\nu} 0}_{=0} \\
&= \sum_{\substack{(\mu,\nu,\lambda)\in \operatorname{Par}^3; \\ |\mu|+|\nu|+|\lambda|=n}} a_{\lambda,\mu,\nu} q_\mu (\mathbf{x}) q_\nu (\mathbf{y}) q_\lambda (\mathbf{z}) . \qquad (13.54.19)
\end{aligned}$$

The same argument (applied to the family $(b_{\lambda,\mu,\nu})_{(\mu,\nu,\lambda)\in \operatorname{Par}^3}$ instead of $(a_{\lambda,\mu,\nu})_{(\mu,\nu,\lambda)\in \operatorname{Par}^3}$) yields

$$\pi_n \left(\sum_{(\mu,\nu,\lambda)\in \operatorname{Par}^3} b_{\lambda,\mu,\nu} q_\mu (\mathbf{x}) q_\nu (\mathbf{y}) q_\lambda (\mathbf{z}) \right) = \sum_{\substack{(\mu,\nu,\lambda)\in \operatorname{Par}^3; \\ |\mu|+|\nu|+|\lambda|=n}} b_{\lambda,\mu,\nu} q_\mu (\mathbf{x}) q_\nu (\mathbf{y}) q_\lambda (\mathbf{z}) . \qquad (13.54.20)$$

Now, applying the map π_n to both sides of the equality (2.5.5), we obtain

$$\begin{aligned}
&\pi_n \left(\sum_{(\mu,\nu,\lambda)\in \operatorname{Par}^3} a_{\lambda,\mu,\nu} q_\mu (\mathbf{x}) q_\nu (\mathbf{y}) q_\lambda (\mathbf{z}) \right) \\
&= \pi_n \left(\sum_{(\mu,\nu,\lambda)\in \operatorname{Par}^3} b_{\lambda,\mu,\nu} q_\mu (\mathbf{x}) q_\nu (\mathbf{y}) q_\lambda (\mathbf{z}) \right) = \sum_{\substack{(\mu,\nu,\lambda)\in \operatorname{Par}^3; \\ |\mu|+|\nu|+|\lambda|=n}} b_{\lambda,\mu,\nu} q_\mu (\mathbf{x}) q_\nu (\mathbf{y}) q_\lambda (\mathbf{z}) .
\end{aligned}$$

Comparing this with (13.54.19), we obtain

$$\sum_{\substack{(\mu,\nu,\lambda)\in \operatorname{Par}^3; \\ |\mu|+|\nu|+|\lambda|=n}} a_{\lambda,\mu,\nu} q_\mu (\mathbf{x}) q_\nu (\mathbf{y}) q_\lambda (\mathbf{z}) = \sum_{\substack{(\mu,\nu,\lambda)\in \operatorname{Par}^3; \\ |\mu|+|\nu|+|\lambda|=n}} b_{\lambda,\mu,\nu} q_\mu (\mathbf{x}) q_\nu (\mathbf{y}) q_\lambda (\mathbf{z}) . \qquad (13.54.21)$$

Notice that both sums appearing in this equality are finite (since there are only finitely many $(\mu, \nu, \lambda) \in \operatorname{Par}^3$ satisfying $|\mu| + |\nu| + |\lambda| = n$). Since the family $(q_\mu (\mathbf{x}) q_\nu (\mathbf{y}) q_\lambda (\mathbf{z}))_{(\mu,\nu,\lambda)\in \operatorname{Par}^3;\ |\mu|+|\nu|+|\lambda|=n}$ is $\mathbf{k}$-linearly independent[588], we can thus conclude from (13.54.21) that

$$a_{\lambda,\mu,\nu} = b_{\lambda,\mu,\nu} \qquad \text{for each } (\mu, \nu, \lambda) \in \operatorname{Par}^3 \text{ satisfying } |\mu| + |\nu| + |\lambda| = n. \qquad (13.54.22)$$

Now, forget that we fixed n. We thus have proven (13.54.22) for each $n \in \mathbb{N}$.

Now, let $(\mu, \nu, \lambda) \in \operatorname{Par}^3$ be arbitrary. Then, $|\mu| + |\nu| + |\lambda| \in \mathbb{N}$. Hence, (13.54.22) (applied to $n = |\mu| + |\nu| + |\lambda|$) yields $a_{\lambda,\mu,\nu} = b_{\lambda,\mu,\nu}$.

Now, forget that we fixed (μ, ν, λ). We thus have proven that $a_{\lambda,\mu,\nu} = b_{\lambda,\mu,\nu}$ for each $(\mu, \nu, \lambda) \in \operatorname{Par}^3$. In other words, $(a_{\lambda,\mu,\nu})_{(\mu,\nu,\lambda)\in \operatorname{Par}^3} = (b_{\lambda,\mu,\nu})_{(\mu,\nu,\lambda)\in \operatorname{Par}^3}$. This solves Exercise 2.5.5 (c).

(b) The solution to Exercise 2.5.5 (b) is analogous to that of Exercise 2.5.5 (c) (the difference being that there are now just two families $\mathbf{x}$ and $\mathbf{y}$ of indeterminates, rather than three families $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{z}$).

---

[588] *Proof.* We know that the family $(q_\mu (\mathbf{x}) q_\nu (\mathbf{y}) q_\lambda (\mathbf{z}))_{(\mu,\nu,\lambda)\in \operatorname{Par}^3}$ is $\mathbf{k}$-linearly independent. Hence, the family $(q_\mu (\mathbf{x}) q_\nu (\mathbf{y}) q_\lambda (\mathbf{z}))_{(\mu,\nu,\lambda)\in \operatorname{Par}^3;\ |\mu|+|\nu|+|\lambda|=n}$ (being a subfamily of it) must also be $\mathbf{k}$-linearly independent.

13.55. **Solution to Exercise 2.5.10.** *Solution to Exercise 2.5.10.* Recall that $(s_\kappa)_{\kappa \in \operatorname{Par}}$ is a basis of the $\mathbf{k}$-module Λ. Hence, $(s_\kappa \otimes s_\lambda)_{(\kappa,\lambda)\in \operatorname{Par}\times\operatorname{Par}}$ is a basis of the $\mathbf{k}$-module $\Lambda\otimes\Lambda$. We will refer to this basis as the "Schur basis" of $\Lambda\otimes\Lambda$.

Now, the first diagram in (1.3.4) commutes when A is set to Λ (since Λ is a bialgebra). In other words,

$$\Delta\circ m = (m\otimes m)\circ(\operatorname{id}\otimes T\otimes\operatorname{id})\circ(\Delta\otimes\Delta) : \Lambda\otimes\Lambda\to\Lambda\otimes\Lambda. \tag{13.55.1}$$

Fix four partitions φ, ψ, κ and λ. Applying both sides of the equality (13.55.1) to $s_\varphi\otimes s_\psi$, we obtain

$$(\Delta\circ m)(s_\varphi\otimes s_\psi) = ((m\otimes m)\circ(\operatorname{id}\otimes T\otimes\operatorname{id})\circ(\Delta\otimes\Delta))(s_\varphi\otimes s_\psi).$$

Comparing the coefficients before $s_\kappa\otimes s_\lambda$ in this equality, we obtain

$$\sum_{\rho\in\operatorname{Par}} c^\rho_{\kappa,\lambda} c^\rho_{\varphi,\psi} = \sum_{(\alpha,\beta,\gamma,\delta)\in\operatorname{Par}^4} c^\kappa_{\alpha,\gamma} c^\lambda_{\beta,\delta} c^\varphi_{\alpha,\beta} c^\psi_{\gamma,\delta}$$

(after a straightforward computation using (2.5.6), (2.5.7) and Corollary 2.5.7). This solves the exercise.

13.56. **Solution to Exercise 2.5.11.** *Solution to Exercise 2.5.11.* Before we step to the solution of this problem, let us make some general observations.

- Every two partitions λ and μ satisfy $s_{\lambda/\mu} = \sum_\nu c^\lambda_{\mu,\nu} s_\nu$, where the sum ranges over all partitions ν (according to Remark 2.5.9). In other words, every two partitions λ and μ satisfy

$$s_{\lambda/\mu} = \sum_{\nu\in\operatorname{Par}} c^\lambda_{\mu,\nu} s_\nu. \tag{13.56.1}$$

- On the other hand, (2.5.6) yields

$$s_\mu s_\nu = \sum_{\lambda\in\operatorname{Par}} c^\lambda_{\mu,\nu} s_\lambda \qquad \text{for any two partitions } \mu \text{ and } \nu. \tag{13.56.2}$$

(a) Let μ be a partition. We have

$$\sum_{\lambda\in\operatorname{Par}} s_\lambda(\mathbf{x}) \underbrace{s_{\lambda/\mu}(\mathbf{y})}_{\substack{=\sum_{\nu\in\operatorname{Par}} c^\lambda_{\mu,\nu} s_\nu(\mathbf{y}) \\ \text{(by (13.56.1), evaluated} \\ \text{at the variable set } \mathbf{y})}} = \sum_{\lambda\in\operatorname{Par}} s_\lambda(\mathbf{x}) \sum_{\nu\in\operatorname{Par}} c^\lambda_{\mu,\nu} s_\nu(\mathbf{y}). \tag{13.56.3}$$

On the other hand, (2.5.1) yields

$$\prod_{i,j=1}^\infty (1-x_iy_j)^{-1} = \sum_{\lambda\in\operatorname{Par}} s_\lambda(\mathbf{x}) s_\lambda(\mathbf{y}) = \sum_{\nu\in\operatorname{Par}} s_\nu(\mathbf{x}) s_\nu(\mathbf{y})$$

(here, we renamed the summation index λ as ν). Multiplying this equality with $s_\mu(\mathbf{x})$, we obtain

$$\begin{aligned} s_\mu(\mathbf{x})\cdot\prod_{i,j=1}^\infty (1-x_iy_j)^{-1} &= s_\mu(\mathbf{x})\cdot\sum_{\nu\in\operatorname{Par}} s_\nu(\mathbf{x}) s_\nu(\mathbf{y}) = \sum_{\nu\in\operatorname{Par}} \underbrace{s_\mu(\mathbf{x}) s_\nu(\mathbf{x})}_{\substack{=\sum_{\lambda\in\operatorname{Par}} c^\lambda_{\mu,\nu} s_\lambda(\mathbf{x}) \\ \text{(by (13.56.2))}}} s_\nu(\mathbf{y}) \\ &= \sum_{\nu\in\operatorname{Par}} \left(\sum_{\lambda\in\operatorname{Par}} c^\lambda_{\mu,\nu} s_\lambda(\mathbf{x})\right) s_\nu(\mathbf{y}) = \sum_{\lambda\in\operatorname{Par}} c^\lambda_{\mu,\nu} s_\lambda(\mathbf{x}) \sum_{\nu\in\operatorname{Par}} s_\nu(\mathbf{y}) \\ &= \sum_{\lambda\in\operatorname{Par}} s_\lambda(\mathbf{x}) \sum_{\nu\in\operatorname{Par}} c^\lambda_{\mu,\nu} s_\nu(\mathbf{y}). \end{aligned}$$

Compared with (13.56.3), this yields

$$\sum_{\lambda\in\operatorname{Par}} s_\lambda(\mathbf{x}) s_{\lambda/\mu}(\mathbf{y}) = s_\mu(\mathbf{x})\cdot\prod_{i,j=1}^\infty (1-x_iy_j)^{-1}.$$

This solves Exercise 2.5.11(a).

(b) Let $\mathbf{k}[[\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{w}]]$ denote the ring

$$\mathbf{k}[[x_1,x_2,x_3,...,y_1,y_2,y_3,...,z_1,z_2,z_3,...,w_1,w_2,w_3,...]].$$

This ring clearly contains $\mathbf{k}[[\mathbf{x},\mathbf{y}]]$ as a subring. We will use the obvious abbreviations for variable sets: $\mathbf{x}=(x_1,x_2,x_3,...)$, $(\mathbf{x},\mathbf{z})=(x_1,x_2,x_3,...,z_1,z_2,z_3,...)$, etc.

Every partition λ satisfies

$$\begin{aligned} s_\lambda(\mathbf{y},\mathbf{x}) &= s_\lambda(\mathbf{x},\mathbf{y}) = \sum_{\mu\subseteq\lambda} s_\mu(\mathbf{x})\,s_{\lambda/\mu}(\mathbf{y}) \qquad \text{(by (2.3.3))} \\ &= \sum_{\mu\in\mathrm{Par}} s_\mu(\mathbf{x})\,s_{\lambda/\mu}(\mathbf{y}) - \sum_{\substack{\mu\in\mathrm{Par};\\ \mu\not\subseteq\lambda}} s_\mu(\mathbf{x})\underbrace{s_{\lambda/\mu}(\mathbf{y})}_{\substack{=0\\ (\text{since } \mu\not\subseteq\lambda)}} = \sum_{\mu\in\mathrm{Par}} s_\mu(\mathbf{x})\,s_{\lambda/\mu}(\mathbf{y}) - \underbrace{\sum_{\substack{\mu\in\mathrm{Par};\\ \mu\not\subseteq\lambda}} s_\mu(\mathbf{x})\,0}_{=0} \\ &= \sum_{\mu\in\mathrm{Par}} s_\mu(\mathbf{x})\,s_{\lambda/\mu}(\mathbf{y}). \end{aligned}$$

Applying this equality to the variables $\mathbf{z}$ and $\mathbf{x}$ instead of $\mathbf{x}$ and $\mathbf{y}$, we obtain

$$s_\lambda(\mathbf{x},\mathbf{z}) = \sum_{\mu\in\mathrm{Par}} s_\mu(\mathbf{z})\,s_{\lambda/\mu}(\mathbf{x}). \tag{13.56.4}$$

Applying this equality to the variables $\mathbf{y}$ and $\mathbf{w}$ instead of $\mathbf{x}$ and $\mathbf{z}$, we obtain

$$s_\lambda(\mathbf{y},\mathbf{w}) = \sum_{\mu\in\mathrm{Par}} s_\mu(\mathbf{w})\,s_{\lambda/\mu}(\mathbf{y}) = \sum_{\nu\in\mathrm{Par}} s_\nu(\mathbf{w})\,s_{\lambda/\nu}(\mathbf{y}) \tag{13.56.5}$$

(here, we renamed the summation index μ as ν) for any partition λ.

Now,

$$\begin{aligned} &\prod_{i,j=1}^{\infty}(1-x_iy_j)^{-1}\prod_{i,j=1}^{\infty}(1-x_iw_j)^{-1}\prod_{i,j=1}^{\infty}(1-z_iy_j)^{-1}\prod_{i,j=1}^{\infty}(1-z_iw_j)^{-1} \\ &= \prod_{\substack{a\in(\mathbf{x},\mathbf{z});\\ b\in(\mathbf{y},\mathbf{w})}}(1-ab)^{-1} = \sum_{\lambda\in\mathrm{Par}} \underbrace{s_\lambda(\mathbf{x},\mathbf{z})}_{\substack{=\sum_{\mu\in\mathrm{Par}} s_\mu(\mathbf{z})s_{\lambda/\mu}(\mathbf{x})\\ (\text{by (13.56.4)})}} \underbrace{s_\lambda(\mathbf{y},\mathbf{w})}_{\substack{=\sum_{\nu\in\mathrm{Par}} s_\nu(\mathbf{w})s_{\lambda/\nu}(\mathbf{y})\\ (\text{by (13.56.5)})}} \\ &\qquad \text{(by (2.5.1), applied to the variable sets } (\mathbf{x},\mathbf{z}) \text{ and } (\mathbf{y},\mathbf{w}) \text{ instead of } \mathbf{x} \text{ and } \mathbf{y}) \\ &= \sum_{\lambda\in\mathrm{Par}}\left(\sum_{\mu\in\mathrm{Par}} s_\mu(\mathbf{z})\,s_{\lambda/\mu}(\mathbf{x})\right)\left(\sum_{\nu\in\mathrm{Par}} s_\nu(\mathbf{w})\,s_{\lambda/\nu}(\mathbf{y})\right) \\ &= \sum_{\mu\in\mathrm{Par}}\sum_{\nu\in\mathrm{Par}} s_\mu(\mathbf{z})\,s_\nu(\mathbf{w})\left(\sum_{\lambda\in\mathrm{Par}} s_{\lambda/\mu}(\mathbf{x})\,s_{\lambda/\nu}(\mathbf{y})\right). \qquad (13.56.6) \end{aligned}$$

On the other hand, every partition κ satisfies

$$\sum_{\lambda\in\mathrm{Par}} s_\lambda(\mathbf{x})\,s_{\lambda/\kappa}(\mathbf{y}) = s_\kappa(\mathbf{x})\cdot\prod_{i,j=1}^{\infty}(1-x_iy_j)^{-1} \tag{13.56.7}$$

(by Exercise 2.5.11(a), applied to κ instead of μ). Applying this to the variable sets $\mathbf{w}$ and $\mathbf{x}$ instead of $\mathbf{x}$ and $\mathbf{y}$, we obtain

$$\begin{aligned}
\sum_{\lambda\in\mathrm{Par}} s_{\lambda}(\mathbf{w})\, s_{\lambda/\kappa}(\mathbf{x}) &= s_{\kappa}(\mathbf{w}) \cdot \underbrace{\prod_{i,j=1}^{\infty} (1-w_i x_j)^{-1}}_{\substack{=\prod_{i,j=1}^{\infty}(1-x_j w_i)^{-1} \\ =\prod_{j,i=1}^{\infty}(1-x_i w_j)^{-1} \\ \text{(here, we renamed the index (i,j) as (j,i))}}} \\
&= s_{\kappa}(\mathbf{w}) \cdot \underbrace{\prod_{j,i=1}^{\infty}}_{=\prod_{i,j=1}^{\infty}} (1-x_i w_j)^{-1} = s_{\kappa}(\mathbf{w}) \cdot \prod_{i,j=1}^{\infty} (1-x_i w_j)^{-1} \qquad (13.56.8)
\end{aligned}$$

for every partition κ. But applying both sides of the identity (13.56.7) to the variable set $\mathbf{z}$ instead of $\mathbf{x}$, and renaming the summation index λ as μ on the left hand side of this equality, we obtain

$$\sum_{\mu\in\mathrm{Par}} s_{\mu}(\mathbf{z})\, s_{\mu/\kappa}(\mathbf{y}) = s_{\kappa}(\mathbf{z}) \cdot \prod_{i,j=1}^{\infty} (1-z_i y_j)^{-1}. \tag{13.56.9}$$

Now, applying (2.5.1) to the variable sets $\mathbf{z}$ and $\mathbf{w}$ instead of $\mathbf{x}$ and $\mathbf{y}$, we obtain

$$\prod_{i,j=1}^{\infty} (1-z_i w_j)^{-1} = \sum_{\lambda\in\mathrm{Par}} s_{\lambda}(\mathbf{z})\, s_{\lambda}(\mathbf{w}) = \sum_{\kappa\in\mathrm{Par}} s_{\kappa}(\mathbf{z})\, s_{\kappa}(\mathbf{w}).$$

Hence,

$$\begin{aligned}
&\prod_{i,j=1}^{\infty} (1-x_i y_j)^{-1} \prod_{i,j=1}^{\infty} (1-x_i w_j)^{-1} \prod_{i,j=1}^{\infty} (1-z_i y_j)^{-1} \underbrace{\prod_{i,j=1}^{\infty} (1-z_i w_j)^{-1}}_{=\sum_{\kappa\in\mathrm{Par}} s_{\kappa}(\mathbf{z}) s_{\kappa}(\mathbf{w})} \\
&= \prod_{i,j=1}^{\infty} (1-x_i y_j)^{-1} \prod_{i,j=1}^{\infty} (1-x_i w_j)^{-1} \prod_{i,j=1}^{\infty} (1-z_i y_j)^{-1} \cdot \sum_{\kappa\in\mathrm{Par}} s_{\kappa}(\mathbf{z})\, s_{\kappa}(\mathbf{w}) \\
&= \prod_{i,j=1}^{\infty} (1-x_i y_j)^{-1} \cdot \sum_{\kappa\in\mathrm{Par}} \underbrace{\left(s_{\kappa}(\mathbf{w}) \cdot \prod_{i,j=1}^{\infty} (1-x_i w_j)^{-1} \right)}_{\substack{=\sum_{\lambda\in\mathrm{Par}} s_{\lambda}(\mathbf{w}) s_{\lambda/\kappa}(\mathbf{x}) \\ \text{(by (13.56.8))}}} \underbrace{\left(s_{\kappa}(\mathbf{z}) \cdot \prod_{i,j=1}^{\infty} (1-z_i y_j)^{-1} \right)}_{\substack{=\sum_{\mu\in\mathrm{Par}} s_{\mu}(\mathbf{z}) s_{\mu/\kappa}(\mathbf{y}) \\ \text{(by (13.56.9))}}} \\
&= \prod_{i,j=1}^{\infty} (1-x_i y_j)^{-1} \cdot \sum_{\kappa\in\mathrm{Par}} \left(\sum_{\lambda\in\mathrm{Par}} s_{\lambda}(\mathbf{w})\, s_{\lambda/\kappa}(\mathbf{x}) \right) \left(\sum_{\mu\in\mathrm{Par}} s_{\mu}(\mathbf{z})\, s_{\mu/\kappa}(\mathbf{y}) \right) \\
&= \sum_{\mu\in\mathrm{Par}} \sum_{\lambda\in\mathrm{Par}} s_{\mu}(\mathbf{z})\, s_{\lambda}(\mathbf{w}) \left(\sum_{\kappa\in\mathrm{Par}} s_{\lambda/\kappa}(\mathbf{x})\, s_{\mu/\kappa}(\mathbf{y}) \right) \cdot \prod_{i,j=1}^{\infty} (1-x_i y_j)^{-1} \\
&= \sum_{\mu\in\mathrm{Par}} \sum_{\nu\in\mathrm{Par}} s_{\mu}(\mathbf{z})\, s_{\nu}(\mathbf{w}) \left(\sum_{\rho\in\mathrm{Par}} s_{\nu/\rho}(\mathbf{x})\, s_{\mu/\rho}(\mathbf{y}) \right) \cdot \prod_{i,j=1}^{\infty} (1-x_i y_j)^{-1}
\end{aligned}$$

(here, we renamed the summation indices λ and κ as ν and ρ, respectively). Comparing this with (13.56.6), we obtain

$$\sum_{\mu\in\operatorname{Par}}\sum_{\nu\in\operatorname{Par}} s_\mu(\mathbf{z})\, s_\nu(\mathbf{w}) \left(\sum_{\lambda\in\operatorname{Par}} s_{\lambda/\mu}(\mathbf{x})\, s_{\lambda/\nu}(\mathbf{y})\right)$$
$$= \sum_{\mu\in\operatorname{Par}}\sum_{\nu\in\operatorname{Par}} s_\mu(\mathbf{z})\, s_\nu(\mathbf{w}) \left(\sum_{\rho\in\operatorname{Par}} s_{\nu/\rho}(\mathbf{x})\, s_{\mu/\rho}(\mathbf{y})\right) \cdot \prod_{i,j=1}^{\infty} (1 - x_i y_j)^{-1}.$$

We can regard this as an identity in the ring $(\mathbf{k}[[\mathbf{x},\mathbf{y}]])[[\mathbf{z},\mathbf{w}]]$ of formal power series in the variables $(\mathbf{z},\mathbf{w}) = (z_1, z_2, z_3, ..., w_1, w_2, w_3, ...)$ over the ring $\mathbf{k}[[\mathbf{x},\mathbf{y}]]$. Extracting the coefficients in front of $s_\alpha(\mathbf{z})\, s_\beta(\mathbf{w})$ in this identity[589], we obtain

$$\sum_{\lambda\in\operatorname{Par}} s_{\lambda/\alpha}(\mathbf{x})\, s_{\lambda/\beta}(\mathbf{y}) = \left(\sum_{\rho\in\operatorname{Par}} s_{\beta/\rho}(\mathbf{x})\, s_{\alpha/\rho}(\mathbf{y})\right) \cdot \prod_{i,j=1}^{\infty} (1 - x_i y_j)^{-1}.$$

This solves Exercise 2.5.11(b).

13.57. **Solution to Exercise 2.5.13.** *Solution to Exercise 2.5.13.* We know that $(s_\lambda)_{\lambda\in\operatorname{Par}}$ is a basis of the $\mathbf{k}$-module Λ, but we can also say something more specific: For every $n \in \mathbb{N}$, the family $(s_\lambda)_{\lambda\in\operatorname{Par}_n}$ is a basis of the $\mathbf{k}$-module Λ_n [590].

The basis $(s_\lambda)_{\lambda\in\operatorname{Par}}$ of Λ is orthonormal with respect to the Hall inner product. In other words,

$$(s_\lambda, s_\mu) = \delta_{\lambda,\mu} \qquad \text{for any partitions } \lambda \text{ and } \mu. \tag{13.57.1}$$

(a) Let n and m be two distinct nonnegative integers. Let $f \in \Lambda_n$ and $g \in \Lambda_m$.

We need to prove that $(f, g) = 0$. Since this equality is $\mathbf{k}$-linear in f, we can WLOG assume that f is an element of the basis $(s_\lambda)_{\lambda\in\operatorname{Par}_n}$ of Λ_n. In other words, we can WLOG assume that $f = s_\lambda$ for some $\lambda \in \operatorname{Par}_n$. Assume this, and similarly assume that $g = s_\mu$ for some $\mu \in \operatorname{Par}_m$. These two partitions λ and μ must be

[589] This notion of "extracting the coefficients" relies on the fact that if two families $(a_{\mu,\nu})_{(\mu,\nu)\in\operatorname{Par}^2} \in \mathbf{k}[[\mathbf{x},\mathbf{y}]]^{\operatorname{Par}^2}$ and $(b_{\mu,\nu})_{(\mu,\nu)\in\operatorname{Par}^2} \in \mathbf{k}[[\mathbf{x},\mathbf{y}]]^{\operatorname{Par}^2}$ satisfy

$$\sum_{(\mu,\nu)\in\operatorname{Par}^2} a_{\mu,\nu} s_\mu(\mathbf{z}) s_\nu(\mathbf{w}) = \sum_{(\mu,\nu)\in\operatorname{Par}^2} b_{\mu,\nu} s_\mu(\mathbf{z}) s_\nu(\mathbf{w})$$

in $(\mathbf{k}[[\mathbf{x},\mathbf{y}]])[[\mathbf{z},\mathbf{w}]]$, then $(a_{\mu,\nu})_{(\mu,\nu)\in\operatorname{Par}^2} = (b_{\mu,\nu})_{(\mu,\nu)\in\operatorname{Par}^2}$. This fact is a consequence of Exercise 2.5.5(b), applied to $q_\lambda = s_\lambda$ (with the base ring $\mathbf{k}$ replaced by $\mathbf{k}[[\mathbf{x},\mathbf{y}]]$, and with the families $\mathbf{x}$ and $\mathbf{y}$ renamed as $\mathbf{z}$ and $\mathbf{w}$).

[590] *Proof.* We have $s_\lambda \in \Lambda_{|\lambda|} = \Lambda_n$ for every $\lambda \in \operatorname{Par}_n$. Thus, $(s_\lambda)_{\lambda\in\operatorname{Par}_n}$ is a family of elements of Λ_n. This family is $\mathbf{k}$-linearly independent (because it is part of the basis $(s_\lambda)_{\lambda\in\operatorname{Par}}$ of Λ). It remains to prove that this family spans the $\mathbf{k}$-module Λ_n.

Let $f \in \Lambda_n$. We can write f in the form $f = \sum_{\lambda\in\operatorname{Par}} a_\lambda s_\lambda$ for some family $(a_\lambda)_{\lambda\in\operatorname{Par}} \in \mathbf{k}^{\operatorname{Par}}$ of scalars. Consider this $(a_\lambda)_{\lambda\in\operatorname{Par}}$ and notice that

$$f = \sum_{\lambda\in\operatorname{Par}} a_\lambda s_\lambda = \sum_{\lambda\in\operatorname{Par}_n} a_\lambda s_\lambda + \sum_{\substack{\lambda\in\operatorname{Par};\\ \lambda\notin\operatorname{Par}_n}} a_\lambda s_\lambda,$$

so that $f - \sum_{\lambda\in\operatorname{Par}_n} a_\lambda s_\lambda = \sum_{\substack{\lambda\in\operatorname{Par};\\ \lambda\notin\operatorname{Par}_n}} a_\lambda s_\lambda$. The left hand side of this latter equality is homogeneous of degree n (since f and all the s_λ with $\lambda \in \operatorname{Par}_n$ are homogeneous of degree n), while the right hand side is a sum of homogeneous elements of degrees different from n. So the only way these two sides can be equal is if they both are 0. In particular, this shows that the left hand side is 0. In other words, $f - \sum_{\lambda\in\operatorname{Par}_n} a_\lambda s_\lambda = 0$, so that $f = \sum_{\lambda\in\operatorname{Par}_n} a_\lambda s_\lambda$. Hence, f is a $\mathbf{k}$-linear combination of the s_λ for $\lambda \in \operatorname{Par}_n$.

Since we have proven this for every $f \in \Lambda_n$, we thus conclude that the family $(s_\lambda)_{\lambda\in\operatorname{Par}_n}$ spans the $\mathbf{k}$-module Λ_n, qed.

distinct (because their sizes differ: $|\lambda| = n \neq m = |\mu|$), and so they satisfy $\delta_{\lambda,\mu} = 0$. Now,

$$\begin{aligned}\Bigg(\underbrace{f}_{=s_\lambda}, \underbrace{g}_{=s_\mu}\Bigg) &= (s_\lambda, s_\mu) = \delta_{\lambda,\mu} \qquad\qquad (\text{by } (13.57.1)) \\ &= 0.\end{aligned}$$

This solves Exercise 2.5.13(a).

(b) Let $n \in \mathbb{N}$ and $f \in \Lambda_n$.

We need to prove that $(h_n, f) = f(1)$. Since this equality is $\mathbf{k}$-linear in f, we can WLOG assume that f is an element of the basis $(s_\lambda)_{\lambda \in \operatorname{Par}_n}$ of Λ_n. In other words, we can WLOG assume that $f = s_\lambda$ for some $\lambda \in \operatorname{Par}_n$. Assume this.

We must be in one of the two cases:

Case 1: We have $\ell(\lambda) \leq 1$.

Case 2: We have $\ell(\lambda) > 1$.

Let us consider Case 1 first. In this case, $\ell(\lambda) \leq 1$. Hence, $\lambda = (n)$ (since $\lambda \in \operatorname{Par}_n$), so that $s_\lambda = s_{(n)}$. Hence, $f = s_\lambda = s_{(n)} = h_n$ and thus $f(1) = h_n(1) = 1$. Compared with

$$\begin{aligned}\Bigg(\underbrace{h_n}_{=s_{(n)}}, \underbrace{f}_{=s_{(n)}}\Bigg) &= \left(s_{(n)}, s_{(n)}\right) = \delta_{(n),(n)} \qquad\qquad (\text{by } (13.57.1)) \\ &= 1,\end{aligned}$$

this yields $(h_n, f) = f(1)$. Hence, $(h_n, f) = f(1)$ is proven in Case 1.

Let us now consider Case 2. In this case, $\ell(\lambda) > 1$, so the partition λ has more than 1 part. Exercise 2.3.8(b) (applied to 1 instead of n) thus yields $s_\lambda(x_1) = 0$. But $s_\lambda(1)$ can be seen as the result of substituting 1 for x_1 in $s_\lambda(x_1)$, and therefore must be 0 as well (since $s_\lambda(x_1) = 0$). Thus, we have $s_\lambda(1) = 0$. Since $f = s_\lambda$, we now have $f(1) = s_\lambda(1) = 0$. Compared with

$$\begin{aligned}\Bigg(\underbrace{h_n}_{=s_{(n)}}, \underbrace{f}_{=s_\lambda}\Bigg) &= \left(s_{(n)}, s_\lambda\right) = \delta_{(n),\lambda} \qquad\qquad (\text{by } (13.57.1)) \\ &= 0 \qquad\qquad (\text{since } (n) \neq \lambda \text{ (because } \ell((n)) = 1 < \ell(\lambda))),\end{aligned}$$

this yields $(h_n, f) = f(1)$. Hence, $(h_n, f) = f(1)$ is proven in Case 2.

Now, $(h_n, f) = f(1)$ is proven in both Cases, which solves Exercise 2.5.13(b).

(c) We know that the Hall inner product has an orthonormal basis (namely, $(s_\lambda)_{\lambda \in \operatorname{Par}}$). Thus, the Hall inner product is symmetric (since any bilinear form that has an orthonormal basis must be symmetric). In other words, $(f, g) = (g, f)$ for all $f \in \Lambda$ and $g \in \Lambda$. This solves Exercise 2.5.13(c).

13.58. **Solution to Exercise 2.5.18.** *Solution to Exercise 2.5.18.* (a) This is a well-known fact; for a proof, see Theorem 5.3 in Keith Conrad, *Universal Identities I*, http://www.math.uconn.edu/~kconrad/blurbs/. Another proof (more complicated, but with the advantage of proving a more general result) can be found in [86, proof of Corollary 0.2].

(b) Consider the endomorphism of the $\mathbf{k}$-module A defined by sending γ_i to β_i for every $i \in I$. This endomorphism is well-defined (since $(\gamma_i)_{i\in I}$ is a basis of A) and surjective (since the family $(\beta_i)_{i\in I}$ spans A), therefore is a $\mathbf{k}$-module isomorphism (according to Exercise 2.5.18(a)). As a consequence, it must send the basis $(\gamma_i)_{i\in I}$ of A to a basis of A (because a $\mathbf{k}$-module isomorphism sends any basis to a basis). Since it sends the basis $(\gamma_i)_{i\in I}$ to $(\beta_i)_{i\in I}$, this yields that $(\beta_i)_{i\in I}$ must be a basis of A. This solves Exercise 2.5.18(b).

13.59. Solution to Exercise 2.5.19.

Solution to Exercise 2.5.19. Fix $n \in \mathbb{N}$. Recall that $v_\lambda \in \Lambda_{|\lambda|}$ for each partition λ. Hence, for each $\lambda \in \operatorname{Par}_n$, we have $v_\lambda \in \Lambda_{|\lambda|} = \Lambda_n$ (since $|\lambda| = n$ (because $\lambda \in \operatorname{Par}_n$)). Thus, $(v_\lambda)_{\lambda \in \operatorname{Par}_n}$ is a family of elements of Λ_n.

But we have assumed that the family $(v_\lambda)_{\lambda \in \operatorname{Par}}$ spans the $\mathbf{k}$-module Λ. Since everything is graded, this yields that the family $(v_\lambda)_{\lambda \in \operatorname{Par}_n}$ spans the $\mathbf{k}$-module Λ_n [591].

Recall that $v_\lambda \in \Lambda_n$ for each $\lambda \in \operatorname{Par}_n$. Hence, we can define the $\mathbf{k}$-linear map

$$\beta : \Lambda_n \to \Lambda_n, \qquad m_\lambda \mapsto v_\lambda \qquad \text{for every } \lambda \in \operatorname{Par}_n$$

(since $(m_\lambda)_{\lambda \in \operatorname{Par}_n}$ is a basis of the $\mathbf{k}$-module Λ_n). Consider this map β. This map β is surjective (since the family $(v_\lambda)_{\lambda \in \operatorname{Par}_n}$ spans the $\mathbf{k}$-module Λ_n) [592].

But Λ_n is a finitely generated $\mathbf{k}$-module (since it has a finite basis $(m_\lambda)_{\lambda \in \operatorname{Par}_n}$). The map β is an endomorphism of this $\mathbf{k}$-module (since $\beta : \Lambda_n \to \Lambda_n$ is $\mathbf{k}$-linear), and is surjective. Thus, Exercise 2.5.18(a) (applied to this endomorphism) yields that β is a $\mathbf{k}$-module isomorphism.

Now, the map β satisfies $\beta(m_\lambda) = v_\lambda$ for each $\lambda \in \operatorname{Par}_n$ (by the definition of β). In other words, $(\beta(m_\lambda))_{\lambda \in \operatorname{Par}_n} = (v_\lambda)_{\lambda \in \operatorname{Par}_n}$.

But the family $(m_\lambda)_{\lambda \in \operatorname{Par}_n}$ is a basis of the $\mathbf{k}$-module Λ_n. Hence, the image of this family $(m_\lambda)_{\lambda \in \operatorname{Par}_n}$ under β is also a basis of the $\mathbf{k}$-module Λ_n (since β is a $\mathbf{k}$-module isomorphism). In other words, the

[591] *Proof.* Here is this argument in more detail:

Let $f \in \Lambda_n$. We shall prove that f is a $\mathbf{k}$-linear combination of the family $(v_\lambda)_{\lambda \in \operatorname{Par}_n}$.

We have $f \in \Lambda_n \subset \Lambda$. Thus, f is a $\mathbf{k}$-linear combination of the family $(v_\lambda)_{\lambda \in \operatorname{Par}}$ (since the family $(v_\lambda)_{\lambda \in \operatorname{Par}}$ spans the $\mathbf{k}$-module Λ). In other words, there exists a family $(c_\lambda)_{\lambda \in \operatorname{Par}} \in \mathbf{k}^{\operatorname{Par}}$ such that (all but finitely many $\lambda \in \operatorname{Par}$ satisfy $c_\lambda = 0$) and $f = \sum_{\lambda \in \operatorname{Par}} c_\lambda v_\lambda$. Consider this $(c_\lambda)_{\lambda \in \operatorname{Par}}$.

Let $\pi_n : \Lambda \to \Lambda_n$ be the projection from the graded $\mathbf{k}$-module Λ onto its n-th graded component. Then, π_n is a $\mathbf{k}$-linear map with the properties that

$$\pi_n(g) = g \qquad \text{for each } g \in \Lambda_n, \tag{13.59.1}$$

and

$$\pi_n(g) = 0 \qquad \text{for each } g \in \Lambda_m \text{ for each } m \in \mathbb{N} \text{ satisfying } m \neq n. \tag{13.59.2}$$

If $\lambda \in \operatorname{Par}$ satisfies $|\lambda| = n$, then $v_\lambda \in \Lambda_{|\lambda|} = \Lambda_n$ (since $|\lambda| = n$) and thus

$$\pi_n(v_\lambda) = v_\lambda \tag{13.59.3}$$

(by (13.59.1), applied to $g = v_\lambda$).

If $\lambda \in \operatorname{Par}$ satisfies $|\lambda| \neq n$, then $v_\lambda \in \Lambda_{|\lambda|}$ and thus

$$\pi_n(v_\lambda) = 0 \tag{13.59.4}$$

(by (13.59.2), applied to $g = v_\lambda$ and $m = |\lambda|$).

Applying (13.59.1) to $g = f$, we obtain $\pi_n(f) = f$, so that

$$\begin{aligned} f &= \pi_n\left(\underbrace{f}_{=\sum_{\lambda \in \operatorname{Par}} c_\lambda v_\lambda}\right) = \pi_n\left(\sum_{\lambda \in \operatorname{Par}} c_\lambda v_\lambda\right) = \sum_{\lambda \in \operatorname{Par}} c_\lambda \pi_n(v_\lambda) \qquad (\text{since } \pi_n \text{ is a } \mathbf{k}\text{-linear map}) \\ &= \sum_{\substack{\lambda \in \operatorname{Par}; \\ |\lambda| = n}} c_\lambda \underbrace{\pi_n(v_\lambda)}_{\substack{= v_\lambda \\ (\text{by } (13.59.3))}} + \sum_{\substack{\lambda \in \operatorname{Par}; \\ |\lambda| \neq n}} c_\lambda \underbrace{\pi_n(v_\lambda)}_{\substack{= 0 \\ (\text{by } (13.59.4))}} \\ &\qquad (\text{since each } \lambda \in \operatorname{Par} \text{ satisfies either } |\lambda| = n \text{ or } |\lambda| \neq n \text{ (but not both)}) \\ &= \sum_{\substack{\lambda \in \operatorname{Par}; \\ |\lambda| = n}} c_\lambda v_\lambda + \underbrace{\sum_{\substack{\lambda \in \operatorname{Par}; \\ |\lambda| \neq n}} c_\lambda 0}_{=0} = \underbrace{\sum_{\substack{\lambda \in \operatorname{Par}; \\ |\lambda| = n}}}_{=\sum_{\lambda \in \operatorname{Par}_n}} c_\lambda v_\lambda = \sum_{\lambda \in \operatorname{Par}_n} c_\lambda v_\lambda. \end{aligned}$$

This shows that f is a $\mathbf{k}$-linear combination of the family $(v_\lambda)_{\lambda \in \operatorname{Par}_n}$.

Now, forget that we fixed f. We thus have proven that every $f \in \Lambda_n$ is a $\mathbf{k}$-linear combination of the family $(v_\lambda)_{\lambda \in \operatorname{Par}_n}$. Therefore, the family $(v_\lambda)_{\lambda \in \operatorname{Par}_n}$ spans the $\mathbf{k}$-module Λ_n.

[592] *Proof.* In slightly more detail: The map β is $\mathbf{k}$-linear; thus, its image $\beta(\Lambda_n)$ is a $\mathbf{k}$-submodule of Λ_n. This $\mathbf{k}$-submodule $\beta(\Lambda_n)$ clearly contains v_λ for each $\lambda \in \operatorname{Par}_n$ (since the definition of β yields $v_\lambda = \beta(m_\lambda)$). Hence, $\beta(\Lambda_n)$ is a $\mathbf{k}$-submodule of Λ_n that contains v_λ for each $\lambda \in \operatorname{Par}_n$. But the only such $\mathbf{k}$-submodule is the whole module Λ_n (since the family $(v_\lambda)_{\lambda \in \operatorname{Par}_n}$ spans the $\mathbf{k}$-module Λ_n). Thus, we conclude that $\beta(\Lambda_n)$ must be the whole module Λ_n. In other words, β is surjective.

family $(v_\lambda)_{\lambda \in \mathrm{Par}_n}$ is a basis of the $\mathbf{k}$-module Λ_n (since the image of the family $(m_\lambda)_{\lambda \in \mathrm{Par}_n}$ under β is $(\beta(m_\lambda))_{\lambda \in \mathrm{Par}_n} = (v_\lambda)_{\lambda \in \mathrm{Par}_n}$).

Forget that we fixed n. We thus have proved that $(v_\lambda)_{\lambda \in \mathrm{Par}_n}$ is a basis of the $\mathbf{k}$-module Λ_n for each $n \in \mathbb{N}$. Hence, the family $(v_\lambda)_{\lambda \in \mathrm{Par}}$ is a graded basis of the $\mathbf{k}$-module $\bigoplus_{n \in \mathbb{N}} \Lambda_n$ (since $\mathrm{Par}_0, \mathrm{Par}_1, \mathrm{Par}_2, \ldots$ is a partition of the set Par). In other words, the family $(v_\lambda)_{\lambda \in \mathrm{Par}}$ is a graded basis of the $\mathbf{k}$-module Λ (since $\Lambda = \bigoplus_{n \in \mathbb{N}} \Lambda_n$). This solves Exercise 2.5.19.

13.60. **Solution to Exercise 2.5.20.** *Solution to Exercise 2.5.20.* (a) This can be solved by following the proof of Corollary 2.5.17, but in doing so one has to be careful about how one obtains the invertibility of A: it is no longer a consequence of A being a transition matrix between two bases (because we do not know in advance that $(u_\lambda)_{\lambda \in \mathrm{Par}}$ is a basis). Instead, one has to argue as in the footnote: The matrices A and B^t are block-diagonal, with each diagonal block corresponding to the partitions of size n for a given $n \in \mathbb{N}$ [593]. In particular, they are block-diagonal matrices with each block being a square matrix of finite size. It is known that if such a matrix is right-invertible, then it is left-invertible[594]; therefore, since A is right-invertible (because $AB^t = I$), we conclude that A is invertible.

(b) We know that for every partition λ, the symmetric functions h_λ and m_λ are two homogeneous elements of Λ, both of degree $|\lambda|$. We also know (from Proposition 2.5.15) that

$$\prod_{i,j=1}^{\infty} (1 - x_i y_j)^{-1} = \sum_{\lambda \in \mathrm{Par}} h_\lambda(\mathbf{x}) m_\lambda(\mathbf{y}).$$

Compared with (2.5.1), this yields $\sum_{\lambda \in \mathrm{Par}} s_\lambda(\mathbf{x}) s_\lambda(\mathbf{y}) = \sum_{\lambda \in \mathrm{Par}} h_\lambda(\mathbf{x}) m_\lambda(\mathbf{y})$. Thus, Exercise 2.5.20(a) (applied to $u_\lambda = h_\lambda$ and $v_\lambda = m_\lambda$) yields that $(h_\lambda)_{\lambda \in \mathrm{Par}}$ and $(m_\lambda)_{\lambda \in \mathrm{Par}}$ are $\mathbf{k}$-bases of Λ, and actually are dual bases with respect to the Hall inner product on Λ. In particular, $(h_\lambda)_{\lambda \in \mathrm{Par}}$ is a $\mathbf{k}$-basis of Λ. This solves Exercise 2.5.20(b).

Remark. We can use Exercise 2.5.20(b) to prove that $(e_\lambda)_{\lambda \in \mathrm{Par}}$ is a $\mathbf{k}$-basis of Λ as well (in a different way than we have done in the proof of Proposition 2.2.10). In fact, let us sketch this proof. We are going to proceed similarly to the above proof of Proposition 2.4.1, but with the roles of the families $(e_n)_{n \geq 1}$ and $(h_n)_{n \geq 1}$ switched.

We know (from Exercise 2.5.20(b)) that $(h_\lambda)_{\lambda \in \mathrm{Par}}$ is a $\mathbf{k}$-basis of Λ. In other words, the family $(h_n)_{n \geq 1}$ is algebraically independent and generates the $\mathbf{k}$-algebra Λ (by parts (b) and (c) of Exercise 2.2.14, applied to $v_n = h_n$ and $v_\lambda = h_\lambda$). Thus, we can define a $\mathbf{k}$-algebra homomorphism $\omega' : \Lambda \to \Lambda$ by setting

$$\omega'(h_n) = e_n \qquad \text{for every } n \geq 1. \tag{13.60.1}$$

[595] The identical form of the two recursions (2.4.6) and (2.4.7) shows that this ω' also satisfies

$$\omega'(e_n) = h_n \qquad \text{for each } n \geq 0. \tag{13.60.2}$$

[596] Combining this with (13.60.1), we conclude that $(\omega' \circ \omega')(h_n) = h_n$ for each $n \geq 1$. Therefore, the two $\mathbf{k}$-algebra homomorphisms $\omega' \circ \omega' : \Lambda \to \Lambda$ and $\mathrm{id} : \Lambda \to \Lambda$ agree on each element of the generating set $\{h_n\}$ of Λ. Hence, they are equal, i.e., we have $\omega' \circ \omega' = \mathrm{id}$. Hence, ω' is an involution, and therefore an automorphism of the $\mathbf{k}$-algebra Λ. In particular, ω' is a $\mathbf{k}$-module isomorphism $\Lambda \to \Lambda$. By multiplicativity and (13.60.1), we obtain

$$\omega'(h_\lambda) = e_\lambda \qquad \text{for every partition } \lambda. \tag{13.60.3}$$

[593] It is here that we are using the assumption that u_λ and v_λ are homogeneous of degree $|\lambda|$.

[594] In fact, it is clearly enough to prove this statement for square matrices of finite size (because block-diagonal matrices can be inverted block-by-block). However, for square matrices, this follows from the fact that a surjective endomorphism of a finitely-generated $\mathbf{k}$-module is an isomorphism (Exercise 2.5.18(a)).

[595] This ω' is, of course, precisely the ω constructed in the proof of Proposition 2.4.1. But we cannot use the ω constructed in the proof of Proposition 2.4.1 at this point, because its construction made use of the fact that $(e_\lambda)_{\lambda \in \mathrm{Par}}$ is a $\mathbf{k}$-basis of Λ (and we want to prove this fact).

[596] The details of this proof proceed just like the proof of (2.4.9), except for the fact that the e_n have traded places with the h_n.

Thus, the image of the basis $(h_\lambda)_{\lambda\in\mathrm{Par}}$ of Λ under the $\mathbf{k}$-module isomorphism ω' is the family $(e_\lambda)_{\lambda\in\mathrm{Par}}$. Therefore, this latter family $(e_\lambda)_{\lambda\in\mathrm{Par}}$ must be a basis of Λ (being the image of a basis under a $\mathbf{k}$-module isomorphism).

13.61. Solution to Exercise 2.5.21.

Solution to Exercise 2.5.21. When $\mathbb{Q}$ is a subring of $\mathbf{k}$, the statement of Exercise 2.5.21 has been proven in (2.5.13). However, there is no immediate way to reuse the proof of (2.5.13) in the general case (because this proof made use of logarithms, and these are only defined when $\mathbb{Q}$ is a subring of $\mathbf{k}$). Nevertheless, it is possible to imitate that proof by defining a notion of "logarithmic derivative" even in the absence of a logarithm. Let us elaborate on this.

Let A be any commutative ring. Whenever $Q \in A[[t]]$ is a formal power series with constant term 1, we define the *logarithmic derivative* of Q to be the power series $\dfrac{Q'}{Q} \in A[[t]]$ (where Q' denotes the derivative of Q, as usual). We denote this logarithmic derivative by $\operatorname{lder} Q$. When $\mathbb{Q}$ is a subring of A, we have $\operatorname{lder} Q = \dfrac{d}{dt}(\log Q)$, but the concept of $\operatorname{lder} Q$ is defined even when log is not. Some authors find it instructive to write $\dfrac{d}{dt}(\log Q)$ for $\operatorname{lder} Q$, but we prefer not to do so, since this might tempt us to write things which make no sense.

It is easy to see that the map

$$\begin{aligned}\{Q \in A[[t]] \ \mid\ Q \text{ has constant term } 1\} &\to A[[t]],\\ Q &\mapsto \operatorname{lder} Q\end{aligned}$$

is continuous (where we equip $A[[t]]$ with the usual topology – i.e., the product topology obtained by regarding the set $A[[t]]$ as a direct product of infinitely many copies of A). Moreover, whenever I is a set and $(Q_i)_{i\in I} \in (A[[t]])^I$ is a family of formal power series with constant term 1 such that the product $\prod_{i\in I} Q_i$ converges (with respect to our topology on $A[[t]]$), then

$$\operatorname{lder}\left(\prod_{i\in I} Q_i\right) = \sum_{i\in I} \operatorname{lder}(Q_i). \tag{13.61.1}$$

[597] This equality can be used as a substitute for the famous property of the logarithm to take products into sums in situations where the logarithm is not defined. Furthermore, we have

$$\operatorname{lder}\left(Q^{-1}\right) = -\operatorname{lder} Q \tag{13.61.2}$$

[597] *Proof of* (13.61.1)*:* Let I be a set, and let $(Q_i)_{i\in I} \in (A[[t]])^I$ be a family of formal power series with constant term 1 such that the product $\prod_{i\in I} Q_i$ converges (with respect to our topology on $A[[t]]$). We need to prove (13.61.1). Since the map

$$\begin{aligned}\{Q \in A[[t]] \ \mid\ Q \text{ has constant term } 1\} &\to A[[t]],\\ Q &\mapsto \operatorname{lder} Q\end{aligned}$$

is continuous (and since infinite products are limits of finite products, and infinite sums are limits of finite sums), we can WLOG assume that the set I is finite. Assuming this, we recall that the product rule for differentiating a product of several power series yields

$$\left(\prod_{i\in I} Q_i\right)' = \sum_{j\in I} Q_j' \left(\prod_{i\in I\setminus\{j\}} Q_i\right).$$

whenever $Q \in A[[t]]$ is a formal power series with constant term 1 (where Q^{-1} means the multiplicative inverse of Q) [598].

Now, set $A = \Lambda$. We have

$$H(t) = \prod_{i=1}^{\infty} (1 - x_i t)^{-1} = \prod_{i \in \{1,2,3,\ldots\}} (1 - x_i t)^{-1},$$

Now, by the definition of $\operatorname{lder}\left(\prod_{i\in I} Q_i\right)$, we have

$$\begin{aligned}
\operatorname{lder}\left(\prod_{i\in I} Q_i\right) &= \frac{\left(\prod_{i\in I} Q_i\right)'}{\prod_{i\in I} Q_i} = \frac{1}{\prod_{i\in I} Q_i} \underbrace{\left(\prod_{i\in I} Q_i\right)'}_{=\sum_{j\in I} Q_j'\left(\prod_{i\in I\setminus\{j\}} Q_i\right)} = \frac{1}{\prod_{i\in I} Q_i} \sum_{j\in I} Q_j' \left(\prod_{i\in I\setminus\{j\}} Q_i\right) \\
&= \sum_{j\in I} Q_j' \cdot \frac{\prod_{i\in I\setminus\{j\}} Q_i}{\prod_{i\in I} Q_i} = \sum_{j\in I} Q_j' \cdot \underbrace{\frac{\prod_{i\in I\setminus\{j\}} Q_i}{Q_j \prod_{i\in I\setminus\{j\}} Q_i}}_{=\frac{1}{Q_j}} \qquad \left(\text{since } \prod_{i\in I} Q_i = Q_j \prod_{i\in I\setminus\{j\}} Q_i \text{ for every } j \in I\right) \\
&= \sum_{j\in I} Q_j' \cdot \frac{1}{Q_j} = \sum_{j\in I} \underbrace{\frac{Q_j'}{Q_j}}_{\substack{=\operatorname{lder}(Q_j) \\ (\text{since } \operatorname{lder}(Q_j) \text{ is} \\ \text{defined as } \frac{Q_j'}{Q_j})}} = \sum_{j\in I} \operatorname{lder}(Q_j) = \sum_{i\in I} \operatorname{lder}(Q_i).
\end{aligned}$$

This proves (13.61.1).

[598] *Proof of* (13.61.2)*:* Let $Q \in A[[t]]$ be a formal power series with constant term 1. Then, the product rule for the derivative of a product of two power series yields $\left(QQ^{-1}\right)' = Q'Q^{-1} + Q\left(Q^{-1}\right)'$, so that $Q'Q^{-1} + Q\left(Q^{-1}\right)' = \Big(\underbrace{QQ^{-1}}_{=1}\Big)' = 1' = 0$ and thus $Q\left(Q^{-1}\right)' = -Q'Q^{-1}$, so that $\frac{\left(Q^{-1}\right)'}{Q^{-1}} = -\frac{Q'}{Q}$. But the definition of $\operatorname{lder}\left(Q^{-1}\right)$ yields $\operatorname{lder}\left(Q^{-1}\right) = \frac{\left(Q^{-1}\right)'}{Q^{-1}} = -\underbrace{\frac{Q'}{Q}}_{=\operatorname{lder} Q} = -\operatorname{lder} Q$. This proves (13.61.2).

so that

$$\begin{aligned}
\operatorname{lder}\left(H\left(t\right)\right) &= \operatorname{lder}\left(\prod_{i\in\{1,2,3,\ldots\}}\left(1-x_it\right)^{-1}\right) = \sum_{i\in\{1,2,3,\ldots\}} \underbrace{\operatorname{lder}\left(\left(1-x_it\right)^{-1}\right)}_{\substack{=-\operatorname{lder}(1-x_it) \\ \text{(by (13.61.2), applied to } Q=1-x_it)}} \\
&\qquad \left(\text{by (13.61.1), applied to } I=\{1,2,3,\ldots\} \text{ and } Q_i=\left(1-x_it\right)^{-1}\right) \\
&= \sum_{i\in\{1,2,3,\ldots\}}\left(-\underbrace{\operatorname{lder}\left(1-x_it\right)}_{\substack{=\dfrac{\left(1-x_it\right)'}{1-x_it} \\ \text{(by the definition of } \operatorname{lder}(1-x_it))}}\right) = \sum_{i\in\{1,2,3,\ldots\}} \underbrace{\left(-\frac{\left(1-x_it\right)'}{1-x_it}\right)}_{\substack{=-\dfrac{-x_i}{1-x_i}t \\ \text{(since } (1-x_it)'=-x_i)}} \\
&= \sum_{i\in\{1,2,3,\ldots\}} \underbrace{\left(-\frac{-x_i}{1-x_it}\right)}_{=x_i\cdot\dfrac{1}{1-x_it}} = \sum_{i\in\{1,2,3,\ldots\}} x_i \cdot \underbrace{\frac{1}{1-x_it}}_{=\sum_{m\geq 0}(x_it)^m=\sum_{m\geq 0}x_i^mt^m} \\
&= \sum_{i\in\{1,2,3,\ldots\}} x_i \cdot \sum_{m\geq 0} x_i^m t^m = \sum_{m\geq 0}\left(\sum_{i\in\{1,2,3,\ldots\}} \underbrace{x_i\cdot x_i^m}_{=x_i^{m+1}}\right)t^m \\
&= \sum_{m\geq 0}\underbrace{\left(\sum_{i\in\{1,2,3,\ldots\}} x_i^{m+1}\right)}_{=p_{m+1}} t^m = \sum_{m\geq 0} p_{m+1}t^m.
\end{aligned}$$

Thus,

$$\begin{aligned}
\sum_{m\geq 0} p_{m+1}t^m = \operatorname{lder}\left(H\left(t\right)\right) &= \frac{\left(H\left(t\right)\right)'}{H\left(t\right)} \qquad \left(\text{by the definition of } \operatorname{lder}\left(H\left(t\right)\right)\right) \\
&= \frac{H'\left(t\right)}{H\left(t\right)}.
\end{aligned}$$

This solves the exercise.

Remark: Another solution of Exercise 2.5.21 proceeds by noticing that the statement of this exercise can be rewritten in the form $\left(\sum_{m\geq 0} p_{m+1}t^m\right)\cdot H\left(t\right) = H'\left(t\right)$, which (by comparison of coefficients) is equivalent to saying that

$$\text{every } m\in\mathbb{N} \text{ satisfies } \sum_{i=0}^{m} p_{i+1}h_{m-i} = \left(m+1\right)h_{m+1}. \tag{13.61.3}$$

But it is clear that once (13.61.3) is proven for $\mathbf{k}=\mathbb{Z}$, it immediately follows that (13.61.3) also holds for all $\mathbf{k}$ (since $\Lambda_{\mathbf{k}} = \mathbf{k}\otimes_{\mathbb{Z}}\Lambda_{\mathbb{Z}}$). Proving (13.61.3) for $\mathbf{k}=\mathbb{Z}$, in turn, boils down to proving (13.61.3) for $\mathbf{k}=\mathbb{Q}$, which we already know how to do (from the proof of (2.5.13)).

13.62. **Solution to Exercise 2.5.22.** *Solution to Exercise 2.5.22.* For every partition λ, define an element $v_\lambda \in \Lambda$ by $v_\lambda = v_{\lambda_1}v_{\lambda_2}\cdots v_{\lambda_{\ell(\lambda)}}$. (This is well-defined, since $\lambda_1,\lambda_2,\ldots,\lambda_{\ell(\lambda)}$ are positive integers whenever λ is a partition.) Thus, Exercise 2.2.14(b) (applied to $A=\Lambda$) shows that the elements $v_1,v_2,v_3,\ldots$ generate the $\mathbf{k}$-algebra Λ if and only if the family $(v_\lambda)_{\lambda\in\operatorname{Par}}$ spans the $\mathbf{k}$-module Λ. Hence, the family $(v_\lambda)_{\lambda\in\operatorname{Par}}$ spans the $\mathbf{k}$-module Λ (since the elements $v_1, v_2, v_3, \ldots$ generate the $\mathbf{k}$-algebra Λ).

Furthermore, v_λ is an element of $\Lambda_{|\lambda|}$ for each partition λ [599]. Hence, Exercise 2.5.19 yields that the family $(v_\lambda)_{\lambda\in\operatorname{Par}}$ is a graded basis of the graded $\mathbf{k}$-module Λ (where the indexing set Par is partitioned into $\operatorname{Par}_0, \operatorname{Par}_1, \operatorname{Par}_2, \ldots$). This solves Exercise 2.5.22(b).

(a) The family $(v_\lambda)_{\lambda\in\operatorname{Par}}$ is a basis of the $\mathbf{k}$-module Λ (since it is a graded basis of the graded $\mathbf{k}$-module Λ). Hence, in particular, this family $(v_\lambda)_{\lambda\in\operatorname{Par}}$ is $\mathbf{k}$-linearly independent.

But Exercise 2.2.14(c) (applied to $A = \Lambda$) yields that the elements $v_1, v_2, v_3, \ldots$ are algebraically independent over $\mathbf{k}$ if and only if the family $(v_\lambda)_{\lambda\in\operatorname{Par}}$ is $\mathbf{k}$-linearly independent. Hence, the elements $v_1, v_2, v_3, \ldots$ are algebraically independent over $\mathbf{k}$ (since the family $(v_\lambda)_{\lambda\in\operatorname{Par}}$ is $\mathbf{k}$-linearly independent). This solves Exercise 2.5.22(a).

13.63. **Solution to Exercise 2.5.23.** *Solution to Exercise 2.5.23.* Let V denote the $\mathbf{k}$-subalgebra of Λ generated by $v_1, v_2, v_3, \ldots$. Thus, V is a $\mathbf{k}$-submodule of Λ.

We shall now show that

$$h_n \in V \qquad \text{for each positive integer } n. \tag{13.63.1}$$

[*Proof of (13.63.1):* We shall prove (13.63.1) by strong induction on n:

Induction step: Let m be a positive integer. Assume that (13.63.1) holds for all $n < m$. We must prove that (13.63.1) holds for $n = m$. In other words, we must prove that $h_m \in V$.

We have assumed that (13.63.1) holds for all $n < m$. In other words, we have

$$h_n \in V \qquad \text{for each positive integer } n < m. \tag{13.63.2}$$

We have assumed that the element $a_{(n)} \in \mathbf{k}$ is invertible for each positive integer n. Applying this to $n = m$, we conclude that the element $a_{(m)} \in \mathbf{k}$ is invertible. Hence, its inverse $a_{(m)}^{-1} \in \mathbf{k}$ is well-defined.

But we have also assumed that each positive integer n satisfies $v_n = \sum_{\lambda\in\operatorname{Par}_n} a_\lambda h_\lambda$. Applying this to $n = m$, we find

$$v_m = \sum_{\lambda\in\operatorname{Par}_m} a_\lambda h_\lambda = a_{(m)} h_{(m)} + \sum_{\substack{\lambda\in\operatorname{Par}_m;\\ \lambda\neq(m)}} a_\lambda h_\lambda$$

(here, we have split off the addend for $\lambda = (m)$ from the sum, since $(m) \in \operatorname{Par}_m$). Hence,

$$a_{(m)} h_{(m)} = v_m - \sum_{\substack{\lambda\in\operatorname{Par}_m;\\ \lambda\neq(m)}} a_\lambda h_\lambda. \tag{13.63.3}$$

[599] *Proof.* Let λ be a partition. Thus, $\lambda = \left(\lambda_1, \lambda_2, \ldots, \lambda_{\ell(\lambda)}\right)$, so that $|\lambda| = \lambda_1 + \lambda_2 + \cdots + \lambda_{\ell(\lambda)}$.

Let $i \in \{1, 2, \ldots, \ell(\lambda)\}$. Then, the number λ_i is a positive integer (by the definition of $\ell(\lambda)$). But recall that $v_n \in \Lambda_n$ for each positive integer n. Applying this to $n = \lambda_i$, we obtain $v_{\lambda_i} \in \Lambda_{\lambda_i}$ (since λ_i is a positive integer). In other words, the symmetric function v_{λ_i} is homogeneous of degree λ_i.

Forget that we fixed i. We thus have showed that the symmetric function v_{λ_i} is homogeneous of degree λ_i for each $i \in \{1, 2, \ldots, \ell(\lambda)\}$. Hence, the product $v_{\lambda_1} v_{\lambda_2} \cdots v_{\lambda_{\ell(\lambda)}}$ of these $\ell(\lambda)$ symmetric functions is homogeneous of degree

$$\begin{aligned} &\lambda_1 + \lambda_2 + \cdots + \lambda_{\ell(\lambda)} && (\text{since } \Lambda \text{ is a graded } \mathbf{k}\text{-algebra}) \\ &= |\lambda| && \left(\text{since } |\lambda| = \lambda_1 + \lambda_2 + \cdots + \lambda_{\ell(\lambda)}\right). \end{aligned}$$

In other words, $v_{\lambda_1} v_{\lambda_2} \cdots v_{\lambda_{\ell(\lambda)}} \in \Lambda_{|\lambda|}$. But our definition of v_λ yields $v_\lambda = v_{\lambda_1} v_{\lambda_2} \cdots v_{\lambda_{\ell(\lambda)}} \in \Lambda_{|\lambda|}$. In other words, v_λ is an element of $\Lambda_{|\lambda|}$. Qed.

But each $\lambda \in \operatorname{Par}_m$ satisfying $\lambda \neq (m)$ satisfies $a_\lambda h_\lambda \in V$ [600]. Hence,

$$\sum_{\substack{\lambda \in \operatorname{Par}_m; \\ \lambda \neq (m)}} \underbrace{a_\lambda h_\lambda}_{\in V} \in \sum_{\substack{\lambda \in \operatorname{Par}_m; \\ \lambda \neq (m)}} V \subset V \qquad (\text{since } V \text{ is a } \mathbf{k}\text{-module}).$$

Furthermore, V contains $v_1, v_2, v_3, \ldots$ (since V is the $\mathbf{k}$-subalgebra of Λ generated by $v_1, v_2, v_3, \ldots$). Thus, in particular, V contains v_m (since m is a positive integer). In other words, $v_m \in V$.

Now, (13.63.3) becomes

$$a_{(m)} h_{(m)} = \underbrace{v_m}_{\in V} - \underbrace{\sum_{\substack{\lambda \in \operatorname{Par}_m; \\ \lambda \neq (m)}} a_\lambda h_\lambda}_{\in V} \in V - V \subset V \qquad (\text{since } V \text{ is a } \mathbf{k}\text{-module}).$$

Now, since $a_{(m)}^{-1}$ is well-defined, we have

$$h_{(m)} = a_{(m)}^{-1} \underbrace{a_{(m)} h_{(m)}}_{\in V} \in a_{(m)}^{-1} V \subset V \qquad (\text{since } V \text{ is a } \mathbf{k}\text{-module}).$$

In view of $h_{(m)} = h_m$, this rewrites as $h_m \in V$. In other words, (13.63.1) holds for $n = m$. This completes the induction step. Thus, (13.63.1) is proved by strong induction.]

Let us now draw some conclusions. We can restate (13.63.1) as follows: The elements $h_1, h_2, h_3, \ldots$ all belong to V. In other words, V contains $h_1, h_2, h_3, \ldots$.

Proposition 2.4.1 shows that the family $(h_n)_{n=1,2,\ldots}$ generates the $\mathbf{k}$-algebra $\Lambda_{\mathbf{k}}$. In other words, the family $(h_n)_{n=1,2,\ldots}$ generates the $\mathbf{k}$-algebra Λ (since $\Lambda_{\mathbf{k}} = \Lambda$). In other words, the elements $h_1, h_2, h_3, \ldots$ generate the $\mathbf{k}$-algebra Λ. Thus, the only $\mathbf{k}$-subalgebra of Λ that contains these elements $h_1, h_2, h_3, \ldots$ is the whole algebra Λ. In other words,

$$(\text{if } B \text{ is any } \mathbf{k}\text{-subalgebra of } \Lambda \text{ that contains } h_1, h_2, h_3, \ldots, \text{ then } B = \Lambda).$$

We can apply this to $B = V$ (since V is a $\mathbf{k}$-subalgebra of Λ that contains $h_1, h_2, h_3, \ldots$), and thus conclude that $V = \Lambda$. In other words, the $\mathbf{k}$-subalgebra of Λ generated by $v_1, v_2, v_3, \ldots$ is Λ (since V was defined to be the $\mathbf{k}$-subalgebra of Λ generated by $v_1, v_2, v_3, \ldots$). In other words, the elements $v_1, v_2, v_3, \ldots$ generate the

[600] *Proof.* Let $\lambda \in \operatorname{Par}_m$ satisfy $\lambda \neq (m)$. We must show that $a_\lambda h_\lambda \in V$.

Write the partition λ in the form $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_\ell)$ with $\ell = \ell(\lambda)$. Then, the definition of h_λ yields $h_\lambda = h_{\lambda_1} h_{\lambda_2} \cdots h_{\lambda_\ell}$. From $\lambda \in \operatorname{Par}_m$, we conclude that $m = |\lambda| = \lambda_1 + \lambda_2 + \cdots + \lambda_\ell$ (since $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_\ell)$).

Assume (for the sake of contradiction) that $\ell = 1$. Then, we have $m = \lambda_1 + \lambda_2 + \cdots + \lambda_\ell = \lambda_1$ (since $\ell = 1$), so that $\lambda_1 = m$. Moreover, from $\ell = 1$, we obtain $(\lambda_1, \lambda_2, \ldots, \lambda_\ell) = (\lambda_1) = (m)$ (since $\lambda_1 = m$), and therefore $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_\ell) = (m)$. This contradicts $\lambda \neq (m)$.

This contradiction shows that our assumption was wrong. Hence, we must have $\ell \neq 1$.

Now, let $i \in \{1, 2, \ldots, \ell\}$ be arbitrary. Then, $i \leq \ell = \ell(\lambda)$. Hence, λ_i is a positive integer.

Also, from $i \in \{1, 2, \ldots, \ell\}$, we obtain $1 \leq i \leq \ell$, so that $\ell \geq 1$. Combining this with $\ell \neq 1$, we obtain $\ell > 1$, so that $\ell \geq 2$. Hence, $2 \leq \ell = \ell(\lambda)$, so that $\lambda_2 > 0$ (by the definition of $\ell(\lambda)$).

But λ is a partition; thus, $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_\ell$. Hence, $\lambda_1 \geq \lambda_i$ (since $1 \leq i \leq \ell$), so that $\lambda_i \leq \lambda_1$. Also,

$$m = |\lambda| = \lambda_1 + \lambda_2 + \lambda_3 + \cdots = \lambda_1 + \underbrace{\lambda_2}_{>0} + \underbrace{(\lambda_3 + \lambda_4 + \lambda_5 + \cdots)}_{\geq 0} > \lambda_1.$$

In other words, $\lambda_1 < m$. Thus, $\lambda_i \leq \lambda_1 < m$. Hence, (13.63.2) (applied to $n = \lambda_i$) yields $h_{\lambda_i} \in V$ (since λ_i is a positive integer).

Forget that we fixed i. We thus have shown that $h_{\lambda_i} \in V$ for each $i \in \{1, 2, \ldots, \ell\}$. In other words, $h_{\lambda_1}, h_{\lambda_2}, \ldots, h_{\lambda_\ell}$ are elements of V. Thus, $h_{\lambda_1} h_{\lambda_2} \cdots h_{\lambda_\ell}$ is a product of elements of V, and therefore must itself be an element of V (since V is a $\mathbf{k}$-subalgebra of Λ). In other words, $h_{\lambda_1} h_{\lambda_2} \cdots h_{\lambda_\ell} \in V$. Now, $h_\lambda = h_{\lambda_1} h_{\lambda_2} \cdots h_{\lambda_\ell} \in V$. Hence, $a_\lambda \underbrace{h_\lambda}_{\in V} \in a_\lambda V \subset V$ (since V is a $\mathbf{k}$-module). Qed.

k-algebra Λ. Moreover, we have $v_n \in \Lambda_n$ for each positive integer n [601]. Thus, Exercise 2.5.22(a) shows that $v_1, v_2, v_3, \ldots$ are algebraically independent over **k**. This completes the solution to Exercise 2.5.23.

13.64. **Solution to Exercise 2.5.24.** *Solution to Exercise 2.5.24.* Corollary 2.5.17(a) says that the bases $(h_\lambda)_{\lambda \in \operatorname{Par}}$ and $(m_\lambda)_{\lambda \in \operatorname{Par}}$ of Λ are dual with respect to the Hall inner product on Λ. In other words,

$$(h_\lambda, m_\mu) = \delta_{\lambda,\mu} \qquad \text{for any } \lambda \in \operatorname{Par} \text{ and } \mu \in \operatorname{Par}. \tag{13.64.1}$$

Proposition 2.4.1 shows that the family $(h_n)_{n=1,2,\ldots}$ generates the **k**-algebra $\Lambda_{\mathbf{k}}$. In other words, the family $(h_n)_{n=1,2,\ldots}$ generates the **k**-algebra Λ (since $\Lambda_{\mathbf{k}} = \Lambda$). In other words, $h_1, h_2, h_3, \ldots$ generate the **k**-algebra Λ. Furthermore, $h_n \in \Lambda_n$ for each positive integer n. Finally, for every partition λ, the symmetric function $h_\lambda \in \Lambda$ is defined by $h_\lambda = h_{\lambda_1} h_{\lambda_2} \cdots h_{\lambda_{\ell(\lambda)}}$. Hence, Exercise 2.5.22(b) (applied to h_n and h_λ instead of v_n and v_λ) shows that the family $(h_\lambda)_{\lambda \in \operatorname{Par}}$ is a graded basis of the graded **k**-module Λ. In other words, for each $n \in \mathbb{N}$,

$$\text{the family } (h_\lambda)_{\lambda \in \operatorname{Par}_n} \text{ is a basis of the } \mathbf{k}\text{-module } \Lambda_n. \tag{13.64.2}$$

Now, let n be a positive integer. Then, $v_n \in \Lambda_n$ (by assumption) and $(n) \in \operatorname{Par}_n$ (since (n) is a partition of n). But (13.64.2) shows that the family $(h_\lambda)_{\lambda \in \operatorname{Par}_n}$ is a basis of the **k**-module Λ_n. Hence, this family spans Λ_n. Thus, v_n is a **k**-linear combination of the family $(h_\lambda)_{\lambda \in \operatorname{Par}_n}$ (since $v_n \in \Lambda_n$). In other words, there exists some family $(b_\lambda)_{\lambda \in \operatorname{Par}_n} \in \mathbf{k}^{\operatorname{Par}_n}$ of scalars such that $v_n = \sum_{\lambda \in \operatorname{Par}_n} b_\lambda h_\lambda$. Consider this family $(b_\lambda)_{\lambda \in \operatorname{Par}_n}$.

Now, recall that the Hall inner product has an orthonormal basis (namely, $(s_\lambda)_{\lambda \in \operatorname{Par}}$), and thus is symmetric. Therefore, every $\mu \in \operatorname{Par}_n$ satisfies

$$\begin{aligned}
(m_\mu, v_n) &= \left(\underbrace{v_n}_{= \sum_{\lambda \in \operatorname{Par}_n} b_\lambda h_\lambda}, m_\mu \right) = \left(\sum_{\lambda \in \operatorname{Par}_n} b_\lambda h_\lambda, m_\mu \right) = \sum_{\lambda \in \operatorname{Par}_n} b_\lambda \underbrace{(h_\lambda, m_\mu)}_{\substack{= \delta_{\lambda,\mu} \\ \text{(by (13.64.1))}}} \\
&\qquad \text{(since the Hall inner product is } \mathbf{k}\text{-bilinear)} \\
&= \sum_{\lambda \in \operatorname{Par}_n} b_\lambda \delta_{\lambda,\mu} = b_\mu \underbrace{\delta_{\mu,\mu}}_{\substack{=1 \\ \text{(since } \mu = \mu)}} + \sum_{\substack{\lambda \in \operatorname{Par}_n; \\ \lambda \neq \mu}} b_\lambda \underbrace{\delta_{\lambda,\mu}}_{\substack{=0 \\ \text{(since } \lambda \neq \mu)}} \\
&\qquad \text{(here, we have split off the addend for } \lambda = \mu \text{ from the sum, since } \mu \in \operatorname{Par}_n) \\
&= b_\mu + \underbrace{\sum_{\substack{\lambda \in \operatorname{Par}_n; \\ \lambda \neq \mu}} b_\lambda 0}_{=0} = b_\mu.
\end{aligned}$$

[601] *Proof.* We know that $h_n \in \Lambda_n$ for each positive integer n. Furthermore, we know that the family $(h_n)_{n=1,2,\ldots}$ generates the **k**-algebra Λ; in other words, $h_1, h_2, h_3, \ldots$ generate the **k**-algebra Λ. Finally, for every partition λ, the symmetric function $h_\lambda \in \Lambda$ is defined by $h_\lambda = h_{\lambda_1} h_{\lambda_2} \cdots h_{\lambda_{\ell(\lambda)}}$. Hence, Exercise 2.5.22(b) (applied to h_n and h_λ instead of v_n and v_λ) shows that the family $(h_\lambda)_{\lambda \in \operatorname{Par}}$ is a graded basis of the graded **k**-module Λ. In other words, for each $n \in \mathbb{N}$,

$$\text{the family } (h_\lambda)_{\lambda \in \operatorname{Par}_n} \text{ is a basis of the } \mathbf{k}\text{-module } \Lambda_n. \tag{13.63.4}$$

Now, let n be a positive integer. Then, (13.63.4) shows that the family $(h_\lambda)_{\lambda \in \operatorname{Par}_n}$ is a basis of the **k**-module Λ_n. Hence,

$$h_\lambda \in \Lambda_n \qquad \text{for each } \lambda \in \operatorname{Par}_n. \tag{13.63.5}$$

But by assumption, we have

$$v_n = \sum_{\lambda \in \operatorname{Par}_n} a_\lambda \underbrace{h_\lambda}_{\substack{\in \Lambda_n \\ \text{(by (13.63.5))}}} \in \sum_{\lambda \in \operatorname{Par}_n} a_\lambda \Lambda_n \subset \Lambda_n \qquad \text{(since } \Lambda_n \text{ is a } \mathbf{k}\text{-module)}.$$

Qed.

Renaming μ as λ in this statement, we obtain the following: Every $\lambda \in \operatorname{Par}_n$ satisfies $(m_\lambda, v_n) = b_\lambda$. Hence, every $\lambda \in \operatorname{Par}_n$ satisfies

$$b_\lambda = (m_\lambda, v_n) = \left(m_\lambda, v_{|\lambda|}\right) \tag{13.64.3}$$

(since $n = |\lambda|$ (because $\lambda \in \operatorname{Par}_n$ and thus $|\lambda| = n$)). Now,

$$v_n = \sum_{\lambda \in \operatorname{Par}_n} \underbrace{b_\lambda}_{\substack{=\left(m_\lambda, v_{|\lambda|}\right) \\ \text{(by (13.64.3))}}} h_\lambda = \sum_{\lambda \in \operatorname{Par}_n} \left(m_\lambda, v_{|\lambda|}\right) h_\lambda.$$

Forget that we fixed n. Thus, we have showed that each positive integer n satisfies $v_n = \sum_{\lambda \in \operatorname{Par}_n} \left(m_\lambda, v_{|\lambda|}\right) h_\lambda$. Moreover, the element $\left(m_{(n)}, v_{|(n)|}\right) \in \mathbf{k}$ is invertible for each positive integer n [602]. Hence, Exercise 2.5.23 (applied to $a_\lambda = \left(m_\lambda, v_{|\lambda|}\right)$) yields that the elements $v_1, v_2, v_3, \ldots$ generate the $\mathbf{k}$-algebra Λ and are algebraically independent over $\mathbf{k}$. This solves Exercise 2.5.24.

13.65. **Solution to Exercise 2.5.25.** *Solution to Exercise 2.5.25.* Corollary 2.5.17(a) yields that $(h_\lambda)_{\lambda \in \operatorname{Par}}$ and $(m_\lambda)_{\lambda \in \operatorname{Par}}$ are dual bases with respect to the Hall inner product on Λ.

Recall the following fundamental fact from linear algebra: If $\mathbf{k}$ is a commutative ring, if A is a $\mathbf{k}$-module, if $(\cdot, \cdot) : A \times A \to \mathbf{k}$ is a symmetric $\mathbf{k}$-bilinear form on A, and if $(u_\lambda)_{\lambda \in L}$ and $(v_\lambda)_{\lambda \in L}$ are two $\mathbf{k}$-bases of A which are dual to each other with respect to the form $(\cdot, \cdot)$ (where L is some indexing set), then every $a \in A$ satisfies

$$a = \sum_{\lambda \in L} (u_\lambda, a)\, v_\lambda.$$

We can apply this fact to $A = \Lambda$, $L = \operatorname{Par}$, $(u_\lambda)_{\lambda \in L} = (h_\lambda)_{\lambda \in \operatorname{Par}}$ and $(v_\lambda)_{\lambda \in L} = (m_\lambda)_{\lambda \in \operatorname{Par}}$ (since the bases $(h_\lambda)_{\lambda \in \operatorname{Par}}$ and $(m_\lambda)_{\lambda \in \operatorname{Par}}$ of Λ are dual to each other with respect to the Hall inner product $(\cdot, \cdot)$). As a result, we obtain that every $a \in \Lambda$ satisfies

$$a = \sum_{\lambda \in \operatorname{Par}} (h_\lambda, a)\, m_\lambda. \tag{13.65.1}$$

Now, let WC denote the set of all weak compositions. Thus, $\beta \in \mathrm{WC}$. Recall (from Section 2.1) the finitary symmetric group $\mathfrak{S}_{(\infty)}$ and its action on the set WC of all weak compositions. Applying the equality (13.65.1) to $a = f$, we obtain

$$\begin{aligned} f &= \sum_{\lambda \in \operatorname{Par}} (h_\lambda, f) \underbrace{m_\lambda}_{\substack{= \sum_{\alpha \in \mathfrak{S}_{(\infty)} \lambda} \mathbf{x}^\alpha \\ \text{(by (2.1.1))}}} = \sum_{\lambda \in \operatorname{Par}} (h_\lambda, f) \sum_{\alpha \in \mathfrak{S}_{(\infty)} \lambda} \mathbf{x}^\alpha \\ &= \underbrace{\sum_{\lambda \in \operatorname{Par}} \sum_{\alpha \in \mathfrak{S}_{(\infty)} \lambda}}_{= \sum_{\alpha \in \mathrm{WC}} \sum_{\substack{\lambda \in \operatorname{Par}; \\ \alpha \in \mathfrak{S}_{(\infty)} \lambda}}} (h_\lambda, f)\, \mathbf{x}^\alpha = \sum_{\alpha \in \mathrm{WC}} \sum_{\substack{\lambda \in \operatorname{Par}; \\ \alpha \in \mathfrak{S}_{(\infty)} \lambda}} (h_\lambda, f)\, \mathbf{x}^\alpha. \end{aligned}$$

Hence,

$$\begin{aligned} \left(\text{the coefficient of } \mathbf{x}^\beta \text{ in } f\right) &= \left(\text{the coefficient of } \mathbf{x}^\beta \text{ in } \sum_{\alpha \in \mathrm{WC}} \sum_{\substack{\lambda \in \operatorname{Par}; \\ \alpha \in \mathfrak{S}_{(\infty)} \lambda}} (h_\lambda, f)\, \mathbf{x}^\alpha \right) \\ &= \sum_{\substack{\lambda \in \operatorname{Par}; \\ \beta \in \mathfrak{S}_{(\infty)} \lambda}} (h_\lambda, f)\,. \end{aligned} \tag{13.65.2}$$

[602] *Proof.* We have assumed that $(p_n, v_n) \in \mathbf{k}$ is invertible for each positive integer n. In other words, $\left(m_{(n)}, v_{|(n)|}\right) \in \mathbf{k}$ is invertible for each positive integer n (since each positive integer n satisfies $m_{(n)} = p_n$ and $|(n)| = n$).

Recall that μ is the partition consisting of the nonzero entries of β (sorted in decreasing order). Hence, the partition μ has the same nonzero entries as the weak composition β (just possibly in a different order). Thus, the partition β can be obtained by rearranging the entries of the weak composition μ. More precisely, we can transform μ into β by rearranging finitely many entries of μ [603]. In other words, there exists a permutation $\sigma \in \mathfrak{S}_{(\infty)}$ such that $\beta = \sigma(\mu)$ (because rearranging finitely many entries of a weak composition is tantamount to applying a permutation in $\mathfrak{S}_{(\infty)}$ to it). In other words, $\beta \in \mathfrak{S}_{(\infty)}\mu$. Hence, μ is a $\lambda \in \operatorname{Par}$ such that $\beta \in \mathfrak{S}_{(\infty)}\lambda$.

On the other hand, recall that every weak composition α lies in the $\mathfrak{S}_{(\infty)}$-orbit of a unique partition λ. In other words, for every weak composition α, there is a unique partition λ such that $\alpha \in \mathfrak{S}_{(\infty)}\lambda$. Applying this to $\alpha = \beta$, we conclude that there is a unique partition λ such that $\beta \in \mathfrak{S}_{(\infty)}\lambda$. In other words, there is a unique $\lambda \in \operatorname{Par}$ such that $\beta \in \mathfrak{S}_{(\infty)}\lambda$. This unique λ must be μ (since μ is a $\lambda \in \operatorname{Par}$ such that $\beta \in \mathfrak{S}_{(\infty)}\lambda$).

So we know that there is a unique $\lambda \in \operatorname{Par}$ such that $\beta \in \mathfrak{S}_{(\infty)}\lambda$, and this unique λ is μ. Hence, the sum $\sum\limits_{\substack{\lambda \in \operatorname{Par};\\ \beta \in \mathfrak{S}_{(\infty)}\lambda}} (h_\lambda, f)$ has exactly one addend, namely the addend for $\lambda = \mu$. Thus, this sum simplifies as follows: $\sum\limits_{\substack{\lambda \in \operatorname{Par};\\ \beta \in \mathfrak{S}_{(\infty)}\lambda}} (h_\lambda, f) = (h_\mu, f)$.

Hence, (13.65.2) becomes

$$\left(\text{the coefficient of } \mathbf{x}^\beta \text{ in } f\right) = \sum_{\substack{\lambda \in \operatorname{Par};\\ \beta \in \mathfrak{S}_{(\infty)}\lambda}} (h_\lambda, f) = (h_\mu, f).$$

Combining this with

$$(f, h_\mu) = (h_\mu, f) \qquad \text{(by Exercise 2.5.13(c), applied to } g = h_\mu\text{)},$$

we obtain

$$(f, h_\mu) = (h_\mu, f) = \left(\text{the coefficient of } \mathbf{x}^\beta \text{ in } f\right).$$

This solves Exercise 2.5.25.

13.66. **Solution to Exercise 2.5.26.** *Solution to Exercise 2.5.26.* Let us first prove three simple claims:

Claim 1: For every $\lambda \in \operatorname{Par}$, we have $p_\lambda \in \Lambda_{|\lambda|}$.

[*Proof of Claim 1:* Let $n = |\lambda|$. Thus, $\lambda \in \operatorname{Par}_n$ (since $\lambda \in \operatorname{Par}$). Hence, we can show (just as in the solution to Exercise 2.2.13(k)) that $p_\lambda \in \Lambda_n$. In view of $n = |\lambda|$, this rewrites as $p_\lambda \in \Lambda_{|\lambda|}$. This proves Claim 1.]

Claim 2: Let $n \in \mathbb{N}$ and $\lambda \in \operatorname{Par}_n$. Then, $(p_\lambda, h_n) = 1$.

[*Proof of Claim 2:* We have $|\lambda| = n$ (since $\lambda \in \operatorname{Par}_n$). Claim 1 yields $p_\lambda \in \Lambda_{|\lambda|} = \Lambda_n$ (since $|\lambda| = n$). Hence, Exercise 2.5.13(b) (applied to $f = p_\lambda$) yields $(h_n, p_\lambda) = p_\lambda(1)$ (where $p_\lambda(1)$ is defined as in Exercise 2.1.2).

Recall that $p_\lambda(1)$ is defined to be the result of substituting $1, 0, 0, 0, \ldots$ for $x_1, x_2, x_3, \ldots$ in p_λ. In other words, $p_\lambda(1) = p_\lambda(1, 0, 0, 0, \ldots)$.

Now, write the partition λ as $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_\ell)$, where $\ell = \ell(\lambda)$. Then, the definition of p_λ yields $p_\lambda = p_{\lambda_1} p_{\lambda_2} \cdots p_{\lambda_\ell} = \prod_{i=1}^{\ell} p_{\lambda_i}$. Substituting $1, 0, 0, 0, \ldots$ for $x_1, x_2, x_3, \ldots$ on both sides of this equality, we obtain

$$p_\lambda(1, 0, 0, 0, \ldots) = \prod_{i=1}^{\ell} \underbrace{p_{\lambda_i}(1, 0, 0, 0, \ldots)}_{\substack{= 1^{\lambda_i} + 0^{\lambda_i} + 0^{\lambda_i} + 0^{\lambda_i} + \cdots \\ (\text{since } p_{\lambda_i} = x_1^{\lambda_i} + x_2^{\lambda_i} + x_3^{\lambda_i} + \cdots)}} = \prod_{i=1}^{\ell} \underbrace{\left(1^{\lambda_i} + 0^{\lambda_i} + 0^{\lambda_i} + 0^{\lambda_i} + \cdots\right)}_{\substack{= 1 + 0 + 0 + 0 + \cdots \\ (\text{since } 1^{\lambda_i} = 1 \text{ and } 0^{\lambda_i} = 0 \\ (\text{because } \lambda_i \text{ is positive (since } i \leq \ell = \ell(\lambda))))}}$$

$$= \prod_{i=1}^{\ell} \underbrace{(1 + 0 + 0 + 0 + \cdots)}_{=1} = 1.$$

[603] Indeed, there is an $i \in \{1, 2, 3, \ldots\}$ such that $\beta_{i+1} = \beta_{i+2} = \beta_{i+3} = \cdots = 0$ and $\mu_{i+1} = \mu_{i+2} = \mu_{i+3} = \cdots = 0$ (since β and μ are weak compositions). We only need to rearrange the first i entries in order to transform μ into β.

Summarizing what we have proved so far, we obtain

$$(h_n, p_\lambda) = p_\lambda (1) = p_\lambda (1, 0, 0, 0, \ldots) = 1.$$

But Exercise 2.5.13(c) (applied to $f = p_\lambda$ and $g = h_n$) yields $(p_\lambda, h_n) = (h_n, p_\lambda) = 1$. This proves Claim 2.]

Claim 3: Let $n \in \mathbb{N}$ and $\lambda \in \operatorname{Par}$ be such that $\lambda \notin \operatorname{Par}_n$. Then, $(p_\lambda, h_n) = 0$.

[*Proof of Claim 3:* We have $|\lambda| \neq n$ (since $\lambda \in \operatorname{Par}$ but $\lambda \notin \operatorname{Par}_n$). Therefore, $|\lambda|$ and n are two distinct nonnegative integers. Moreover, $p_\lambda \in \Lambda_{|\lambda|}$ (by Claim 1) and $h_n \in \Lambda_n$. Hence, Exercise 2.5.13(a) (applied to $|\lambda|$, n, p_λ and h_n instead of n, m, f and g) yields $(p_\lambda, h_n) = 0$. This proves Claim 3.]

Corollary 2.5.17(b) yields that $(p_\lambda)_{\lambda\in\operatorname{Par}}$ and $\left(z_\lambda^{-1} p_\lambda\right)_{\lambda\in\operatorname{Par}}$ are dual bases with respect to the Hall inner product on Λ. Recall the following fundamental fact from linear algebra: If $\mathbf{k}$ is a commutative ring, if A is a $\mathbf{k}$-module, if $(\cdot, \cdot) : A \times A \to \mathbf{k}$ is a symmetric $\mathbf{k}$-bilinear form on A, and if $(u_\lambda)_{\lambda\in L}$ and $(v_\lambda)_{\lambda\in L}$ are two $\mathbf{k}$-bases of A which are dual to each other with respect to the form $(\cdot, \cdot)$ (where L is some indexing set), then every $a \in A$ satisfies

$$a = \sum_{\lambda\in L} (u_\lambda, a) v_\lambda.$$

We can apply this fact to $A = \Lambda$, $L = \operatorname{Par}$, $(u_\lambda)_{\lambda\in L} = (p_\lambda)_{\lambda\in\operatorname{Par}}$ and $(v_\lambda)_{\lambda\in L} = \left(z_\lambda^{-1} p_\lambda\right)_{\lambda\in\operatorname{Par}}$ (since the bases $(p_\lambda)_{\lambda\in\operatorname{Par}}$ and $\left(z_\lambda^{-1} p_\lambda\right)_{\lambda\in\operatorname{Par}}$ of Λ are dual to each other with respect to the Hall inner product $(\cdot, \cdot)$). As a result, we obtain that every $a \in \Lambda$ satisfies

$$a = \sum_{\lambda\in\operatorname{Par}} (p_\lambda, a) z_\lambda^{-1} p_\lambda. \tag{13.66.1}$$

Now, let $n \in \mathbb{N}$. Applying (13.66.1) to $a = h_n$, we find

$$\begin{aligned} h_n &= \sum_{\lambda\in\operatorname{Par}} (p_\lambda, h_n) z_\lambda^{-1} p_\lambda = \underbrace{\sum_{\substack{\lambda\in\operatorname{Par};\\ \lambda\in\operatorname{Par}_n}}}_{\substack{=\sum_{\lambda\in\operatorname{Par}_n}\\ (\text{since } \operatorname{Par}_n\subset\operatorname{Par})}} \underbrace{(p_\lambda, h_n)}_{\substack{=1\\ (\text{by Claim 2})}} z_\lambda^{-1} p_\lambda + \sum_{\substack{\lambda\in\operatorname{Par};\\ \lambda\notin\operatorname{Par}_n}} \underbrace{(p_\lambda, h_n)}_{\substack{=0\\ (\text{by Claim 3})}} z_\lambda^{-1} p_\lambda \\ &\qquad \left(\begin{array}{c} \text{since each } \lambda \in \operatorname{Par} \text{ satisfies either } \lambda \in \operatorname{Par}_n \text{ or } \lambda \notin \operatorname{Par}_n \\ \text{(but not both at the same time)} \end{array}\right) \\ &= \sum_{\lambda\in\operatorname{Par}_n} z_\lambda^{-1} p_\lambda + \underbrace{\sum_{\substack{\lambda\in\operatorname{Par};\\ \lambda\notin\operatorname{Par}_n}} 0 z_\lambda^{-1} p_\lambda}_{=0} = \sum_{\lambda\in\operatorname{Par}_n} z_\lambda^{-1} p_\lambda. \end{aligned}$$

This proves (2.5.17).

Recall the map $\omega : \Lambda \to \Lambda$ introduced in Definition 2.4.2. Applying this map ω to both sides of the equality (2.5.17), we obtain

$$\begin{aligned} \omega(h_n) &= \omega\left(\sum_{\lambda\in\operatorname{Par}_n} z_\lambda^{-1} p_\lambda\right) = \sum_{\lambda\in\operatorname{Par}_n} z_\lambda^{-1} \underbrace{\omega(p_\lambda)}_{\substack{=(-1)^{|\lambda|-\ell(\lambda)} p_\lambda\\ (\text{by } (2.4.14))}} \qquad (\text{since the map } \omega \text{ is } \mathbf{k}\text{-linear}) \\ &= \sum_{\lambda\in\operatorname{Par}_n} \underbrace{z_\lambda^{-1} (-1)^{|\lambda|-\ell(\lambda)}}_{=(-1)^{|\lambda|-\ell(\lambda)} z_\lambda^{-1}} p_\lambda = \sum_{\lambda\in\operatorname{Par}_n} (-1)^{|\lambda|-\ell(\lambda)} z_\lambda^{-1} p_\lambda. \end{aligned}$$

But Proposition 2.4.3(b) yields $\omega(h_n) = e_n$. Comparing these two equalities, we obtain

$$e_n = \sum_{\lambda\in\operatorname{Par}_n} (-1)^{|\lambda|-\ell(\lambda)} z_\lambda^{-1} p_\lambda.$$

This proves (2.5.18). Thus, the solution to Exercise 2.5.26 is complete.

13.67. **Solution to Exercise 2.7.5.** *Solution to Exercise 2.7.5.* We recall that the cells of λ/μ are the cells $(p,q) \in \{1,2,3,...\}^2$ satisfying $\mu_p < q \leq \lambda_p$.

(a) In order to solve Exercise 2.7.5(a), we need to prove the following two claims:

$$\text{(if } \lambda/\mu \text{ is a horizontal strip, then every } i \in \{1,2,3,...\} \text{ satisfies } \mu_i \geq \lambda_{i+1}) \tag{13.67.1}$$

and

$$\text{(if every } i \in \{1,2,3,...\} \text{ satisfies } \mu_i \geq \lambda_{i+1}, \text{ then } \lambda/\mu \text{ is a horizontal strip)}. \tag{13.67.2}$$

Proof of (13.67.1)*:* Assume that λ/μ is a horizontal strip. In other words, no two cells of λ/μ lie in the same column.

Let $i \in \{1,2,3,...\}$. Assume (for the sake of contradiction) that $\mu_i < \lambda_{i+1}$. Hence, $\lambda_{i+1} > \mu_i \geq 0$. Also, $\mu \subseteq \lambda$ yields $\mu_i \leq \lambda_i$. Hence, $\mu_{i+1} \leq \mu_i < \lambda_{i+1}$. Now, (i, λ_{i+1}) is a cell of λ/μ (since $\mu_i < \lambda_{i+1} \leq \lambda_i$). Also, $(i+1, \lambda_{i+1})$ is a cell of λ/μ (since $\mu_{i+1} < \lambda_{i+1} \leq \lambda_{i+1}$). Thus, (i, λ_{i+1}) and $(i+1, \lambda_{i+1})$ are two cells of λ/μ lying in the same column. This contradicts the fact that no two cells of λ/μ lie in the same column. This contradiction shows that our assumption (that $\mu_i < \lambda_{i+1}$) was wrong. Hence, $\mu_i \geq \lambda_{i+1}$. This proves (13.67.1).

Proof of (13.67.2)*:* Assume that every $i \in \{1,2,3,...\}$ satisfies $\mu_i \geq \lambda_{i+1}$.

Now, let us (for the sake of contradiction) assume that there exist two distinct cells of λ/μ which lie in the same column. Let c and d be two such cells. Thus, c and d are two distinct cells of λ/μ which lie in the same column.

Write c and d in the forms $c = (p_c, q_c)$ and $d = (p_d, q_d)$ for some positive integers p_c, q_c, p_d, q_d. We have $\mu_{p_c} < q_c \leq \lambda_{p_c}$ (since $(p_c, q_c) = c$ is a cell of λ/μ) and $\mu_{p_d} < q_d \leq \lambda_{p_d}$ (similarly). The cells c and d lie in the same column; in other words, $q_c = q_d$ (because the cell $c = (p_c, q_c)$ lies in column q_c, and the cell $d = (p_d, q_d)$ lies in column q_d).

Our situation so far is symmetric with respect to interchanging c with d. Hence, we can WLOG assume that $p_c \leq p_d$ (because otherwise, we can switch c with d). Assume this.

If we had $p_c = p_d$, then we would have $c = \left(\underbrace{p_c}_{=p_d}, \underbrace{q_c}_{=q_d}\right) = (p_d, q_d) = d$, which would contradict the fact that c and d are distinct. Hence, we cannot have $p_c = p_d$. Thus, we have $p_c \neq p_d$. Combined with $p_c \leq p_d$, this yields $p_c < p_d$. Thus, $p_c \leq p_d - 1$ (since p_c and p_d are integers), so that $p_c + 1 \leq p_d$.

Since λ is a partition, we have $\lambda_u \geq \lambda_v$ for any positive integers u and v satisfying $u \leq v$. Applying this to $u = p_c + 1$ and $v = p_d$, we obtain $\lambda_{p_c+1} \geq \lambda_{p_d}$ (since $p_c + 1 \leq p_d$). But recall that every $i \in \{1,2,3,...\}$ satisfies $\mu_i \geq \lambda_{i+1}$. Applying this to $i = p_c$, we obtain $\mu_{p_c} \geq \lambda_{p_c+1}$. Now, recall that $\mu_{p_c} < q_c$, so that $q_c > \mu_{p_c} \geq \lambda_{p_c+1} \geq \lambda_{p_d} \geq q_d$ (since $q_d \leq \lambda_{p_d}$). This contradicts $q_c = q_d$. This contradiction shows that our assumption (that there exist two distinct cells of λ/μ which lie in the same column) was wrong. Hence, no two cells of λ/μ lie in the same column. In other words, λ/μ is a horizontal strip. This proves (13.67.2).

Now, both (13.67.1) and (13.67.2) are proven. Thus, Exercise 2.7.5(a) is solved.

(b) We have the following equivalence of statements:

$$\begin{aligned}
& (\lambda/\mu \text{ is a vertical strip}) \\
\Longleftrightarrow\ & (\text{no two cells of } \lambda/\mu \text{ lie in the same row}) \qquad (\text{by the definition of a ``vertical strip''}) \\
\Longleftrightarrow\ & (\text{for every } i \in \{1,2,3,\ldots\}, \text{ no two cells of } \lambda/\mu \text{ lie in row } i) \\
\Longleftrightarrow\ & \left(\text{for every } i \in \{1,2,3,\ldots\}, \underbrace{\text{the number of cells of } \lambda/\mu \text{ in row } i}_{=\lambda_i - \mu_i} \text{ is } \leq 1\right) \\
\Longleftrightarrow\ & (\text{for every } i \in \{1,2,3,\ldots\}, \text{ we have } \lambda_i - \mu_i \leq 1) \\
\Longleftrightarrow\ & (\text{for every } i \in \{1,2,3,\ldots\}, \text{ we have } \lambda_i \leq \mu_i + 1).
\end{aligned}$$

This solves Exercise 2.7.5(b).

13.68. **Solution to Exercise 2.7.6.** *Solution to Exercise 2.7.6.* (a) We use the notation $f(a_1, a_2, \ldots, a_k)$ defined in Exercise 2.1.2 whenever $a_1, a_2, \ldots, a_k$ are elements of a commutative $\mathbf{k}$-algebra A and $f \in R(\mathbf{x})$. In particular, $f(1)$ is a well-defined element of $\mathbf{k}$ for every $f \in R(\mathbf{x})$. Every $f \in R(\mathbf{x})$ satisfies

$$f(1) = (\text{the result of substituting } 1, 0, 0, 0, \ldots \text{ for } x_1, x_2, x_3, \ldots \text{ in } f). \tag{13.68.1}$$

We have $s_{\lambda/\mu} = \sum_T \mathbf{x}^{\mathrm{cont}(T)}$, where T runs through all column-strict tableaux of shape λ/μ. In other words, $s_{\lambda/\mu} = \sum_{\substack{T \text{ is a column-strict} \\ \text{tableau of shape } \lambda/\mu}} \mathbf{x}^{\mathrm{cont}(T)}$. Hence,

$$\begin{aligned} s_{\lambda/\mu}(1) &= \left(\sum_{\substack{T \text{ is a column-strict} \\ \text{tableau of shape } \lambda/\mu}} \mathbf{x}^{\mathrm{cont}(T)} \right)(1) \\ &= \left(\text{the result of substituting } 1, 0, 0, 0, \ldots \text{ for } x_1, x_2, x_3, \ldots \text{ in } \sum_{\substack{T \text{ is a column-strict} \\ \text{tableau of shape } \lambda/\mu}} \mathbf{x}^{\mathrm{cont}(T)} \right) \end{aligned} \tag{13.68.2}$$

(by (13.68.1)).

But the substitution $1, 0, 0, 0, \ldots$ for $x_1, x_2, x_3, \ldots$ has the following effect on any given monomial $\mathbf{x}^\alpha$:

- if none of the indeterminates $x_2, x_3, x_4, \ldots$ occur in this monomial $\mathbf{x}^\alpha$, then the monomial $\mathbf{x}^\alpha$ goes to 1;
- otherwise, the monomial $\mathbf{x}^\alpha$ goes to 0.

Hence, applying this substitution to $\sum_{\substack{T \text{ is a column-strict} \\ \text{tableau of shape } \lambda/\mu}} \mathbf{x}^{\mathrm{cont}(T)}$ yields a sum of 1's over all column-strict tableaux T of shape λ/μ having the property that none of the indeterminates $x_2, x_3, x_4, \ldots$ occur in this monomial $\mathbf{x}^{\mathrm{cont}(T)}$. In other words,

$$\begin{aligned} &\left(\text{the result of substituting } 1, 0, 0, 0, \ldots \text{ for } x_1, x_2, x_3, \ldots \text{ in } \sum_{\substack{T \text{ is a column-strict} \\ \text{tableau of shape } \lambda/\mu}} \mathbf{x}^{\mathrm{cont}(T)} \right) \\ &= \sum_{\substack{T \text{ is a column-strict tableau of shape } \lambda/\mu; \\ \text{none of the indeterminates } x_2, x_3, x_4, \ldots \\ \text{occur in the monomial } \mathbf{x}^{\mathrm{cont}(T)}}} 1. \end{aligned}$$

Therefore, (13.68.2) rewrites as

$$s_{\lambda/\mu}(1) = \sum_{\substack{T \text{ is a column-strict tableau of shape } \lambda/\mu; \\ \text{none of the indeterminates } x_2, x_3, x_4, \ldots \\ \text{occur in the monomial } \mathbf{x}^{\mathrm{cont}(T)}}} 1.$$

This rewrites as

$$s_{\lambda/\mu}(1) = \sum_{\substack{T \text{ is a column-strict tableau of shape } \lambda/\mu; \\ \text{all entries of } T \text{ are } 1}} 1 \tag{13.68.3}$$

(because for a column-strict tableau T, saying that none of the indeterminates $x_2, x_3, x_4, \ldots$ occur in the monomial $\mathbf{x}^{\mathrm{cont}(T)}$ is equivalent to saying that all entries of T are 1).

In order to solve Exercise 2.7.6(a), we need to prove the following two statements:

$$\left(\text{if } \lambda/\mu \text{ is a horizontal } n\text{-strip, then } \left(h_n, s_{\lambda/\mu}\right) = 1 \right) \tag{13.68.4}$$

and

$$\left(\text{if } \lambda/\mu \text{ is not a horizontal } n\text{-strip, then } \left(h_n, s_{\lambda/\mu}\right) = 0 \right). \tag{13.68.5}$$

Proof of (13.68.4)*:* Assume that λ/μ is a horizontal n-strip. Thus, in particular, we have $|\lambda/\mu| = n$. Hence, $s_{\lambda/\mu} \in \Lambda_{|\lambda/\mu|} = \Lambda_n$ (since $|\lambda/\mu| = n$). Thus, Exercise 2.5.13(b) (applied to $f = s_{\lambda/\mu}$) yields

$$(13.68.6) \qquad \left(h_n, s_{\lambda/\mu}\right) = s_{\lambda/\mu}(1) = \sum_{\substack{T \text{ is a column-strict tableau of shape } \lambda/\mu; \\ \text{all entries of } T \text{ are } 1}} 1 \qquad \text{(by (13.68.3))}.$$

Recall that λ/μ is a horizontal n-strip, therefore a horizontal strip. In other words, no two cells of λ/μ lie in the same column. Hence, if we fill every cell of the Ferrers diagram of λ/μ with a 1, we obtain a column-strict tableau T of shape λ/μ having the property that all entries of T are 1. Therefore, such a tableau exists. It is also clearly unique (because requiring that all entries be 1 does not leave any freedom in choosing the entries). Therefore, there exists exactly one such tableau. This shows that the sum on the right hand side of (13.68.6) has exactly one addend, and therefore equals 1 (since the addend is 1). Hence, (13.68.6) rewrites as $\left(h_n, s_{\lambda/\mu}\right) = 1$, and thus (13.68.4) is proven.

Proof of (13.68.5)*:* Assume that λ/μ is not a horizontal n-strip. We need to show that $\left(h_n, s_{\lambda/\mu}\right) = 0$.

If $|\lambda/\mu| \neq n$, then Exercise 2.5.13(a) (applied to $m = |\lambda/\mu|$, $f = h_n$ and $g = s_{\lambda/\mu}$) yields $\left(h_n, s_{\lambda/\mu}\right) = 0$ (since $h_n \in \Lambda_n$ and $s_{\lambda/\mu} \in \Lambda_{|\lambda/\mu|}$). Thus, $\left(h_n, s_{\lambda/\mu}\right) = 0$ is proven if $|\lambda/\mu| \neq n$. Hence, for the rest of the proof of $\left(h_n, s_{\lambda/\mu}\right) = 0$, we WLOG assume that we don't have $|\lambda/\mu| \neq n$. Thus, $|\lambda/\mu| = n$. We can thus prove (13.68.6) just as we did before (in the proof of (13.68.4)).

But if λ/μ were a horizontal strip, then λ/μ would be a horizontal n-strip (since $|\lambda/\mu| = n$), which would contradict our assumption that λ/μ is not a horizontal n-strip. Hence, λ/μ cannot be a horizontal strip. As a consequence, there must be two cells of λ/μ which lie in the same column. If T is a column-strict tableau of shape λ/μ, then these two cells must be filled with two different entries in T (because the entries of a column-strict tableau are strictly increasing top-to-bottom down columns, and hence all entries in any given column must be distinct), which is impossible if all entries of T are to be 1. Therefore, there exists no column-strict tableau T of shape λ/μ such that all entries of T are 1. Hence, the sum on the right hand side of (13.68.6) is empty, and therefore equals 0. So (13.68.6) rewrites as $\left(h_n, s_{\lambda/\mu}\right) = 0$, and thus (13.68.5) is proven.

Now that both (13.68.4) and (13.68.5) are proven, Exercise 2.7.6(a) is solved.

(b) Notice first that

$$(13.68.7) \qquad \left(h_n, s_{\lambda/\mu}\right) = c^{\lambda}_{\mu,(n)} \qquad \text{for every } n \in \mathbb{N},\ \lambda \in \operatorname{Par} \text{ and } \mu \in \operatorname{Par}.$$

[604]

Recall that any two partitions μ and ν satisfy

$$(13.68.8) \qquad s_\mu s_\nu = \sum_{\lambda \in \operatorname{Par}} c^{\lambda}_{\mu,\nu} s_\lambda = \sum_{\tau \in \operatorname{Par}} c^{\tau}_{\mu,\nu} s_\tau$$

(here, we renamed the summation index λ as τ).

[604] *Proof.* Let $n \in \mathbb{N}$, $\lambda \in \operatorname{Par}$ and $\mu \in \operatorname{Par}$. From Remark 2.5.9, we know that $s_{\lambda/\mu} = \sum_\nu c^{\lambda}_{\mu,\nu} s_\nu$, where the sum is over all $\nu \in \operatorname{Par}$. Thus, for every given partition τ, the s_τ-coordinate of $s_{\lambda/\mu}$ in the basis $\left(s_\nu\right)_{\nu \in \operatorname{Par}}$ of Λ equals $c^{\lambda}_{\mu,\tau}$. But since $\left(s_\nu\right)_{\nu \in \operatorname{Par}}$ is an orthonormal basis of the $\mathbf{k}$-module Λ with respect to the Hall inner product, this s_τ-coordinate also equals $\left(s_\tau, s_{\lambda/\mu}\right)$. Comparing these two expressions for this s_τ-coordinate, we obtain $c^{\lambda}_{\mu,\tau} = \left(s_\tau, s_{\lambda/\mu}\right)$. Applying this to $\tau = (n)$, we obtain $c^{\lambda}_{\mu,(n)} = \left(\underbrace{s_{(n)}}_{=h_n}, s_{\lambda/\mu}\right) = \left(h_n, s_{\lambda/\mu}\right)$, which proves (13.68.7).

Let $n \in \mathbb{N}$. Let λ be a partition. We have $s_{(n)} = h_n$, so that $h_n = s_{(n)}$. Now,

$$s_\lambda \underbrace{h_n}_{=s_{(n)}} = s_\lambda s_{(n)} = \sum_{\tau \in \operatorname{Par}} \underbrace{c^\tau_{\lambda,(n)}}_{\substack{=\left(h_n, s_{\tau/\lambda}\right) \\ \text{(because (13.68.7) (applied to } \tau \\ \text{and } \lambda \text{ instead of } \lambda \text{ and } \mu\text{) yields} \\ \left(h_n, s_{\tau/\lambda}\right) = c^\tau_{\lambda,(n)})}} s_\tau \qquad \text{(by (13.68.8), applied to } \mu = \lambda \text{ and } \nu = (n))$$

$$= \sum_{\tau \in \operatorname{Par}} \left(h_n, s_{\tau/\lambda}\right) s_\tau = \sum_{\substack{\tau \in \operatorname{Par}; \\ \lambda \subseteq \tau}} \left(h_n, s_{\tau/\lambda}\right) s_\tau$$

$$\left(\begin{array}{c} \text{here, we have ridden the sum of all its addends in which } \lambda \not\subseteq \tau; \\ \text{these addends were zero (because if } \lambda \not\subseteq \tau \text{, then } \left(h_n, \underbrace{s_{\tau/\lambda}}_{=0}\right) s_\tau = 0) \end{array}\right)$$

$$= \sum_{\substack{\tau \in \operatorname{Par}; \\ \lambda \subseteq \tau; \\ \tau/\lambda \text{ is a horizontal } n\text{-strip}}} \underbrace{\left(h_n, s_{\tau/\lambda}\right)}_{\substack{=1 \\ \text{(by (13.68.4), applied} \\ \text{to } \tau \text{ and } \lambda \text{ instead of } \lambda \text{ and } \mu)}} s_\tau + \sum_{\substack{\tau \in \operatorname{Par}; \\ \lambda \subseteq \tau; \\ \tau/\lambda \text{ is not a horizontal } n\text{-strip}}} \underbrace{\left(h_n, s_{\tau/\lambda}\right)}_{\substack{=0 \\ \text{(by (13.68.5), applied} \\ \text{to } \tau \text{ and } \lambda \text{ instead of } \lambda \text{ and } \mu)}} s_\tau$$

$$= \sum_{\substack{\tau \in \operatorname{Par}; \\ \lambda \subseteq \tau; \\ \tau/\lambda \text{ is a horizontal } n\text{-strip}}} s_\tau + \underbrace{\sum_{\substack{\tau \in \operatorname{Par}; \\ \lambda \subseteq \tau; \\ \tau/\lambda \text{ is not a horizontal } n\text{-strip}}} 0 s_\tau}_{=0}$$

$$= \sum_{\substack{\tau \in \operatorname{Par}; \\ \lambda \subseteq \tau; \\ \tau/\lambda \text{ is a horizontal } n\text{-strip}}} s_\tau = \sum_{\substack{\lambda^+ \in \operatorname{Par}; \\ \lambda \subseteq \lambda^+; \\ \lambda^+/\lambda \text{ is a horizontal } n\text{-strip}}} s_{\lambda^+}$$

$$\left(\text{here, we renamed the summation index } \tau \text{ as } \lambda^+\right)$$

$$= \sum_{\substack{\lambda^+ : \lambda^+/\lambda \text{ is a} \\ \text{horizontal } n\text{-strip}}} s_{\lambda^+}.$$

This proves (2.7.1). Thus, Exercise 2.7.6(b) is solved.

13.69. **Solution to Exercise 2.7.7.** *Solution to Exercise 2.7.7.* Let us use the notations introduced in the proof of Theorem 2.5.1. In particular, we use the words "letter" and "positive integer" as synonyms.

We shall use the following lemma:

Lemma 13.69.1. *Let P be a column-strict tableau, and let j and j' be two letters. Applying RS-insertion to the tableau P and the letter j yields a new column-strict tableau P' and a corner cell c. Applying RS-insertion to the tableau P' and the letter j' yields a new column-strict tableau P'' and a corner cell c'.*

(a) *Assume that $j \leq j'$. Then, the cell c' is in the same row as the cell c or in a row further up; it is also in a column further right than c.*

(b) *Assume instead that $j > j'$. Then, the cell c' is in a row further down than the cell c; it is also in the same column as c or in a column further left.*

Lemma 13.69.1 is part of the Row bumping lemma that appeared in our proof of Theorem 2.5.1.

We shall first concentrate on proving (2.7.1).

Alternative proof of (2.7.1). Let λ be a partition, and let $n \in \mathbb{N}$. The definition of s_λ yields

$$s_\lambda = \sum_{\substack{T \text{ is a column-strict} \\ \text{tableau of shape } \lambda}} \mathbf{x}^{\operatorname{cont}(T)}.$$

The definition of h_n yields

$$h_n = \sum_{i_1 \leq i_2 \leq \cdots \leq i_n} x_{i_1} x_{i_2} \cdots x_{i_n}. \tag{13.69.1}$$

Multiplying these two identities, we obtain

$$\begin{aligned} s_\lambda h_n &= \sum_{\substack{T \text{ is a column-strict} \\ \text{tableau of shape } \lambda}} \mathbf{x}^{\operatorname{cont}(T)} \sum_{i_1 \leq i_2 \leq \cdots \leq i_n} x_{i_1} x_{i_2} \cdots x_{i_n} \\ &= \sum_{(T,(i_1,i_2,\ldots,i_n)) \in \mathbf{A}} \mathbf{x}^{\operatorname{cont}(T)} x_{i_1} x_{i_2} \cdots x_{i_n}, \qquad (13.69.2) \end{aligned}$$

where $\mathbf{A}$ is the set of all pairs $(T, (i_1, i_2, \ldots, i_n))$ of a column-strict tableau T of shape λ and an n-tuple $(i_1, i_2, \ldots, i_n)$ of positive integers satisfying $i_1 \leq i_2 \leq \cdots \leq i_n$. Consider this set $\mathbf{A}$.

On the other hand, every partition λ^+ satisfies

$$\begin{aligned} s_{\lambda^+} &= \sum_{\substack{T \text{ is a column-strict} \\ \text{tableau of shape } \lambda^+}} \mathbf{x}^{\operatorname{cont}(T)} \qquad \left(\text{by the definition of } s_{\lambda^+}\right) \\ &= \sum_{\substack{S \text{ is a column-strict} \\ \text{tableau of shape } \lambda^+}} \mathbf{x}^{\operatorname{cont}(S)} \qquad \left(\text{here, we renamed the summation index } T \text{ as } S\right). \qquad (13.69.3) \end{aligned}$$

Hence,

$$\begin{aligned} \sum_{\substack{\lambda^+ : \lambda^+/\lambda \text{ is a} \\ \text{horizontal } n\text{-strip}}} \underbrace{s_{\lambda^+}}_{\substack{= \sum\limits_{\substack{S \text{ is a column-strict} \\ \text{tableau of shape } \lambda^+}} \mathbf{x}^{\operatorname{cont}(S)} \\ (\text{by (13.69.3)})}} &= \sum_{\substack{\lambda^+ : \lambda^+/\lambda \text{ is a} \\ \text{horizontal } n\text{-strip}}} \sum_{\substack{S \text{ is a column-strict} \\ \text{tableau of shape } \lambda^+}} \mathbf{x}^{\operatorname{cont}(S)} \\ &= \sum_{(\lambda^+, S) \in \mathbf{B}} \mathbf{x}^{\operatorname{cont}(S)}, \qquad (13.69.4) \end{aligned}$$

where $\mathbf{B}$ is the set of all pairs (λ^+, S) of a partition λ^+ and a column-strict tableau S of shape λ^+ such that λ^+/λ is a horizontal n-strip. Consider this set $\mathbf{B}$.

We shall now prove that there exists a bijection $\mathbf{i} : \mathbf{A} \to \mathbf{B}$ which has the property that

$$\mathbf{x}^{\operatorname{cont}(T)} x_{i_1} x_{i_2} \cdots x_{i_n} = \mathbf{x}^{\operatorname{cont}(S)} \tag{13.69.5}$$

whenever some $(T, (i_1, i_2, \ldots, i_n)) \in \mathbf{A}$ and $(\lambda^+, S) \in \mathbf{B}$ satisfy $\mathbf{i}((T, (i_1, i_2, \ldots, i_n))) = (\lambda^+, S)$. Once this will be proven, it will immediately follow that $\sum_{(T,(i_1,i_2,\ldots,i_n)) \in \mathbf{A}} \mathbf{x}^{\operatorname{cont}(T)} x_{i_1} x_{i_2} \cdots x_{i_n} = \sum_{(\lambda^+,S) \in \mathbf{B}} \mathbf{x}^{\operatorname{cont}(S)}$, and therefore (13.69.2) will become

$$s_\lambda h_n = \sum_{(T,(i_1,i_2,\ldots,i_n)) \in \mathbf{A}} \mathbf{x}^{\operatorname{cont}(T)} x_{i_1} x_{i_2} \cdots x_{i_n} = \sum_{(\lambda^+,S) \in \mathbf{B}} \mathbf{x}^{\operatorname{cont}(S)} = \sum_{\substack{\lambda^+ : \lambda^+/\lambda \text{ is a} \\ \text{horizontal } n\text{-strip}}} s_{\lambda^+} \qquad \left(\text{by (13.69.4)}\right),$$

and thus (2.7.1) will be proven. Hence, in order to complete the proof of (2.7.1), it is enough to prove that there exists a bijection $\mathbf{i} : \mathbf{A} \to \mathbf{B}$ which has the property (13.69.5).

We construct such a bijection $\mathbf{i} : \mathbf{A} \to \mathbf{B}$ explicitly. Namely, for every $(T, (i_1, i_2, \ldots, i_n)) \in \mathbf{A}$, we define $\mathbf{i}((T, (i_1, i_2, \ldots, i_n)))$ as follows: Construct a sequence $(T_0, T_1, \ldots, T_n)$ of column-strict tableaux recursively: We set $T_0 = T$. For every $k \in \{1, 2, \ldots, n\}$, if T_{k-1} is already defined, we let T_k be the column-strict tableau obtained by applying RS-insertion to the tableau T_{k-1} and the letter i_k. (This RS-insertion also returns a corner cell, but we do not care about it.) Thus, a sequence $(T_0, T_1, \ldots, T_n)$ is defined. We now set $S = T_n$, and let λ^+ be the shape of S. It is easy to see that λ^+/λ is a horizontal n-strip[605]. Thus, $(\lambda^+, S) \in \mathbf{B}$. Now,

[605] *Proof.* Notice that $i_1 \leq i_2 \leq \cdots \leq i_n$ (since $(T, (i_1, i_2, \ldots, i_n)) \in \mathbf{A}$). The tableau S has been obtained from T by applying RS-insertion n times, using the letters $i_1, i_2, \ldots, i_n$ in this order. Each time that we have applied RS-insertion, the shape of our tableau has grown by a new cell. According to Lemma 13.69.1(a), each of these cells (except for the first one) lies in a column further right than the previous one (because the letters $i_1, i_2, \ldots, i_n$ that we inserted satisfy $i_1 \leq i_2 \leq \cdots \leq i_n$). Therefore, no two of these cells lie in the same column. Since these cells are precisely the cells of λ^+/λ (because λ^+ is the shape of S, while λ is the shape of T), this means that no two cells of λ^+/λ lie in the same column. In other words, λ^+/λ is a

set $\mathbf{i}((T,(i_1,i_2,\ldots,i_n))) = (\lambda^+, S)$. We have therefore defined a map $\mathbf{i} : \mathbf{A} \to \mathbf{B}$. It remains to prove that this map $\mathbf{i}$ is a bijection and satisfies (13.69.5).

Proving that the map $\mathbf{i}$ satisfies (13.69.5) is easy[606]. It remains to show that $\mathbf{i}$ is a bijection. We will achieve this by constructing an inverse map.

Indeed, let us define a map $\mathbf{r} : \mathbf{B} \to \mathbf{A}$. For every $(\lambda^+, S) \in \mathbf{B}$, we define $\mathbf{r}((\lambda^+, S))$ as follows: We know that λ^+/λ is a horizontal n-strip (since $(\lambda^+, S) \in \mathbf{B}$). We can thus uniquely label the n cells of λ^+/λ by $c_1, c_2, \ldots, c_n$ from right to left. Consider these cells $c_1, c_2, \ldots, c_n$. Construct a sequence $(S_0, S_1, \ldots, S_n)$ of column-strict tableaux and a sequence $(j_1, j_2, \ldots, j_n)$ of positive integers recursively: We set $S_0 = S$. For every $k \in \{1, 2, \ldots, n\}$, if S_{k-1} is already defined, we apply reverse bumping to the tableau S_{k-1} and its corner cell c_k. We denote the resulting tableau by S_k, and the resulting letter by j_k. [607] Thus, two sequences $(S_0, S_1, \ldots, S_n)$ and $(j_1, j_2, \ldots, j_n)$ are defined. We now set $T = S_n$ and $(i_1, i_2, \ldots, i_n) = (j_n, j_{n-1}, \ldots, j_1)$. So the tableau T is obtained from S by successively applying reverse bumping using

horizontal strip. Since λ^+/λ has precisely n cells (because we have applied RS-insertion exactly n times, gaining precisely one cell every time), this yields that λ^+/λ is a horizontal n-strip, qed.

[606] *Proof.* Assume that some $(T,(i_1,i_2,\ldots,i_n)) \in \mathbf{A}$ and $\left(\lambda^+, S\right) \in \mathbf{B}$ satisfy $\mathbf{i}((T,(i_1,i_2,\ldots,i_n))) = \left(\lambda^+, S\right)$. We need to show that (13.69.5) holds.

According to the definition of $\mathbf{i}((T,(i_1,i_2,\ldots,i_n)))$, the tableau S is obtained by successively applying RS-insertion to the tableau T using the letters $i_1, i_2, \ldots, i_n$. But whenever a tableau V results from applying RS-insertion to a column-strict tableau U and a letter j, the multiset of entries of V is obtained from the multiset of entries of U by tossing in the letter j. Thus, the multiset of entries of S is obtained from the multiset of entries of T by tossing in the n letters $i_1, i_2, \ldots, i_n$ (because S is obtained by successively applying RS-insertion to the tableau T using the letters $i_1, i_2, \ldots, i_n$). Hence, $\mathbf{x}^{\operatorname{cont}(S)} = \mathbf{x}^{\operatorname{cont}(T)} x_{i_1} x_{i_2} \cdots x_{i_n}$. Thus, (13.69.5) is proven.

[607] In order to verify that this definition makes sense, we need to check that c_k is a corner cell of S_{k-1} for each $k \in \{1, 2, \ldots, n\}$. Let us sketch a proof of this fact.

In fact, we shall show a stronger claim:

> *Claim CC:* For each $k \in \{1, 2, \ldots, n\}$, the shape of the tableau S_{k-1} is obtained from λ^+ by removing the corner cells $c_1, c_2, \ldots, c_{k-1}$, and the cell c_k is a corner cell of S_{k-1}.

Proof. We use induction over k.

Induction base: The shape of the tableau S_0 is λ^+ (since $S_0 = S$ has shape λ^+), which is clearly the shape obtained from λ^+ by removing the corner cells $c_1, c_2, \ldots, c_0$. Moreover, the cell c_1 is the rightmost cell of λ^+/λ (since we labelled the cells of λ^+/λ by $c_1, c_2, \ldots, c_n$ from right to left), and thus has no cells of λ^+ to its right in its row; but it also has no cells of λ^+ below it in its column (since any such cell would have to belong to λ^+/λ, but λ^+/λ is a horizontal strip and therefore cannot have two cells in the same column). Thus, c_1 is a corner cell of $S = S_0$. Therefore, Claim CC is proven for $k = 1$. This completes the induction base.

Induction step: Fix some $K \in \{1, 2, \ldots, n-1\}$. Assume that Claim CC is proven for $k = K$. We need to show that Claim CC holds for $k = K + 1$.

We know that Claim CC is proven for $k = K$. In other words, the shape of the tableau S_{K-1} is obtained from λ^+ by removing the corner cells $c_1, c_2, \ldots, c_{K-1}$, and the cell c_K is a corner cell of S_{K-1}. Now, the shape of the tableau S_K is obtained from the shape of S_{K-1} by removing the corner cell c_K (because S_K is obtained by applying reverse bumping to the tableau S_{K-1} and its corner cell c_K), and thus is obtained from λ^+ by removing the corner cells $c_1, c_2, \ldots, c_K$ (since the shape of the tableau S_{K-1} is obtained from λ^+ by removing the corner cells $c_1, c_2, \ldots, c_{K-1}$). Hence, $c_1, c_2, \ldots, c_K$ are not cells of S_K, but c_{K+1} is a cell of S_K (since c_{K+1} is a cell of λ^+/λ, therefore is a cell of λ^+, and is not among the K cells $c_1, c_2, \ldots, c_K$ that were removed). We shall now show that the cell c_{K+1} is a corner cell of S_K.

Indeed, assume the contrary. Thus, c_{K+1} is a cell of S_K, but not a corner cell. Hence, there exists a cell d of S_K that is either the bottom neighbor or the right neighbor of c_{K+1} (that is, it either lies in the same column as c_{K+1} but one step further down, or lies in the same row as c_{K+1} but one step further right). Consider this cell d.

Recall that the cells of λ^+/λ have been labelled $c_1, c_2, \ldots, c_n$ from right to left. Therefore, the cell c_{K+1} belongs to λ^+/λ. Hence, c_{K+1} is not a cell of λ. Therefore, d is not a cell of λ either (since d is either the bottom neighbor or the right neighbor of c_{K+1}). But d is a cell of S_K, thus a cell of λ^+ (since the shape of the tableau S_K is obtained from λ^+ by removing some cells). Hence, the cell d must also belong to λ^+/λ (since d is not a cell of λ). If d was the bottom neighbor of c_{K+1}, we would thus conclude that c_{K+1} and d are two cells of λ^+/λ that lie in the same column; but this is impossible (since λ^+/λ is a horizontal strip and thus cannot have two cells in the same column). Hence, d cannot be the bottom neighbor of c_{K+1}. Thus, d must be the right neighbor of c_{K+1} (since d is either the bottom neighbor or the right neighbor of c_{K+1}). Thus, d lies further right than c_{K+1}. But all cells of λ^+/λ that lie further right than c_{K+1} are $c_1, c_2, \ldots, c_K$ (since the cells of λ^+/λ have been labelled $c_1, c_2, \ldots, c_n$ from right to left). Hence, any cell of λ^+/λ that lies further right than c_{K+1} must be one of $c_1, c_2, \ldots, c_K$. Thus, d must be one of $c_1, c_2, \ldots, c_K$ (since d is a cell of λ^+/λ that lies further right than c_{K+1}). Hence, d cannot be a cell of S_K (since $c_1, c_2, \ldots, c_K$ are not cells of S_K). This contradicts the fact that d is a cell of S_K.

This contradiction shows that our assumption was wrong. Hence, the cell c_{K+1} is a corner cell of S_K.

Thus we have shown that the shape of the tableau S_K is obtained from λ^+ by removing the corner cells $c_1, c_2, \ldots, c_K$, and the cell c_{K+1} is a corner cell of S_K. This proves Claim CC for $k = K + 1$. Thus, the induction step is finished. Claim CC is thus proven. We hence conclude that our definition makes sense.

the corner cells $c_1, c_2, \ldots, c_n$ (in this order), and $j_1, j_2, \ldots, j_n$ are the letters that are obtained from the reverse-bumping procedure. Since reverse bumping is the inverse map to RS-insertion, this yields that we can obtain S back from T by successively applying RS-insertion using the letters $j_n, j_{n-1}, \ldots, j_1$ (that is, the letters $i_1, i_2, \ldots, i_n$), and that these successive RS-insertion steps recover the corner cells $c_n, c_{n-1}, \ldots, c_1$ in this order. Now, it is easy to see that the tableau T has shape λ (because in passing from S to T, we lost the corner cells $c_n, c_{n-1}, \ldots, c_1$, which are exactly the cells of λ^+/λ) and the letters $i_1, i_2, \ldots, i_n$ satisfy $i_1 \leq i_2 \leq \cdots \leq i_n$ [608]. Hence, $(T, (i_1, i_2, \ldots, i_n)) \in \mathbf{A}$. We set $\mathbf{r}((\lambda^+, S)) = (T, (i_1, i_2, \ldots, i_n))$. The map $\mathbf{r} : \mathbf{B} \to \mathbf{A}$ is thus defined. It is now easy to prove that the maps $\mathbf{i}$ and $\mathbf{r}$ are mutually inverse (since RS-insertion and reverse bumping are inverse maps), and thus $\mathbf{i}$ is a bijection. Thus, there exists a bijection $\mathbf{i} : \mathbf{A} \to \mathbf{B}$ which has the property (13.69.5). This completes our proof of (2.7.1). □

The proof of (2.7.2) is almost entirely analogous. We give it for the sake of completeness (but most of it is copypasted material from the proof above).

Alternative proof of (2.7.2). Let λ be a partition, and let $n \in \mathbb{N}$. The definition of s_λ yields

$$s_\lambda = \sum_{\substack{T \text{ is a column-strict} \\ \text{tableau of shape } \lambda}} \mathbf{x}^{\operatorname{cont}(T)}.$$

The definition of h_n yields

$$e_n = \sum_{i_1 < i_2 < \cdots < i_n} x_{i_1} x_{i_2} \cdots x_{i_n} = \sum_{i_n > i_{n-1} > \cdots > i_1} x_{i_n} x_{i_{n-1}} \cdots x_{i_1} = \sum_{i_1 > i_2 > \cdots > i_n} x_{i_1} x_{i_2} \cdots x_{i_n} \tag{13.69.6}$$

(here, we substituted $(i_1, i_2, \ldots, i_n)$ for $(i_n, i_{n-1}, \ldots, i_1)$ in the sum). Multiplying these two identities, we obtain

$$\begin{aligned} s_\lambda e_n &= \sum_{\substack{T \text{ is a column-strict} \\ \text{tableau of shape } \lambda}} \mathbf{x}^{\operatorname{cont}(T)} \sum_{i_1 > i_2 > \cdots > i_n} x_{i_1} x_{i_2} \cdots x_{i_n} \\ &= \sum_{(T, (i_1, i_2, \ldots, i_n)) \in \mathbf{A}} \mathbf{x}^{\operatorname{cont}(T)} x_{i_1} x_{i_2} \cdots x_{i_n}, \qquad (13.69.7) \end{aligned}$$

where $\mathbf{A}$ is the set of all pairs $(T, (i_1, i_2, \ldots, i_n))$ of a column-strict tableau T of shape λ and an n-tuple $(i_1, i_2, \ldots, i_n)$ of positive integers satisfying $i_1 > i_2 > \cdots > i_n$. Consider this set $\mathbf{A}$.

On the other hand, every partition λ^+ satisfies

$$\begin{aligned} s_{\lambda^+} &= \sum_{\substack{T \text{ is a column-strict} \\ \text{tableau of shape } \lambda^+}} \mathbf{x}^{\operatorname{cont}(T)} && \left(\text{by the definition of } s_{\lambda^+}\right) \\ &= \sum_{\substack{S \text{ is a column-strict} \\ \text{tableau of shape } \lambda^+}} \mathbf{x}^{\operatorname{cont}(S)} && \left(\text{here, we renamed the summation index } T \text{ as } S\right). \qquad (13.69.8) \end{aligned}$$

[608] *Proof.* Assume the contrary. Then, we don't have $i_1 \leq i_2 \leq \cdots \leq i_n$. In other words, we don't have $j_n \leq j_{n-1} \leq \cdots \leq j_1$ (since $(i_1, i_2, \ldots, i_n) = (j_n, j_{n-1}, \ldots, j_1)$). Hence, there exists a $k \in \{2, 3, \ldots, n\}$ such that $j_k > j_{k-1}$. Consider this k. Recall that the tableau S_k and the letter j_k were obtained by applying reverse bumping to the tableau S_{k-1} and its corner cell c_k, while the tableau S_{k-1} and the letter j_{k-1} were obtained by applying reverse bumping to the tableau S_{k-2} and its corner cell c_{k-1}. Since reverse bumping is the inverse map to RS-insertion, this entails that conversely, the tableau S_{k-1} and its corner cell c_k are obtained by applying RS-insertion to the tableau S_k and the letter j_k, and the tableau S_{k-2} and its corner cell c_{k-1} are obtained by applying RS-insertion to the tableau S_{k-1} and the letter j_{k-1}. Hence, Lemma 13.69.1(b) (applied to S_k, j_k, j_{k-1}, S_{k-1}, c_k, S_{k-2} and c_{k-1} instead of P, j, j', P', c, P'' and c') yields that the cell c_{k-1} is in the same column as c_k or in a column further left. But this contradicts the fact that c_{k-1} lies in a column further right than c_k (since the cells of λ^+/λ were labelled by $c_1, c_2, \ldots, c_n$ **from right to left**, and lie in different columns). This contradiction completes our proof.

Hence,

$$\sum_{\substack{\lambda^+ : \lambda^+/\lambda \text{ is a} \\ \text{vertical } n\text{-strip}}} \underbrace{s_{\lambda^+}}_{\substack{= \sum_{\substack{S \text{ is a column-strict} \\ \text{tableau of shape } \lambda^+}} \mathbf{x}^{\operatorname{cont}(S)} \\ \text{(by (13.69.8))}}} = \sum_{\substack{\lambda^+ : \lambda^+/\lambda \text{ is a} \\ \text{vertical } n\text{-strip}}} \sum_{\substack{S \text{ is a column-strict} \\ \text{tableau of shape } \lambda^+}} \mathbf{x}^{\operatorname{cont}(S)}$$

$$= \sum_{(\lambda^+, S) \in \mathbf{B}} \mathbf{x}^{\operatorname{cont}(S)}, \qquad (13.69.9)$$

where $\mathbf{B}$ is the set of all pairs (λ^+, S) of a partition λ^+ and a column-strict tableau S of shape λ^+ such that λ^+/λ is a vertical n-strip. Consider this set $\mathbf{B}$.

We shall now prove that there exists a bijection $\mathbf{i} : \mathbf{A} \to \mathbf{B}$ which has the property that

$$\mathbf{x}^{\operatorname{cont}(T)} x_{i_1} x_{i_2} \cdots x_{i_n} = \mathbf{x}^{\operatorname{cont}(S)} \qquad (13.69.10)$$

whenever some $(T, (i_1, i_2, \ldots, i_n)) \in \mathbf{A}$ and $(\lambda^+, S) \in \mathbf{B}$ satisfy $\mathbf{i}((T, (i_1, i_2, \ldots, i_n))) = (\lambda^+, S)$. Once this will be proven, it will immediately follow that $\sum_{(T,(i_1,i_2,\ldots,i_n)) \in \mathbf{A}} \mathbf{x}^{\operatorname{cont}(T)} x_{i_1} x_{i_2} \cdots x_{i_n} = \sum_{(\lambda^+,S)\in\mathbf{B}} \mathbf{x}^{\operatorname{cont}(S)}$, and therefore (13.69.7) will become

$$s_\lambda e_n = \sum_{(T,(i_1,i_2,\ldots,i_n))\in\mathbf{A}} \mathbf{x}^{\operatorname{cont}(T)} x_{i_1} x_{i_2} \cdots x_{i_n} = \sum_{(\lambda^+,S)\in\mathbf{B}} \mathbf{x}^{\operatorname{cont}(S)} = \sum_{\substack{\lambda^+ : \lambda^+/\lambda \text{ is a} \\ \text{vertical } n\text{-strip}}} s_{\lambda^+} \qquad \text{(by (13.69.9))},$$

and thus (2.7.2) will be proven. Hence, in order to complete the proof of (2.7.2), it is enough to prove that there exists a bijection $\mathbf{i} : \mathbf{A} \to \mathbf{B}$ which has the property (13.69.10).

We construct such a bijection $\mathbf{i} : \mathbf{A} \to \mathbf{B}$ explicitly. Namely, for every $(T, (i_1, i_2, \ldots, i_n)) \in \mathbf{A}$, we define $\mathbf{i}((T, (i_1, i_2, \ldots, i_n)))$ as follows: Construct a sequence $(T_0, T_1, \ldots, T_n)$ of column-strict tableaux recursively: We set $T_0 = T$. For every $k \in \{1, 2, \ldots, n\}$, if T_{k-1} is already defined, we let T_k be the column-strict tableau obtained by applying RS-insertion to the tableau T_{k-1} and the letter i_k. (This RS-insertion also returns a corner cell, but we do not care about it.) Thus, a sequence $(T_0, T_1, \ldots, T_n)$ is defined. We now set $S = T_n$, and let λ^+ be the shape of S. It is easy to see that λ^+/λ is a vertical n-strip[609]. Thus, $(\lambda^+, S) \in \mathbf{B}$. Now, set $\mathbf{i}((T, (i_1, i_2, \ldots, i_n))) = (\lambda^+, S)$. We have therefore defined a map $\mathbf{i} : \mathbf{A} \to \mathbf{B}$. It remains to prove that this map $\mathbf{i}$ is a bijection and satisfies (13.69.10).

Proving that the map $\mathbf{i}$ satisfies (13.69.10) is easy[610]. It remains to show that $\mathbf{i}$ is a bijection. We will achieve this by constructing an inverse map.

Indeed, let us define a map $\mathbf{r} : \mathbf{B} \to \mathbf{A}$. For every $(\lambda^+, S) \in \mathbf{B}$, we define $\mathbf{r}((\lambda^+, S))$ as follows: We know that λ^+/λ is a vertical n-strip (since $(\lambda^+, S) \in \mathbf{B}$). We can thus uniquely label the n cells of λ^+/λ by $c_1, c_2, \ldots, c_n$ from bottom to top. Consider these cells $c_1, c_2, \ldots, c_n$. Construct a sequence $(S_0, S_1, \ldots, S_n)$ of column-strict tableaux and a sequence $(j_1, j_2, \ldots, j_n)$ of positive integers recursively: We set $S_0 = S$. For every $k \in \{1, 2, \ldots, n\}$, if S_{k-1} is already defined, we apply reverse bumping to the tableau S_{k-1} and its corner cell c_k. We denote the resulting tableau by S_k, and the resulting letter by j_k. [611] Thus,

[609] *Proof.* Notice that $i_1 > i_2 > \cdots > i_n$ (since $(T, (i_1, i_2, \ldots, i_n)) \in \mathbf{A}$). The tableau S has been obtained from T by applying RS-insertion n times, using the letters $i_1, i_2, \ldots, i_n$ in this order. Each time that we have applied RS-insertion, the shape of our tableau has grown by a new cell. According to Lemma 13.69.1(b), each of these cells (except for the first one) lies in a row further down than the previous one (because the letters $i_1, i_2, \ldots, i_n$ that we inserted satisfy $i_1 > i_2 > \cdots > i_n$). Therefore, no two of these cells lie in the same row. Since these cells are precisely the cells of λ^+/λ (because λ^+ is the shape of S, while λ is the shape of T), this means that no two cells of λ^+/λ lie in the same row. In other words, λ^+/λ is a vertical strip. Since λ^+/λ has precisely n cells (because we have applied RS-insertion exactly n times, gaining precisely one cell every time), this yields that λ^+/λ is a vertical n-strip, qed.

[610] *Proof.* Assume that some $(T, (i_1, i_2, \ldots, i_n)) \in \mathbf{A}$ and $(\lambda^+, S) \in \mathbf{B}$ satisfy $\mathbf{i}((T, (i_1, i_2, \ldots, i_n))) = (\lambda^+, S)$. We need to show that (13.69.10) holds.

According to the definition of $\mathbf{i}((T, (i_1, i_2, \ldots, i_n)))$, the tableau S is obtained by successively applying RS-insertion to the tableau T using the letters $i_1, i_2, \ldots, i_n$. But whenever a tableau V results from applying RS-insertion to a column-strict tableau U and a letter j, the multiset of entries of V is obtained from the multiset of entries of U by tossing in the letter j. Thus, the multiset of entries of S is obtained from the multiset of entries of T by tossing in the n letters $i_1, i_2, \ldots, i_n$ (because S is obtained by successively applying RS-insertion to the tableau T using the letters $i_1, i_2, \ldots, i_n$). Hence, $\mathbf{x}^{\operatorname{cont}(S)} = \mathbf{x}^{\operatorname{cont}(T)} x_{i_1} x_{i_2} \cdots x_{i_n}$. Thus, (13.69.10) is proven.

[611] One again has to check that this is well-defined. The proof is very similar to the analogous argument in our proof of (2.7.1), and is left to the reader.

two sequences $(S_0, S_1, \ldots, S_n)$ and $(j_1, j_2, \ldots, j_n)$ are defined. We now set $T = S_n$ and $(i_1, i_2, \ldots, i_n) = (j_n, j_{n-1}, \ldots, j_1)$. So the tableau T is obtained from S by successively applying reverse bumping using the corner cells $c_1, c_2, \ldots, c_n$ (in this order), and $j_1, j_2, \ldots, j_n$ are the letters that are obtained from the reverse-bumping procedure. Since reverse bumping is the inverse map to RS-insertion, this yields that we can obtain S back from T by successively applying RS-insertion using the letters $j_n, j_{n-1}, \ldots, j_1$ (that is, the letters $i_1, i_2, \ldots, i_n$), and that these successive RS-insertion steps recover the corner cells $c_n, c_{n-1}, \ldots, c_1$ in this order. Now, it is easy to see that the tableau T has shape λ (because in passing from S to T, we lost the corner cells $c_n, c_{n-1}, \ldots, c_1$, which are exactly the cells of λ^+/λ) and the letters $i_1, i_2, \ldots, i_n$ satisfy $i_1 > i_2 > \cdots > i_n$ [612]. Hence, $(T, (i_1, i_2, \ldots, i_n)) \in \mathbf{A}$. We set $\mathbf{r}((\lambda^+, S)) = (T, (i_1, i_2, \ldots, i_n))$. The map $\mathbf{r} : \mathbf{B} \to \mathbf{A}$ is thus defined. It is now easy to prove that the maps $\mathbf{i}$ and $\mathbf{r}$ are mutually inverse (since RS-insertion and reverse bumping are inverse maps), and thus $\mathbf{i}$ is a bijection. Thus, there exists a bijection $\mathbf{i} : \mathbf{A} \to \mathbf{B}$ which has the property (13.69.10). This completes our proof of (2.7.2). □

Now, both (2.7.1) and (2.7.2) are proven. Hence, Theorem 2.7.1 is proven again.

13.70. **Solution to Exercise 2.7.8.** *Solution to Exercise 2.7.8.* Before we start solving any specific part of this exercise, let us state some general properties of determinants (and prove some of them):

- Every $m \in \mathbb{N}$ and every matrix $(\alpha_{i,j})_{i,j=1,2,\ldots,m} \in A^{m\times m}$ satisfy

$$\det\left((\alpha_{i,j})_{i,j=1,2,\ldots,m}\right) = \sum_{\sigma\in\mathfrak{S}_m} (-1)^\sigma \prod_{i=1}^m \alpha_{i,\sigma(i)}. \tag{13.70.1}$$

(This is simply the explicit formula for the determinant of a matrix as a sum over permutations.)

- If a positive integer m and a matrix $(\alpha_{i,j})_{i,j=1,2,\ldots,m} \in A^{m\times m}$ are such that every $j \in \{1, 2, ..., m-1\}$ satisfies $\alpha_{m,j} = 0$, then

$$\det\left((\alpha_{i,j})_{i,j=1,2,\ldots,m}\right) = \alpha_{m,m} \cdot \det\left((\alpha_{i,j})_{i,j=1,2,\ldots,m-1}\right). \tag{13.70.2}$$

[613]

- Every positive integer m and every matrix $(\alpha_{i,j})_{i,j=1,2,\ldots,m} \in A^{m\times m}$ satisfy

$$\det\left((\alpha_{i,j}\alpha_{m,m} - \alpha_{i,m}\alpha_{m,j})_{i,j=1,2,\ldots,m-1}\right) = \alpha_{m,m}^{m-2} \cdot \det\left((\alpha_{i,j})_{i,j=1,2,\ldots,m}\right) \tag{13.70.3}$$

if $\alpha_{m,m}$ is an invertible element of A.

Proof of (13.70.3)*:* Let m be a positive integer. Let $(\alpha_{i,j})_{i,j=1,2,\ldots,m} \in A^{m\times m}$ be a matrix such that $\alpha_{m,m}$ is an invertible element of A.

[612] *Proof.* Assume the contrary. Then, we don't have $i_1 > i_2 > \cdots > i_n$. In other words, we don't have $j_n > j_{n-1} > \cdots > j_1$ (since $(i_1, i_2, \ldots, i_n) = (j_n, j_{n-1}, \ldots, j_1)$). Hence, there exists a $k \in \{2, 3, \ldots, n\}$ such that $j_k \leq j_{k-1}$. Consider this k. Recall that the tableau S_k and the letter j_k were obtained by applying reverse bumping to the tableau S_{k-1} and its corner cell c_k, while the tableau S_{k-1} and the letter j_{k-1} were obtained by applying reverse bumping to the tableau S_{k-2} and its corner cell c_{k-1}. Since reverse bumping is the inverse map to RS-insertion, this entails that conversely, the tableau S_{k-1} and its corner cell c_k are obtained by applying RS-insertion to the tableau S_k and the letter j_k, and the tableau S_{k-2} and its corner cell c_{k-1} are obtained by applying RS-insertion to the tableau S_{k-1} and the letter j_{k-1}. Hence, Lemma 13.69.1(a) (applied to S_k, j_k, j_{k-1}, S_{k-1}, c_k, S_{k-2} and c_{k-1} instead of P, j, j', P', c, P'' and c') yields that the cell c_{k-1} is in the same row as c_k or in a row further up. But this contradicts the fact that c_{k-1} lies in a row further down than c_k (since the cells of λ^+/λ were labelled by $c_1, c_2, \ldots, c_n$ **from bottom to top**, and lie in different rows). This contradiction completes our proof.

[613] In fact, the condition that every $j \in \{1, 2, ..., m-1\}$ satisfies $\alpha_{m,j} = 0$ means that all entries of the m-th row of the $m \times m$-matrix $(\alpha_{i,j})_{i,j=1,2,\ldots,m}$ are zeroes apart from (possibly) the last entry. Hence, applying Laplace expansion to the determinant $\det\left((\alpha_{i,j})_{i,j=1,2,\ldots,m}\right)$ of this matrix yields a sum of products, all of which are zero apart from

$$\alpha_{m,m} \cdot \underbrace{\left(\text{the } (m,m)\text{-th cofactor of the matrix } (\alpha_{i,j})_{i,j=1,2,\ldots,m}\right)}_{=(-1)^{m+m}\cdot\det\left((\alpha_{i,j})_{i,j=1,2,\ldots,m-1}\right)=\det\left((\alpha_{i,j})_{i,j=1,2,\ldots,m-1}\right)} = \alpha_{m,m} \cdot \det\left((\alpha_{i,j})_{i,j=1,2,\ldots,m-1}\right).$$

For every $(i,j) \in \{1,2,...,m\}^2$, define an element $\beta_{i,j}$ of A by

$$\beta_{i,j} = \begin{cases} \alpha_{m,m}, & \text{if } i = j; \\ -\alpha_{m,j}, & \text{if } i = m \text{ and } j \neq m \end{cases} . \tag{13.70.4}$$

Then, the matrix $(\beta_{i,j})_{i,j=1,2,...,m}$ is lower-triangular. Since the determinant of a lower-triangular matrix equals the product of its diagonal entries, we thus have

$$\det\left((\beta_{i,j})_{i,j=1,2,...,m}\right) = \prod_{i=1}^{m} \underbrace{\beta_{i,i}}_{\substack{=\alpha_{m,m} \\ \text{(this is} \\ \text{easily seen)}}} = \prod_{i=1}^{m} \alpha_{m,m} = \alpha_{m,m}^m . \tag{13.70.5}$$

But since the determinant of a product of two square matrices equals the product of their determinants, we have

$$\begin{aligned} \det\left((\alpha_{i,j})_{i,j=1,2,...,m} \cdot (\beta_{i,j})_{i,j=1,2,...,m}\right) &= \det\left((\alpha_{i,j})_{i,j=1,2,...,m}\right) \cdot \underbrace{\det\left((\beta_{i,j})_{i,j=1,2,...,m}\right)}_{\substack{=\alpha_{m,m}^m \\ \text{(by (13.70.5))}}} \\ &= \det\left((\alpha_{i,j})_{i,j=1,2,...,m}\right) \cdot \alpha_{m,m}^m . \end{aligned}$$

Since $(\alpha_{i,j})_{i,j=1,2,...,m} \cdot (\beta_{i,j})_{i,j=1,2,...,m} = \left(\sum_{k=1}^m \alpha_{i,k}\beta_{k,j}\right)_{i,j=1,2,...,m}$, this rewrites as

$$\det\left(\left(\sum_{k=1}^m \alpha_{i,k}\beta_{k,j}\right)_{i,j=1,2,...,m}\right) = \det\left((\alpha_{i,j})_{i,j=1,2,...,m}\right) \cdot \alpha_{m,m}^m . \tag{13.70.6}$$

Now, for every $(i,j) \in \{1,2,...,m\}^2$, define an element $\gamma_{i,j}$ of A by

$$\gamma_{i,j} = \sum_{k=1}^m \alpha_{i,k}\beta_{k,j} . \tag{13.70.7}$$

Then,

$$\det\left((\gamma_{i,j})_{i,j=1,2,...,m}\right) = \det\left(\left(\sum_{k=1}^m \alpha_{i,k}\beta_{k,j}\right)_{i,j=1,2,...,m}\right) = \det\left((\alpha_{i,j})_{i,j=1,2,...,m}\right) \cdot \alpha_{m,m}^m \tag{13.70.8}$$

(by (13.70.6)).

However, for every (i,j), we can simplify the expression for $\gamma_{i,j}$ given by (13.70.7) by plugging in the definition of $\beta_{k,j}$ (which guarantees that no more than two of the terms of the sum will be nonzero). We obtain

$$\gamma_{i,j} = \begin{cases} \alpha_{i,j}\alpha_{m,m} - \alpha_{i,m}\alpha_{m,j}, & \text{if } i \neq m \text{ and } j \neq m; \\ 0, & \text{if } i = m \text{ and } j \neq m; \\ \alpha_{i,m}\alpha_{m,m}, & \text{if } j = m \end{cases} . \tag{13.70.9}$$

In particular, this shows that every $j \in \{1,2,...,m-1\}$ satisfies $\gamma_{m,j} = 0$. Hence, (13.70.2) (applied to $\gamma_{i,j}$ instead of $\alpha_{i,j}$) yields

$$\begin{aligned} \det\left((\gamma_{i,j})_{i,j=1,2,...,m}\right) &= \underbrace{\gamma_{m,m}}_{\substack{=\alpha_{m,m}\alpha_{m,m} \\ \text{(by (13.70.9), since } m=m)}} \cdot \det\left(\left(\underbrace{\gamma_{i,j}}_{\substack{=\alpha_{i,j}\alpha_{m,m}-\alpha_{i,m}\alpha_{m,j} \\ \text{(by (13.70.9), since } i\neq m \text{ and } j\neq m)}}\right)_{i,j=1,2,...,m-1}\right) \\ &= \alpha_{m,m}\alpha_{m,m} \cdot \det\left((\alpha_{i,j}\alpha_{m,m} - \alpha_{i,m}\alpha_{m,j})_{i,j=1,2,...,m-1}\right) \\ &= \alpha_{m,m}^2 \cdot \det\left((\alpha_{i,j}\alpha_{m,m} - \alpha_{i,m}\alpha_{m,j})_{i,j=1,2,...,m-1}\right) . \end{aligned}$$

Compared with (13.70.8), this yields

$$\alpha_{m,m}^2 \cdot \det\left(\left(\alpha_{i,j}\alpha_{m,m} - \alpha_{i,m}\alpha_{m,j}\right)_{i,j=1,2,\ldots,m-1}\right) = \det\left(\left(\alpha_{i,j}\right)_{i,j=1,2,\ldots,m}\right) \cdot \alpha_{m,m}^m.$$

We can divide both sides of this equality by $\alpha_{m,m}^2$ (since $\alpha_{m,m}$ is invertible in A), and thus obtain

$$\det\left(\left(\alpha_{i,j}\alpha_{m,m} - \alpha_{i,m}\alpha_{m,j}\right)_{i,j=1,2,\ldots,m-1}\right) = \alpha_{m,m}^{m-2} \cdot \det\left(\left(\alpha_{i,j}\right)_{i,j=1,2,\ldots,m}\right).$$

This proves (13.70.3).

[*Remark:* The equality (13.70.3) holds even without requiring that $\alpha_{m,m}$ be invertible, if we have $m \geq 2$ (of course, if m is not ≥ 2, then the $\alpha_{m,m}^{m-2}$ on the right hand side of (13.70.3) does not make sense unless $\alpha_{m,m}$ is invertible). There are several ways to see why this is so. One of these ways proceeds as follows: First of all, one should notice that our above proof of (13.70.3) works without requiring that $\alpha_{m,m}$ be invertible, as long as $\alpha_{m,m}$ is a non-zero-divisor[614] in A and we have $m \geq 2$. However, for any fixed $m \geq 2$, the equality (13.70.3) is a polynomial identity in the elements $\alpha_{i,j}$ of A; thus, it suffices to prove it when $\alpha_{i,j}$ are distinct indeterminates $X_{i,j}$ in the polynomial ring $\mathbb{Z}\left[X_{i,j} \mid (i,j) \in \{1,2,...,m\}^2\right]$. But in this case, $\alpha_{m,m}$ is clearly a non-zero-divisor, and so our above proof applies. (An alternative approach would be to replace $\alpha_{m,m}$ by $X + \alpha_{m,m}$ in the polynomial ring $A[X]$; again, $X + \alpha_{m,m}$ is a non-zero-divisor even if $\alpha_{m,m}$ is not.)]

Now, rather than solve parts (a) and (b) of the exercise separately, we are going to prove a result from which both of these parts will easily follow:

Proposition 13.70.1. *Let A be a commutative ring. Let $n \in \mathbb{N}$. For every $i \in \{1,2,...,n\}$, let a_i, b_i, c_i and d_i be four elements of A. Assume that $a_i d_j - b_i c_j$ is an invertible element of A for every $i \in \{1,2,...,n\}$ and $j \in \{1,2,...,n\}$. Then,*

$$\det\left(\left(\frac{1}{a_i d_j - b_i c_j}\right)_{i,j=1,2,\ldots,n}\right) = \frac{\prod_{1\leq j<i\leq n}\left(\left(a_i b_j - a_j b_i\right)\left(c_j d_i - c_i d_j\right)\right)}{\prod_{(i,j)\in\{1,2,\ldots,n\}^2}\left(a_i d_j - b_i c_j\right)}$$

Proof of Proposition 13.70.1. We prove this by induction over n. The base case ($n = 0$) is obvious, as it claims an equality between the determinant of a 0×0-matrix (defined to be 1) and the ratio of two empty products (thus $\frac{1}{1} = 1$). For the induction step, we fix some positive integer m, and we set out to prove the equality

$$\det\left(\left(\frac{1}{a_i d_j - b_i c_j}\right)_{i,j=1,2,\ldots,m}\right) = \frac{\prod_{1\leq j<i\leq m}\left(\left(a_i b_j - a_j b_i\right)\left(c_j d_i - c_i d_j\right)\right)}{\prod_{(i,j)\in\{1,2,\ldots,m\}^2}\left(a_i d_j - b_i c_j\right)}, \tag{13.70.10}$$

assuming that we already know

$$\det\left(\left(\frac{1}{a_i d_j - b_i c_j}\right)_{i,j=1,2,\ldots,m-1}\right) = \frac{\prod_{1\leq j<i\leq m-1}\left(\left(a_i b_j - a_j b_i\right)\left(c_j d_i - c_i d_j\right)\right)}{\prod_{(i,j)\in\{1,2,\ldots,m-1\}^2}\left(a_i d_j - b_i c_j\right)} \tag{13.70.11}$$

to be true.

Now, we know that $\frac{1}{a_m d_m - b_m c_m}$ is an invertible element of A (because it is the inverse of $a_m d_m - b_m c_m$). Thus, (13.70.3) (applied to $\alpha_{i,j} = \frac{1}{a_i d_j - b_i c_j}$) yields

$$\begin{aligned}&\det\left(\left(\frac{1}{a_i d_j - b_i c_j} \cdot \frac{1}{a_m d_m - b_m c_m} - \frac{1}{a_i d_m - b_i c_m} \cdot \frac{1}{a_m d_j - b_m c_j}\right)_{i,j=1,2,\ldots,m-1}\right)\\ &= \left(\frac{1}{a_m d_m - b_m c_m}\right)^{m-2} \cdot \det\left(\left(\frac{1}{a_i d_j - b_i c_j}\right)_{i,j=1,2,\ldots,m}\right).\end{aligned}$$

[614] A *non-zero-divisor* in a commutative ring B means an element $b \in B$ such that every element $c \in B$ satisfying $bc = 0$ must satisfy $c = 0$.

Solving this for $\det\left(\left(\frac{1}{a_id_j-b_ic_j}\right)_{i,j=1,2,...,m}\right)$, we obtain

$$\det\left(\left(\frac{1}{a_id_j-b_ic_j}\right)_{i,j=1,2,...,m}\right)$$
(13.70.12)
$$=(a_md_m-b_mc_m)^{m-2}\cdot\det\left(\left(\frac{1}{a_id_j-b_ic_j}\cdot\frac{1}{a_md_m-b_mc_m}-\frac{1}{a_id_m-b_ic_m}\cdot\frac{1}{a_md_j-b_mc_j}\right)_{i,j=1,2,...,m-1}\right).$$

But straightforward computations show that every $(i,j)\in\{1,2,...,m\}^2$ satisfy

$$\frac{1}{a_id_j-b_ic_j}\cdot\frac{1}{a_md_m-b_mc_m}-\frac{1}{a_id_m-b_ic_m}\cdot\frac{1}{a_md_j-b_mc_j}$$
$$=\frac{c_jd_m-c_md_j}{(a_md_m-b_mc_m)(a_md_j-b_mc_j)}\cdot\frac{a_mb_i-a_ib_m}{a_id_m-b_ic_m}\cdot\frac{1}{a_id_j-b_ic_j}.$$

Hence, the matrix

$$\left(\frac{1}{a_id_j-b_ic_j}\cdot\frac{1}{a_md_m-b_mc_m}-\frac{1}{a_id_m-b_ic_m}\cdot\frac{1}{a_md_j-b_mc_j}\right)_{i,j=1,2,...,m-1}$$

can be rewritten as

$$\left(\frac{c_jd_m-c_md_j}{(a_md_m-b_mc_m)(a_md_j-b_mc_j)}\cdot\frac{a_mb_i-a_ib_m}{a_id_m-b_ic_m}\cdot\frac{1}{a_id_j-b_ic_j}\right)_{i,j=1,2,...,m-1}.$$

This means that this matrix can be obtained from the matrix $\left(\frac{1}{a_id_j-b_ic_j}\right)_{i,j=1,2,...,m-1}$ by multiplying every row with $\frac{a_mb_i-a_ib_m}{a_id_m-b_ic_m}$, where i is the index of this row, and then multiplying every column with $\frac{c_jd_m-c_md_j}{(a_md_m-b_mc_m)(a_md_j-b_mc_j)}$, where j is the index of the column. As a consequence, the determinant of this matrix is

$$\left(\prod_{j=1}^{m-1}\frac{c_jd_m-c_md_j}{(a_md_m-b_mc_m)(a_md_j-b_mc_j)}\right)\cdot\left(\prod_{i=1}^{m-1}\frac{a_mb_i-a_ib_m}{a_id_m-b_ic_m}\right)\cdot\det\left(\left(\frac{1}{a_id_j-b_ic_j}\right)_{i,j=1,2,...,m-1}\right)$$

(because when a row of a matrix is multiplied by a scalar, the determinant of the matrix gets multiplied by the same scalar, and the same rule holds for columns). Hence, we have shown that

$$\det\left(\left(\frac{1}{a_id_j-b_ic_j}\cdot\frac{1}{a_md_m-b_mc_m}-\frac{1}{a_id_m-b_ic_m}\cdot\frac{1}{a_md_j-b_mc_j}\right)_{i,j=1,2,...,m-1}\right)$$
$$=\left(\prod_{j=1}^{m-1}\frac{c_jd_m-c_md_j}{(a_md_m-b_mc_m)(a_md_j-b_mc_j)}\right)\cdot\left(\prod_{i=1}^{m-1}\frac{a_mb_i-a_ib_m}{a_id_m-b_ic_m}\right)\cdot\underbrace{\det\left(\left(\frac{1}{a_id_j-b_ic_j}\right)_{i,j=1,2,...,m-1}\right)}_{\substack{=\frac{\prod_{1\leq j<i\leq m-1}((a_ib_j-a_jb_i)(c_jd_i-c_id_j))}{\prod_{(i,j)\in\{1,2,...,m-1\}^2}(a_id_j-b_ic_j)}\\ \text{(by (13.70.11))}}}$$
$$=\left(\prod_{j=1}^{m-1}\frac{c_jd_m-c_md_j}{(a_md_m-b_mc_m)(a_md_j-b_mc_j)}\right)\cdot\left(\prod_{i=1}^{m-1}\frac{a_mb_i-a_ib_m}{a_id_m-b_ic_m}\right)\cdot\frac{\prod_{1\leq j<i\leq m-1}((a_ib_j-a_jb_i)(c_jd_i-c_id_j))}{\prod_{(i,j)\in\{1,2,...,m-1\}^2}(a_id_j-b_ic_j)}.$$

We can now plug this into (13.70.12) and make straightforward simplifications (splitting apart and pulling together products), and obtain (13.70.10). Thus, the induction step is complete, and Proposition 13.70.1 is proven.

Now, let us solve the exercise.

(a) *First solution of Exercise 2.7.8(a):* Exercise 2.7.8(a) follows from Proposition 13.70.1 (applied to a_i, 1, b_i and 1 instead of a_i, b_i, c_i and d_i).

Second solution of Exercise 2.7.8(a) (sketched): The statement of Exercise 2.7.8(a) is an identity between two rational functions in the variables $a_1, a_2, ..., a_n, b_1, b_2, ..., b_n$, and is easily seen to be equivalent to a polynomial identity in these variables (by multiplying both sides through with the common denominator $\prod_{(i,j)\in\{1,2,...,n\}^2} (a_i - b_j)$). It is well-known that such identities need only be checked on complex numbers to ensure that they hold for any elements of any ring. So we only need to prove the statement of Exercise 2.7.8(a) in the case when $A = \mathbb{C}$. But the statement is well-known in this case (see, e.g., [69, Lemma 5.15.3 (Lemma 4.48 in the arXiv version)] and various other sources for the proof).

(b) *First solution of Exercise 2.7.8(b):* Exercise 2.7.8(b) follows from Proposition 13.70.1 (applied to 1, a_i, b_i and 1 instead of a_i, b_i, c_i and d_i).

Second solution of Exercise 2.7.8(b): Just as in the Second solution of Exercise 2.7.8(a), we can see that it is enough to solve Exercise 2.7.8(b) in the case when $A = \mathbb{C}$. But in this case, the statement of this exercise is well-known, and proven, e.g., in [69, Corollary 5.15.4 (Corollary 4.49 in the arXiv version)] and [44, Cauchy's lemma, p. 18].

Remark: One could also derive the statement of Exercise 2.7.8(b) from Exercise 2.7.8(a) by applying the latter to $1/a_i$ instead of a_i, after first WLOG assuming that the a_i are invertible (but one needs to justify this WLOG assumption).

(c) *Alternative proof of Theorem 2.5.1.* We are going to show that, for every $n \in \mathbb{N}$, we have

$$\prod_{i,j=1}^{n} (1 - x_i y_j)^{-1} = \sum_{\lambda\in\mathrm{Par}} s_\lambda (x_1, x_2, ..., x_n) s_\lambda (y_1, y_2, ..., y_n) \tag{13.70.13}$$

in the ring $\mathbf{k}[[x_1, x_2, ..., x_n, y_1, y_2, ..., y_n]]$. Once this is proven, Theorem 2.5.1 will easily follow.[615]

So let $n \in \mathbb{N}$. For every $i \in \{1, 2, ..., n\}$ and $j \in \{1, 2, ..., n\}$, the element $1 - x_i y_j$ of the ring $\mathbf{k}[[x_1, x_2, ..., x_n, y_1, y_2, ..., y_n]]$ is invertible. Hence, Exercise 2.7.8(b) (applied to $A = \mathbf{k}[[x_1, x_2, ..., x_n, y_1, y_2, ..., y_n]]$, $a_i = x_i$ and $b_j = y_j$) yields

$$\det\left(\left(\frac{1}{1 - x_i y_j}\right)_{i,j=1,2,...,n}\right) = \frac{\prod_{1\le j<i\le n} ((x_i - x_j)(y_i - y_j))}{\prod_{(i,j)\in\{1,2,...,n\}^2} (1 - x_i y_j)} = \frac{\prod_{1\le i<j\le n} ((x_i - x_j)(y_i - y_j))}{\prod_{(i,j)\in\{1,2,...,n\}^2} (1 - x_i y_j)}$$

$$= \left(\prod_{1\le i<j\le n} (x_i - x_j)\right) \cdot \left(\prod_{1\le i<j\le n} (y_i - y_j)\right) \cdot \prod_{i,j=1}^{n} \frac{1}{1 - x_i y_j} \tag{13.70.14}$$

in the ring $\mathbf{k}[[x_1, x_2, ..., x_n, y_1, y_2, ..., y_n]]$.

We will use the notations of Definition 2.6.2 and Proposition 2.6.4 (but we do not require $\mathbf{k}$ to be $\mathbb{Z}$ or a field of characteristic not equal to 2 as was done in Proposition 2.6.4). We have $a_\rho = \prod_{1\le i<j\le n} (x_i - x_j)$ (as

[615] *Proof.* Assume that (13.70.13) is proven for all $n \in \mathbb{N}$. We need to show that

$$\prod_{i,j=1}^{\infty} (1 - x_i y_j)^{-1} = \sum_{\lambda\in\mathrm{Par}} s_\lambda (\mathbf{x}) s_\lambda (\mathbf{y})$$

in the ring $\mathbf{k}[[\mathbf{x}, \mathbf{y}]] = \mathbf{k}[[x_1, x_2, x_3, ..., y_1, y_2, y_3, ...]]$. In order to do so, it is clearly enough to prove that for any two weak compositions α and β, the coefficient of $\mathbf{x}^\alpha \mathbf{y}^\beta$ in $\prod_{i,j=1}^{\infty} (1 - x_i y_j)^{-1}$ equals the coefficient of $\mathbf{x}^\alpha \mathbf{y}^\beta$ in $\sum_{\lambda\in\mathrm{Par}} s_\lambda (\mathbf{x}) s_\lambda (\mathbf{y})$. So fix two weak compositions α and β. Write α and β as $\alpha = (\alpha_1, \alpha_2, \alpha_3, ...)$ and $\beta = (\beta_1, \beta_2, \beta_3, ...)$. Choose some $n \in \mathbb{N}$ such that every integer $m > n$ satisfies $\alpha_m = \beta_m = 0$. (Such an m clearly exists, since α and β are finitely supported.) Then, the coefficient of $\mathbf{x}^\alpha \mathbf{y}^\beta$ in $\prod_{i,j=1}^{\infty} (1 - x_i y_j)^{-1}$ equals the coefficient of $x_1^{\alpha_1} x_2^{\alpha_2} ... x_n^{\alpha_n} y_1^{\beta_1} y_2^{\beta_2} ... y_n^{\beta_n}$ in $\prod_{i,j=1}^{n} (1 - x_i y_j)^{-1}$, whereas the coefficient of $\mathbf{x}^\alpha \mathbf{y}^\beta$ in $\sum_{\lambda\in\mathrm{Par}} s_\lambda (\mathbf{x}) s_\lambda (\mathbf{y})$ equals the coefficient of $x_1^{\alpha_1} x_2^{\alpha_2} ... x_n^{\alpha_n} y_1^{\beta_1} y_2^{\beta_2} ... y_n^{\beta_n}$ in $\sum_{\lambda\in\mathrm{Par}} s_\lambda (x_1, x_2, ..., x_n) s_\lambda (y_1, y_2, ..., y_n)$. Since the coefficients of $x_1^{\alpha_1} x_2^{\alpha_2} ... x_n^{\alpha_n} y_1^{\beta_1} y_2^{\beta_2} ... y_n^{\beta_n}$ in $\prod_{i,j=1}^{n} (1 - x_i y_j)^{-1}$ and in $\sum_{\lambda\in\mathrm{Par}} s_\lambda (x_1, x_2, ..., x_n) s_\lambda (y_1, y_2, ..., y_n)$ are equal (because of (13.70.13)), this shows that the coefficients of $\mathbf{x}^\alpha \mathbf{y}^\beta$ in $\prod_{i,j=1}^{\infty} (1 - x_i y_j)^{-1}$ and in $\sum_{\lambda\in\mathrm{Par}} s_\lambda (\mathbf{x}) s_\lambda (\mathbf{y})$ are equal; but this is exactly what we need to prove. Thus, Theorem 2.5.1 follows if (13.70.13) is proven.

proven in the proof of Proposition 2.6.4). Thus, $a_\rho(x_1,x_2,...,x_n) = \prod_{1\leq i<j\leq n}(x_i-x_j)$ and $a_\rho(y_1,y_2,...,y_n) = \prod_{1\leq i<j\leq n}(y_i-y_j)$. Thus, (13.70.14) becomes

$$\begin{aligned}
&\det\left(\left(\frac{1}{1-x_iy_j}\right)_{i,j=1,2,...,n}\right)\\
&=\underbrace{\left(\prod_{1\leq i<j\leq n}(x_i-x_j)\right)}_{=a_\rho(x_1,x_2,...,x_n)}\cdot\underbrace{\left(\prod_{1\leq i<j\leq n}(y_i-y_j)\right)}_{=a_\rho(y_1,y_2,...,y_n)}\cdot\prod_{i,j=1}^{n}\underbrace{\frac{1}{1-x_iy_j}}_{=(1-x_iy_j)^{-1}}\\
&=a_\rho(x_1,x_2,...,x_n)\cdot a_\rho(y_1,y_2,...,y_n)\cdot\prod_{i,j=1}^{n}(1-x_iy_j)^{-1}. \qquad (13.70.15)
\end{aligned}$$

On the other hand, (13.70.1) (applied to $m=n$ and $\alpha_{i,j}=\frac{1}{1-x_iy_j}$) yields

$$\begin{aligned}
\det\left(\left(\frac{1}{1-x_iy_j}\right)_{i,j=1,2,...,n}\right) &= \sum_{\sigma\in S_n}(-1)^\sigma\prod_{i=1}^{n}\underbrace{\frac{1}{1-x_iy_{\sigma(i)}}}_{\substack{=\sum_{k\in\mathbb{N}}\left(x_iy_{\sigma(i)}\right)^k\\ \text{(by the formula for the geometric series)}}}\\
&=\sum_{\sigma\in S_n}(-1)^\sigma\underbrace{\prod_{i=1}^{n}\sum_{k\in\mathbb{N}}\left(x_iy_{\sigma(i)}\right)^k}_{\substack{=\sum_{(k_1,k_2,...,k_n)\in\mathbb{N}^n}\left(x_1y_{\sigma(1)}\right)^{k_1}\left(x_2y_{\sigma(2)}\right)^{k_2}...\left(x_ny_{\sigma(n)}\right)^{k_n}\\ \text{(by the product rule)}}}\\
&=\sum_{\sigma\in S_n}(-1)^\sigma\sum_{(k_1,k_2,...,k_n)\in\mathbb{N}^n}\underbrace{\left(x_1y_{\sigma(1)}\right)^{k_1}\left(x_2y_{\sigma(2)}\right)^{k_2}...\left(x_ny_{\sigma(n)}\right)^{k_n}}_{\substack{=\left(x_1^{k_1}y_{\sigma(1)}^{k_1}\right)\left(x_2^{k_2}y_{\sigma(2)}^{k_2}\right)...\left(x_n^{k_n}y_{\sigma(n)}^{k_n}\right)\\ =\left(x_1^{k_1}x_2^{k_2}...x_n^{k_n}\right)\left(y_{\sigma(1)}^{k_1}y_{\sigma(2)}^{k_2}...y_{\sigma(n)}^{k_n}\right)}}\\
&=\sum_{\sigma\in S_n}(-1)^\sigma\sum_{(k_1,k_2,...,k_n)\in\mathbb{N}^n}\left(x_1^{k_1}x_2^{k_2}...x_n^{k_n}\right)\left(y_{\sigma(1)}^{k_1}y_{\sigma(2)}^{k_2}...y_{\sigma(n)}^{k_n}\right)\\
&=\sum_{(k_1,k_2,...,k_n)\in\mathbb{N}^n}x_1^{k_1}x_2^{k_2}...x_n^{k_n}\sum_{\sigma\in S_n}(-1)^\sigma y_{\sigma(1)}^{k_1}y_{\sigma(2)}^{k_2}...y_{\sigma(n)}^{k_n}. \qquad (13.70.16)
\end{aligned}$$

But every $(k_1,k_2,...,k_n)\in\mathbb{N}^n$ satisfies

$$\begin{aligned}
\det\left(\left(y_j^{k_i}\right)_{i,j=1,2,...,n}\right) &= \sum_{\sigma\in S_n}(-1)^\sigma\underbrace{\prod_{i=1}^{n}y_{\sigma(i)}^{k_i}}_{=y_{\sigma(1)}^{k_1}y_{\sigma(2)}^{k_2}...y_{\sigma(n)}^{k_n}} \qquad \left(\text{by (13.70.1), applied to } m=n \text{ and } \alpha_{i,j}=y_j^{k_i}\right)\\
&=\sum_{\sigma\in S_n}(-1)^\sigma y_{\sigma(1)}^{k_1}y_{\sigma(2)}^{k_2}...y_{\sigma(n)}^{k_n}. \qquad (13.70.17)
\end{aligned}$$

Hence, (13.70.16) becomes

$$\begin{aligned}
\det\left(\left(\frac{1}{1-x_iy_j}\right)_{i,j=1,2,\ldots,n}\right) &= \sum_{(k_1,k_2,\ldots,k_n)\in\mathbb{N}^n} x_1^{k_1}x_2^{k_2}\ldots x_n^{k_n} \underbrace{\sum_{\sigma\in S_n}(-1)^\sigma y_{\sigma(1)}^{k_1}y_{\sigma(2)}^{k_2}\ldots y_{\sigma(n)}^{k_n}}_{\substack{=\det\left(\left(y_j^{k_i}\right)_{i,j=1,2,\ldots,n}\right)\\ \text{(by (13.70.17))}}} \\
&= \sum_{(k_1,k_2,\ldots,k_n)\in\mathbb{N}^n} x_1^{k_1}x_2^{k_2}\ldots x_n^{k_n} \det\left(\left(y_j^{k_i}\right)_{i,j=1,2,\ldots,n}\right) \\
&= \sum_{\substack{(k_1,k_2,\ldots,k_n)\in\mathbb{N}^n;\\ \text{the integers } k_1,\ k_2,\ \ldots,\ k_n\\ \text{are distinct}}} x_1^{k_1}x_2^{k_2}\ldots x_n^{k_n} \det\left(\left(y_j^{k_i}\right)_{i,j=1,2,\ldots,n}\right). \qquad (13.70.18)
\end{aligned}$$

(Here, we have removed from our sum all addends in which the integers k_1, k_2, ..., k_n are not distinct. This did not change the value of the sum, because all these addends are zero[616].)

On the other hand, every $(k_1, k_2, ..., k_n) \in \mathbb{N}^n$ satisfies

$$\begin{aligned}
\det\left(\left(x_j^{k_i}\right)_{i,j=1,2,\ldots,n}\right) &= \sum_{\sigma\in S_n}(-1)^\sigma \underbrace{\prod_{i=1}^n x_{\sigma(i)}^{k_i}}_{=x_{\sigma(1)}^{k_1}x_{\sigma(2)}^{k_2}\ldots x_{\sigma(n)}^{k_n}} \qquad \left(\text{by (13.70.1), applied to } m=n \text{ and } \alpha_{i,j}=x_j^{k_i}\right) \\
&= \sum_{\sigma\in S_n}(-1)^\sigma x_{\sigma(1)}^{k_1}x_{\sigma(2)}^{k_2}\ldots x_{\sigma(n)}^{k_n}.
\end{aligned}$$

[616] In fact, if $(k_1, k_2, ..., k_n) \in \mathbb{N}^n$ is such that the integers k_1, k_2, ..., k_n are **not** distinct, then the matrix $\left(y_j^{k_i}\right)_{i,j=1,2,\ldots,n}$ has two equal rows, which causes its determinant $\det\left(\left(y_j^{k_i}\right)_{i,j=1,2,\ldots,n}\right)$ to be 0, and thus the addend corresponding to this $(k_1, k_2, ..., k_n) \in \mathbb{N}^n$ is

$$x_1^{k_1}x_2^{k_2}\ldots x_n^{k_n} \underbrace{\det\left(\left(y_j^{k_i}\right)_{i,j=1,2,\ldots,n}\right)}_{=0} = 0.$$

Thus,

$$\begin{aligned}
&\sum_{\substack{(k_1,k_2,\ldots,k_n)\in\mathbb{N}^n;\\ k_1>k_2>\ldots>k_n}} \underbrace{\det\left(\left(x_j^{k_i}\right)_{i,j=1,2,\ldots,n}\right)}_{=\sum_{\sigma\in S_n}(-1)^\sigma x_{\sigma(1)}^{k_1}x_{\sigma(2)}^{k_2}\ldots x_{\sigma(n)}^{k_n}} \det\left(\left(y_j^{k_i}\right)_{i,j=1,2,\ldots,n}\right) \\
&= \sum_{\substack{(k_1,k_2,\ldots,k_n)\in\mathbb{N}^n;\\ k_1>k_2>\ldots>k_n}} \sum_{\sigma\in S_n} (-1)^\sigma x_{\sigma(1)}^{k_1}x_{\sigma(2)}^{k_2}\ldots x_{\sigma(n)}^{k_n} \det\left(\left(y_j^{k_i}\right)_{i,j=1,2,\ldots,n}\right) \\
&= \sum_{\sigma\in S_n} (-1)^\sigma \underbrace{\sum_{\substack{(k_1,k_2,\ldots,k_n)\in\mathbb{N}^n;\\ k_1>k_2>\ldots>k_n}} x_{\sigma(1)}^{k_1}x_{\sigma(2)}^{k_2}\ldots x_{\sigma(n)}^{k_n} \det\left(\left(y_j^{k_i}\right)_{i,j=1,2,\ldots,n}\right)}_{\substack{=\sum_{\substack{(k_1,k_2,\ldots,k_n)\in\mathbb{N}^n;\\ k_{\sigma(1)}>k_{\sigma(2)}>\ldots>k_{\sigma(n)}}} x_{\sigma(1)}^{k_{\sigma(1)}}x_{\sigma(2)}^{k_{\sigma(2)}}\ldots x_{\sigma(n)}^{k_{\sigma(n)}} \det\left(\left(y_j^{k_{\sigma(i)}}\right)_{i,j=1,2,\ldots,n}\right)\\ \text{(here, we substituted } \left(k_{\sigma(1)},k_{\sigma(2)},\ldots,k_{\sigma(n)}\right) \text{ for } (k_1,k_2,\ldots,k_n) \text{ in the sum}\\ \text{(this is allowed, since } \sigma \text{ is a permutation))}}} \\
&= \sum_{\sigma\in S_n} (-1)^\sigma \sum_{\substack{(k_1,k_2,\ldots,k_n)\in\mathbb{N}^n;\\ k_{\sigma(1)}>k_{\sigma(2)}>\ldots>k_{\sigma(n)}}} \underbrace{x_{\sigma(1)}^{k_{\sigma(1)}}x_{\sigma(2)}^{k_{\sigma(2)}}\ldots x_{\sigma(n)}^{k_{\sigma(n)}}}_{\substack{=x_1^{k_1}x_2^{k_2}\ldots x_n^{k_n}\\ \text{(since } \sigma \text{ is a permutation)}}} \underbrace{\det\left(\left(y_j^{k_{\sigma(i)}}\right)_{i,j=1,2,\ldots,n}\right)}_{\substack{=(-1)^\sigma \det\left(\left(y_j^{k_i}\right)_{i,j=1,2,\ldots,n}\right)\\ \text{(since permuting the rows of a matrix}\\ \text{multiplies its determinant by the sign}\\ \text{of the permutation)}}} \\
&= \sum_{\sigma\in S_n} (-1)^\sigma \sum_{\substack{(k_1,k_2,\ldots,k_n)\in\mathbb{N}^n;\\ k_{\sigma(1)}>k_{\sigma(2)}>\ldots>k_{\sigma(n)}}} x_1^{k_1}x_2^{k_2}\ldots x_n^{k_n} (-1)^\sigma \det\left(\left(y_j^{k_i}\right)_{i,j=1,2,\ldots,n}\right) \\
&= \sum_{\sigma\in S_n} \underbrace{(-1)^\sigma (-1)^\sigma}_{=((-1)^\sigma)^2=1} \sum_{\substack{(k_1,k_2,\ldots,k_n)\in\mathbb{N}^n;\\ k_{\sigma(1)}>k_{\sigma(2)}>\ldots>k_{\sigma(n)}}} x_1^{k_1}x_2^{k_2}\ldots x_n^{k_n} \det\left(\left(y_j^{k_i}\right)_{i,j=1,2,\ldots,n}\right) \\
&= \underbrace{\sum_{\sigma\in S_n} \sum_{\substack{(k_1,k_2,\ldots,k_n)\in\mathbb{N}^n;\\ k_{\sigma(1)}>k_{\sigma(2)}>\ldots>k_{\sigma(n)}}}}_{=\sum_{(k_1,k_2,\ldots,k_n)\in\mathbb{N}^n} \sum_{\substack{\sigma\in S_n;\\ k_{\sigma(1)}>k_{\sigma(2)}>\ldots>k_{\sigma(n)}}}} x_1^{k_1}x_2^{k_2}\ldots x_n^{k_n} \det\left(\left(y_j^{k_i}\right)_{i,j=1,2,\ldots,n}\right) \\
&= \sum_{(k_1,k_2,\ldots,k_n)\in\mathbb{N}^n} \sum_{\substack{\sigma\in S_n;\\ k_{\sigma(1)}>k_{\sigma(2)}>\ldots>k_{\sigma(n)}}} x_1^{k_1}x_2^{k_2}\ldots x_n^{k_n} \det\left(\left(y_j^{k_i}\right)_{i,j=1,2,\ldots,n}\right) \\
&= \sum_{\substack{(k_1,k_2,\ldots,k_n)\in\mathbb{N}^n;\\ \text{the integers } k_1,\ k_2,\ \ldots,\ k_n\\ \text{are distinct}}} \underbrace{\sum_{\substack{\sigma\in S_n;\\ k_{\sigma(1)}>k_{\sigma(2)}>\ldots>k_{\sigma(n)}}} x_1^{k_1}x_2^{k_2}\ldots x_n^{k_n} \det\left(\left(y_j^{k_i}\right)_{i,j=1,2,\ldots,n}\right)}_{\substack{=x_1^{k_1}x_2^{k_2}\ldots x_n^{k_n} \det\left(\left(y_j^{k_i}\right)_{i,j=1,2,\ldots,n}\right)\\ \text{(because there exists exactly one } \sigma\in S_n \text{ such that } k_{\sigma(1)}>k_{\sigma(2)}>\ldots>k_{\sigma(n)}\\ \text{(since the integers } k_1,\ k_2,\ \ldots,\ k_n \text{ are distinct, and thus}\\ \text{there is exactly one permutation which arranges these integers}\\ \text{in decreasing order))}}} \\
&\qquad + \sum_{\substack{(k_1,k_2,\ldots,k_n)\in\mathbb{N}^n;\\ \text{the integers } k_1,\ k_2,\ \ldots,\ k_n\\ \text{are not distinct}}} \sum_{\substack{\sigma\in S_n;\\ k_{\sigma(1)}>k_{\sigma(2)}>\ldots>k_{\sigma(n)}}} x_1^{k_1}x_2^{k_2}\ldots x_n^{k_n} \underbrace{\det\left(\left(y_j^{k_i}\right)_{i,j=1,2,\ldots,n}\right)}_{\substack{=0\\ \text{(because the matrix } \left(y_j^{k_i}\right)_{i,j=1,2,\ldots,n}\\ \text{has two equal rows (since the integers}\\ k_1,\ k_2,\ \ldots,\ k_n \text{ are not distinct,}\\ \text{i. e., there are two equal among them))}}}
\end{aligned}$$

$$= \sum_{\substack{(k_1,k_2,\ldots,k_n)\in\mathbb{N}^n;\\ \text{the integers } k_1,\ k_2,\ \ldots,\ k_n\\ \text{are distinct}}} x_1^{k_1}x_2^{k_2}\ldots x_n^{k_n}\det\left(\left(y_j^{k_i}\right)_{i,j=1,2,\ldots,n}\right)$$
$$+ \underbrace{\sum_{\substack{(k_1,k_2,\ldots,k_n)\in\mathbb{N}^n;\\ \text{the integers } k_1,\ k_2,\ \ldots,\ k_n\\ \text{are not distinct}}} \sum_{\substack{\sigma\in S_n;\\ k_{\sigma(1)}>k_{\sigma(2)}>\ldots>k_{\sigma(n)}}} x_1^{k_1}x_2^{k_2}\ldots x_n^{k_n}0}_{=0}$$
$$= \sum_{\substack{(k_1,k_2,\ldots,k_n)\in\mathbb{N}^n;\\ \text{the integers } k_1,\ k_2,\ \ldots,\ k_n\\ \text{are distinct}}} x_1^{k_1}x_2^{k_2}\ldots x_n^{k_n}\det\left(\left(y_j^{k_i}\right)_{i,j=1,2,\ldots,n}\right).$$

Compared with (13.70.18), this yields

$$\det\left(\left(\frac{1}{1-x_iy_j}\right)_{i,j=1,2,\ldots,n}\right)$$
$$= \sum_{\substack{(k_1,k_2,\ldots,k_n)\in\mathbb{N}^n;\\ k_1>k_2>\ldots>k_n}} \det\left(\left(x_j^{k_i}\right)_{i,j=1,2,\ldots,n}\right)\det\left(\left(y_j^{k_i}\right)_{i,j=1,2,\ldots,n}\right)$$
$$= \underbrace{\sum_{\substack{(\lambda_1,\lambda_2,\ldots,\lambda_n)\in\mathbb{N}^n;\\ \lambda_1\geq\lambda_2\geq\ldots\geq\lambda_n}}}_{=\sum_{\substack{\lambda=(\lambda_1,\lambda_2,\ldots,\lambda_n)\in\mathbb{N}^n\\ \text{is a partition with at most } n \text{ parts}}}} \det\left(\left(x_j^{\lambda_i+n-i}\right)_{i,j=1,2,\ldots,n}\right)\det\left(\left(y_j^{\lambda_i+n-i}\right)_{i,j=1,2,\ldots,n}\right)$$
$$\left(\begin{array}{c}\text{here, we substituted } (\lambda_1+n-1,\lambda_2+n-2,\ldots,\lambda_n+n-n) \text{ for } (k_1,k_2,\ldots,k_n)\\ \text{in the sum (since the map}\\ \{(\lambda_1,\lambda_2,\ldots,\lambda_n)\in\mathbb{N}^n \mid \lambda_1\geq\lambda_2\geq\ldots\geq\lambda_n\}\to\{(k_1,k_2,\ldots,k_n)\in\mathbb{N}^n \mid k_1>k_2>\ldots>k_n\},\\ (\lambda_1,\lambda_2,\ldots,\lambda_n)\mapsto(\lambda_1+n-1,\lambda_2+n-2,\ldots,\lambda_n+n-n)\\ \text{is a bijection)}\end{array}\right)$$

(13.70.19)
$$= \sum_{\substack{\lambda=(\lambda_1,\lambda_2,\ldots,\lambda_n)\in\mathbb{N}^n\\ \text{is a partition with at most } n \text{ parts}}} \det\left(\left(x_j^{\lambda_i+n-i}\right)_{i,j=1,2,\ldots,n}\right)\det\left(\left(y_j^{\lambda_i+n-i}\right)_{i,j=1,2,\ldots,n}\right).$$

But whenever $\lambda=(\lambda_1,\lambda_2,\ldots,\lambda_n)$ is a partition with at most n parts, we have

$$\underbrace{\lambda}_{=(\lambda_1,\lambda_2,\ldots,\lambda_n)} + \underbrace{\rho}_{=(n-1,n-2,\ldots,0)} = (\lambda_1,\lambda_2,\ldots,\lambda_n)+(n-1,n-2,\ldots,0)$$
$$= (\lambda_1+n-1,\lambda_2+n-2,\ldots,\lambda_n+n-n)$$

and thus

$$\begin{aligned}
a_{\lambda+\rho} &= a_{(\lambda_1+n-1,\lambda_2+n-2,\ldots,\lambda_n+n-n)} = \det\left(\underbrace{\left(x_i^{\lambda_j+n-j}\right)_{i,j=1,2,\ldots,n}}_{=\left(\left(x_j^{\lambda_i+n-i}\right)_{i,j=1,2,\ldots,n}\right)^T}\right) \\
&\qquad \left(\text{by the definition of } a_{(\alpha_1,\alpha_2,\ldots,\alpha_n)}\right) \\
&= \det\left(\left(\left(x_j^{\lambda_i+n-i}\right)_{i,j=1,2,\ldots,n}\right)^T\right) = \det\left(\left(x_j^{\lambda_i+n-i}\right)_{i,j=1,2,\ldots,n}\right) \\
&\qquad (\text{since the determinant of a matrix is preserved under transposition}) .
\end{aligned}$$

Hence, we get the two equalities $a_{\lambda+\rho}(x_1,x_2,\ldots,x_n) = \det\left(\left(x_j^{\lambda_i+n-i}\right)_{i,j=1,2,\ldots,n}\right)$ and $a_{\lambda+\rho}(y_1,y_2,\ldots,y_n) = \det\left(\left(y_j^{\lambda_i+n-i}\right)_{i,j=1,2,\ldots,n}\right)$ for any partition $\lambda = (\lambda_1,\lambda_2,\ldots,\lambda_n)$ with at most n parts. Thus, (13.70.19) becomes

$$\begin{aligned}
&\det\left(\left(\frac{1}{1-x_iy_j}\right)_{i,j=1,2,\ldots,n}\right) \\
&= \sum_{\substack{\lambda=(\lambda_1,\lambda_2,\ldots,\lambda_n)\in\mathbb{N}^n \\ \text{is a partition with at most } n \text{ parts}}} \underbrace{\det\left(\left(x_j^{\lambda_i+n-i}\right)_{i,j=1,2,\ldots,n}\right)}_{=a_{\lambda+\rho}(x_1,x_2,\ldots,x_n)} \underbrace{\det\left(\left(y_j^{\lambda_i+n-i}\right)_{i,j=1,2,\ldots,n}\right)}_{=a_{\lambda+\rho}(y_1,y_2,\ldots,y_n)} \\
&= \sum_{\substack{\lambda=(\lambda_1,\lambda_2,\ldots,\lambda_n)\in\mathbb{N}^n \\ \text{is a partition with at most } n \text{ parts}}} a_{\lambda+\rho}(x_1,x_2,\ldots,x_n)\, a_{\lambda+\rho}(y_1,y_2,\ldots,y_n) \\
&= \sum_{\lambda \text{ is a partition with at most } n \text{ parts}} a_{\lambda+\rho}(x_1,x_2,\ldots,x_n)\, a_{\lambda+\rho}(y_1,y_2,\ldots,y_n) .
\end{aligned}$$

Compared to (13.70.14), this yields

$$\begin{aligned}
&\sum_{\lambda \text{ is a partition with at most } n \text{ parts}} a_{\lambda+\rho}(x_1,x_2,\ldots,x_n)\, a_{\lambda+\rho}(y_1,y_2,\ldots,y_n) \\
&= \underbrace{\left(\prod_{1\leq i<j\leq n} (x_i - x_j)\right)}_{=a_\rho(x_1,x_2,\ldots,x_n)} \cdot \underbrace{\left(\prod_{1\leq i<j\leq n} (y_i - y_j)\right)}_{=a_\rho(y_1,y_2,\ldots,y_n)} \cdot \prod_{i,j=1}^{n} \underbrace{\frac{1}{1-x_iy_j}}_{=(1-x_iy_j)^{-1}} \\
&= a_\rho(x_1,x_2,\ldots,x_n) \cdot a_\rho(y_1,y_2,\ldots,y_n) \cdot \prod_{i,j=1}^{n} (1-x_iy_j)^{-1} . \qquad (13.70.20)
\end{aligned}$$

Now, Corollary 2.6.7 says that $s_\lambda(x_1,x_2,\ldots,x_n) = \frac{a_{\lambda+\rho}}{a_\rho}$ for every partition λ with at most n parts. Thus, $s_\lambda(x_1,x_2,\ldots,x_n) = \frac{a_{\lambda+\rho}(x_1,x_2,\ldots,x_n)}{a_\rho(x_1,x_2,\ldots,x_n)}$ and $s_\lambda(y_1,y_2,\ldots,y_n) = \frac{a_{\lambda+\rho}(y_1,y_2,\ldots,y_n)}{a_\rho(y_1,y_2,\ldots,y_n)}$ for every partition λ

with at most n parts. Hence,

$$\begin{aligned}
&\sum_{\lambda \text{ is a partition with at most } n \text{ parts}} \underbrace{s_\lambda\left(x_1,x_2,...,x_n\right)}_{=\dfrac{a_{\lambda+\rho}\left(x_1,x_2,...,x_n\right)}{a_\rho\left(x_1,x_2,...,x_n\right)}} \underbrace{s_\lambda\left(y_1,y_2,...,y_n\right)}_{=\dfrac{a_{\lambda+\rho}\left(y_1,y_2,...,y_n\right)}{a_\rho\left(y_1,y_2,...,y_n\right)}} \\
&= \sum_{\lambda \text{ is a partition with at most } n \text{ parts}} \frac{a_{\lambda+\rho}\left(x_1,x_2,...,x_n\right)}{a_\rho\left(x_1,x_2,...,x_n\right)} \cdot \frac{a_{\lambda+\rho}\left(y_1,y_2,...,y_n\right)}{a_\rho\left(y_1,y_2,...,y_n\right)} \\
&= \frac{1}{a_\rho\left(x_1,x_2,...,x_n\right)\cdot a_\rho\left(y_1,y_2,...,y_n\right)} \underbrace{\sum_{\lambda \text{ is a partition with at most } n \text{ parts}} a_{\lambda+\rho}\left(x_1,x_2,...,x_n\right) a_{\lambda+\rho}\left(y_1,y_2,...,y_n\right)}_{\substack{=a_\rho(x_1,x_2,...,x_n)\cdot a_\rho(y_1,y_2,...,y_n)\cdot \prod_{i,j=1}^{n}(1-x_iy_j)^{-1} \\ \text{(by (13.70.20))}}} \\
&= \frac{1}{a_\rho\left(x_1,x_2,...,x_n\right)\cdot a_\rho\left(y_1,y_2,...,y_n\right)} \cdot a_\rho\left(x_1,x_2,...,x_n\right)\cdot a_\rho\left(y_1,y_2,...,y_n\right)\cdot \prod_{i,j=1}^{n}\left(1-x_iy_j\right)^{-1}
\end{aligned}$$

(13.70.21)
$$= \prod_{i,j=1}^{n}\left(1-x_iy_j\right)^{-1}.$$

We are almost done. Let us now notice that every partition λ with more than n parts satisfies

$$s_\lambda\left(x_1,x_2,...,x_n\right) = 0 \qquad (13.70.22)$$

(according to Exercise 2.3.8(b)). Thus, in the sum $\sum_{\lambda\in\operatorname{Par}} s_\lambda\left(x_1,x_2,...,x_n\right) s_\lambda\left(y_1,y_2,...,y_n\right)$, all addends corresponding to partitions λ having more than n parts are 0. We can therefore remove these addends from the sum. Hence,

$$\begin{aligned}
&\sum_{\lambda\in\operatorname{Par}} s_\lambda\left(x_1,x_2,...,x_n\right) s_\lambda\left(y_1,y_2,...,y_n\right) \\
&= \sum_{\lambda \text{ is a partition with at most } n \text{ parts}} s_\lambda\left(x_1,x_2,...,x_n\right) s_\lambda\left(y_1,y_2,...,y_n\right) \\
&= \prod_{i,j=1}^{n}\left(1-x_iy_j\right)^{-1} \qquad \text{(by (13.70.21))}.
\end{aligned}$$

Thus, we have proven that (13.70.13) holds for every $n\in\mathbb{N}$. As we explained above, this yields that Theorem 2.5.1 is true.

Remark: In our above solution, we solved Exercise 2.7.8(b) first, and then used it to prove Theorem 2.5.1. It is also possible (more or less by treading the above proof backwards) to conversely derive the statement of Exercise 2.7.8(b) from Theorem 2.5.1 instead (though Exercise 2.7.8(b) is usually considered a more elementary fact than Theorem 2.5.1).

13.71. **Solution to Exercise 2.7.9.** *Solution to Exercise 2.7.9.* Let us first check that we have

$$h_u h_v = \sum_{i=0}^{v} s_{(u+i,v-i)} \qquad (13.71.1)$$

for any two nonnegative integers u and v satisfying $u\geq v$.

Proof of (13.71.1)*:* Let u and v be nonnegative integers satisfying $u\geq v$. We WLOG assume that $u>0$ (otherwise, $u=0$, and $u\geq v$ forces $v=0$, so that (13.71.1) can be checked immediately). We have $h_u = s_{(u)}$

and thus

$$h_u h_v = s_{(u)} h_v = \sum_{\substack{\lambda^+ \ : \ \lambda^+/(u) \text{ is a} \\ \text{horizontal } v\text{-strip}}} s_{\lambda^+} \tag{13.71.2}$$

(by (2.7.1), applied to $n = v$ and $\lambda = (u)$).

Now, fix a partition λ^+ such that $\lambda^+ / (u)$ is a horizontal v-strip. Then, the skew diagram $\lambda^+ / (u)$ has at most one cell in column 1 (since $\lambda^+ / (u)$ is a horizontal v-strip and thus contains no two cells in the same column). Hence, the partition λ^+ has at most 2 rows (because otherwise, the skew diagram $\lambda^+ / (u)$ would contain at least two cells in column 1, which would contradict the fact that the skew diagram $\lambda^+ / (u)$ has at most one cell in column 1). Thus, the partition λ^+ has the form (p, q) for two nonnegative integers p and q satisfying $p \geq q$. Consider these two integers p and q. Since $\lambda^+ / (u)$ is a horizontal v-strip, we have $|\lambda^+ / (u)| = v$ and thus $v = |\lambda^+ / (u)| = \left| \underbrace{\lambda^+}_{=(p,q)} \right| - \underbrace{|(u)|}_{=u} = \underbrace{|(p,q)|}_{=p+q} - u = p + q - u$. Moreover, $\lambda^+ / (u)$ is a horizontal v-strip, so that $(u) \subseteq \lambda^+ = (p, q)$. Thus, $u \leq p$. Hence, there exists an $i \in \mathbb{N}$ such that $p = u + i$. Consider this i. Now, $v = \underbrace{p}_{=u+i} + q - u = u + i + q - u = q + i$, so that $q = v - i$ and thus $v - i = q \geq 0$, so that $i \leq v$ and thus $i \in \{0, 1, ..., v\}$ (since $i \in \mathbb{N}$). Hence, $\lambda^+ = \left(\underbrace{p}_{=u+i}, \underbrace{q}_{=v-i} \right) = (u + i, v - i)$. We have thus shown that the partition λ^+ has the form $\lambda^+ = (u + i, v - i)$ for some $i \in \{0, 1, ..., v\}$.

Now forget that we fixed λ^+. We thus have proven that every partition λ^+ such that $\lambda^+ / (u)$ is a horizontal v-strip has the form $\lambda^+ = (u + i, v - i)$ for some $i \in \{0, 1, ..., v\}$. Conversely, it is clear that for every $i \in \{0, 1, ..., v\}$, the weak composition $(u + i, v - i)$ is a partition λ^+ such that $\lambda^+ / (u)$ is a horizontal v-strip. Combining the previous two sentences, we conclude that the partitions λ^+ such that $\lambda^+ / (u)$ is a horizontal v-strip are precisely the weak compositions of the form $(u + i, v - i)$ for $i \in \{0, 1, ..., v\}$. Therefore, $\sum_{\substack{\lambda^+ \ : \ \lambda^+/(u) \text{ is a} \\ \text{horizontal } v\text{-strip}}} s_{\lambda^+} = \sum_{i \in \{0,1,...,v\}} s_{(u+i,v-i)} = \sum_{i=0}^{v} s_{(u+i,v-i)}$. Now, (13.71.2) becomes

$$h_u h_v = \sum_{\substack{\lambda^+ \ : \ \lambda^+/(u) \text{ is a} \\ \text{horizontal } v\text{-strip}}} s_{\lambda^+} = \sum_{i=0}^{v} s_{(u+i,v-i)}.$$

This proves (13.71.1).

Now, fix two integers a and b satisfying $a \geq b \geq 0$. We need to prove that $s_{(a,b)} = h_a h_b - h_{a+1} h_{b-1}$.

If $b = 0$, then proving $s_{(a,b)} = h_a h_b - h_{a+1} h_{b-1}$ is very easy[617]. Hence, for the rest of the proof of $s_{(a,b)} = h_a h_b - h_{a+1} h_{b-1}$, we assume WLOG that we don't have $b = 0$. Thus, $b \geq 1$, so that $b - 1 \geq 0$. Clearly, $a \geq b$ shows that $a + 1 \geq b + 1 \geq b - 1$, so that we can apply (13.71.1) to $u = a + 1$ and $v = b - 1$.

[617] In fact, assume that $b = 0$. Then,

$$s_{(a,b)} = s_{(a,0)} = s_{(a)} = h_a$$

and

$$h_a \underbrace{h_b}_{=h_0=1} - h_{a+1} \underbrace{h_{b-1}}_{=h_{-1}=0} = h_a 1 - h_{a+1} 0 = h_a.$$

Hence, $s_{(a,b)} = h_a = h_a h_b - h_{a+1} h_{b-1}$, qed.

Now,

$$\begin{aligned}
&\underbrace{h_a h_b}_{\substack{=\sum_{i=0}^{b} s_{(a+i,b-i)} \\ \text{(by (13.71.1), applied} \\ \text{to } u=a \text{ and } v=b)}} - \underbrace{h_{a+1} h_{b-1}}_{\substack{=\sum_{i=0}^{b-1} s_{(a+1+i,b-1-i)} \\ \text{(by (13.71.1), applied} \\ \text{to } u=a+1 \text{ and } v=b-1)}} \\
&= \underbrace{\sum_{i=0}^{b} s_{(a+i,b-i)}}_{=s_{(a+0,b-0)}+\sum_{i=1}^{b} s_{(a+i,b-i)}} - \sum_{i=0}^{b-1} \underbrace{s_{(a+1+i,b-1-i)}}_{=s_{(a+(i+1),b-(i+1))}} = s_{(a+0,b-0)} + \sum_{i=1}^{b} s_{(a+i,b-i)} - \sum_{i=0}^{b-1} s_{(a+(i+1),b-(i+1))} \\
&= s_{(a+0,b-0)} + \sum_{i=1}^{b} s_{(a+i,b-i)} - \sum_{i=1}^{b} s_{(a+i,b-i)} \\
&\qquad \text{(here, we have substituted } i \text{ for } i+1 \text{ in the second sum)} \\
&= s_{(a+0,b-0)} = s_{(a,b)}.
\end{aligned}$$

We thus have proven $s_{(a,b)} = h_a h_b - h_{a+1} h_{b-1}$. The exercise is solved.

13.72. **Solution to Exercise 2.7.10.** *Solution to Exercise 2.7.10.* (a) We shall prove the statement of Exercise 2.7.10(a) by induction over the length $\ell(\mu)$ of μ.

The *induction base* (i.e., the case $\ell(\mu) = 0$) is trivial. For the *induction step*, we fix a positive integer L and assume (as the induction hypothesis) that Exercise 2.7.10(a) has been solved for all μ satisfying $\ell(\mu) = L - 1$. We now have to solve Exercise 2.7.10(a) for every partition μ satisfying $\ell(\mu) = L$.

So let μ be a partition satisfying $\ell(\mu) = L$. Write μ in the form $(\mu_1, \mu_2, \ldots, \mu_L)$. Let $\overline{\mu}$ be the partition $(\mu_1, \mu_2, \ldots, \mu_{L-1})$ (this is well-defined since L is positive). Then, $\ell(\overline{\mu}) = L - 1$, and hence (by the induction hypothesis) we can apply Exercise 2.7.10(a) to $\overline{\mu}$ instead of μ. As a result, we obtain $h_{\overline{\mu}} = \sum_\lambda K_{\lambda,\overline{\mu}} s_\lambda$, where the sum ranges over all partitions λ.

But the definition of h_μ yields $h_\mu = h_{\mu_1} h_{\mu_2} \ldots h_{\mu_L}$; similarly, $h_{\overline{\mu}} = h_{\mu_1} h_{\mu_2} \ldots h_{\mu_{L-1}}$. Hence,

$$\begin{aligned}
h_\mu &= h_{\mu_1} h_{\mu_2} \ldots h_{\mu_L} = \underbrace{h_{\mu_1} h_{\mu_2} \ldots h_{\mu_{L-1}}}_{=h_{\overline{\mu}}=\sum_\lambda K_{\lambda,\overline{\mu}} s_\lambda} h_{\mu_L} = \left(\sum_\lambda K_{\lambda,\overline{\mu}} s_\lambda \right) h_{\mu_L} \\
&= \sum_\lambda K_{\lambda,\overline{\mu}} \underbrace{s_\lambda h_{\mu_L}}_{\substack{= \sum_{\lambda^+ \,:\, \lambda^+/\lambda \text{ is a horizontal } \mu_L\text{-strip}} s_{\lambda^+} \\ \text{(by (2.7.1), applied to } n=\mu_L)}} = \sum_\lambda K_{\lambda,\overline{\mu}} \sum_{\substack{\lambda^+ \,:\, \lambda^+/\lambda \text{ is a horizontal} \\ \mu_L\text{-strip}}} s_{\lambda^+} \\
&= \underbrace{\sum_\lambda \sum_{\substack{\lambda^+ \,:\, \lambda^+/\lambda \text{ is a horizontal} \\ \mu_L\text{-strip}}}}_{=\sum_{\lambda^+} \sum_{\lambda \,:\, \lambda^+/\lambda \text{ is a horizontal } \mu_L\text{-strip}}} K_{\lambda,\overline{\mu}} s_{\lambda^+} = \sum_{\lambda^+} \sum_{\substack{\lambda \,:\, \lambda^+/\lambda \text{ is a horizontal} \\ \mu_L\text{-strip}}} K_{\lambda,\overline{\mu}} s_{\lambda^+}
\end{aligned}$$

$$= \sum_\lambda \sum_{\substack{\lambda^- \,:\, \lambda/\lambda^- \text{ is a horizontal} \\ \mu_L\text{-strip}}} K_{\lambda^-,\overline{\mu}} s_\lambda \tag{13.72.1}$$

$$\left(\text{here, we renamed the summation indices } \lambda^+ \text{ and } \lambda \text{ as } \lambda \text{ and } \lambda^-\right),$$

where all summation indices are supposed to be partitions.

Now, let us fix a partition λ. We shall show that

$$\sum_{\substack{\lambda^- \,:\, \lambda/\lambda^- \text{ is a horizontal} \\ \mu_L\text{-strip}}} K_{\lambda^-,\overline{\mu}} = K_{\lambda,\mu}. \tag{13.72.2}$$

Proof of (13.72.2)*:* For every partition λ^-, the number $K_{\lambda^-,\overline{\mu}}$ is the number of column-strict tableaux S of shape λ^- having $\operatorname{cont}(S) = \overline{\mu}$. Hence, the sum $\sum\limits_{\substack{\lambda^- \ : \ \lambda/\lambda^- \text{ is a horizontal} \\ \mu_L\text{-strip}}} K_{\lambda^-,\overline{\mu}}$ is the number of all pairs (λ^-, S), where:

- λ^- is a partition such that λ/λ^- is a horizontal μ_L-strip;
- S is a column-strict tableau of shape λ^- having $\operatorname{cont}(S) = \overline{\mu}$.

We will refer to such pairs (λ^-, S) as (λ, μ)-*last-step pairs*. So the sum $\sum\limits_{\substack{\lambda^- \ : \ \lambda/\lambda^- \text{ is a horizontal} \\ \mu_L\text{-strip}}} K_{\lambda^-,\overline{\mu}}$ is the number of (λ, μ)-last-step pairs. On the other hand, $K_{\lambda,\mu}$ is the number of all column-strict tableaux T of shape λ having $\operatorname{cont}(T) = \mu$. We now will construct a bijection between the (λ, μ)-last-step pairs and the column-strict tableaux T of shape λ having $\operatorname{cont}(T) = \mu$.

Indeed, let us first define a map Φ from the set of all (λ, μ)-last-step pairs to the set of all column-strict tableaux T of shape λ having $\operatorname{cont}(T) = \mu$. Namely, let (λ^-, S) be a (λ, μ)-last-step pair. Then, λ^- is a partition such that λ/λ^- is a horizontal μ_L-strip, whereas S is a column-strict tableau of shape λ^- having $\operatorname{cont}(S) = \overline{\mu}$. In particular, all entries of S are $< L$ (since $\operatorname{cont}(S) = \overline{\mu} = (\mu_1, \mu_2, \ldots, \mu_{L-1})$). Now, we can extend the column-strict tableau S of shape λ^- to a tableau of shape λ by filling the number L into all cells of the skew shape λ/λ^-. The resulting tableau is a column-strict tableau T of shape λ having $\operatorname{cont}(T) = \mu$ (indeed, its column-strictness follows from the fact that λ/λ^- is a horizontal μ_L-strip whereas all entries of S are $< L$; and the property $\operatorname{cont}(T) = \mu$ follows from the facts that $\operatorname{cont}(S) = \overline{\mu}$ and $|\lambda/\lambda^-| = \mu_L$). We define $\Phi(\lambda^-, S)$ to be this tableau. Thus we have defined a map Φ.

Conversely, let us define a map Ψ from the set of all column-strict tableaux T of shape λ having $\operatorname{cont}(T) = \mu$ to the set of all (λ, μ)-last-step pairs. Namely, let T be a column-strict tableau of shape λ having $\operatorname{cont}(T) = \mu$. Then, all entries of T are $\leq L$ (since $\operatorname{cont}(T) = \mu = (\mu_1, \mu_2, \ldots, \mu_L)$), and the cells containing the entries L form a horizontal strip (since T is column-strict). Hence, if we remove all entries L from T (along with their cells), the result will be a column-strict tableau S of some shape λ^- such that λ^- is a partition (because all entries of T were $\leq L$, so the entries we removed were maximal), λ/λ^- is a horizontal μ_L-strip (since we removed a total of μ_L entries and they formed a horizontal strip), and $\operatorname{cont}(S) = \overline{\mu}$. Consider these S and λ^-, and define $\Psi(T)$ to be the pair (λ^-, S); it is clear that this $\Psi(T)$ is a (λ, μ)-last-step pair. Hence, we have defined a map Ψ.

It is fairly obvious that the maps Φ and Ψ are mutually inverse. Hence, they are bijections; in particular, Φ is a bijection. Thus, we have a bijection between the set of all (λ, μ)-last-step pairs and the set of all column-strict tableaux T of shape λ having $\operatorname{cont}(T) = \mu$. As a consequence, the number of all (λ, μ)-last-step pairs equals the number of all column-strict tableaux T of shape λ having $\operatorname{cont}(T) = \mu$. In other words, $\sum\limits_{\substack{\lambda^- \ : \ \lambda/\lambda^- \text{ is a horizontal} \\ \mu_L\text{-strip}}} K_{\lambda^-,\overline{\mu}}$ equals $K_{\lambda,\mu}$ (because $\sum\limits_{\substack{\lambda^- \ : \ \lambda/\lambda^- \text{ is a horizontal} \\ \mu_L\text{-strip}}} K_{\lambda^-,\overline{\mu}}$ is the number of (λ, μ)-last-step pairs, whereas $K_{\lambda,\mu}$ is the number of all column-strict tableaux T of shape λ having $\operatorname{cont}(T) = \mu$). This proves (13.72.2).

Now, (13.72.1) becomes

$$h_\mu = \sum_\lambda \underbrace{\sum_{\substack{\lambda^- \ : \ \lambda/\lambda^- \text{ is a horizontal} \\ \mu_L\text{-strip}}} K_{\lambda^-,\overline{\mu}}}_{\substack{=K_{\lambda,\mu} \\ \text{(by (13.72.2))}}} s_\lambda = \sum_\lambda K_{\lambda,\mu} s_\lambda.$$

Hence, Exercise 2.7.10(a) is solved for our partition μ. Thus, Exercise 2.7.10(a) is solved for every partition μ satisfying $\ell(\mu) = L$. This completes the induction step, and thus Exercise 2.7.10(a) is solved by induction.

(b) Let us work in the power series ring $\mathbf{k}[[\mathbf{x}, \mathbf{y}]] := \mathbf{k}[[x_1, x_2, \ldots, y_1, y_2, \ldots]]$. Every partition μ satisfies $h_\mu(\mathbf{x}) = \sum_\lambda K_{\lambda,\mu} s_\lambda(\mathbf{x})$, where the sum ranges over all partitions λ (because Exercise 2.7.10(a) yields $h_\mu = \sum_\lambda K_{\lambda,\mu} s_\lambda$ in Λ). In other words, every partition μ satisfies $h_\mu(\mathbf{x}) = \sum_{\lambda \in \mathrm{Par}} K_{\lambda,\mu} s_\lambda(\mathbf{x})$. On the other hand, every partition λ satisfies $s_\lambda = \sum_\mu K_{\lambda,\mu} m_\mu$, where the sum ranges over all partitions μ (this was shown in the proof of Proposition 2.2.10). In other words, every partition λ satisfies $s_\lambda = \sum_{\mu \in \mathrm{Par}} K_{\lambda,\mu} m_\mu$,

so that

$$s_\lambda(\mathbf{y}) = \sum_{\mu\in\mathrm{Par}} K_{\lambda,\mu} m_\mu(\mathbf{y}) . \tag{13.72.3}$$

Now,

$$\begin{aligned}
\prod_{i,j=1}^{\infty} (1 - x_i y_j)^{-1} &= \sum_{\lambda\in\mathrm{Par}} h_\lambda(\mathbf{x}) m_\lambda(\mathbf{y}) \qquad (\text{by } (2.5.11)) \\
&= \sum_{\mu\in\mathrm{Par}} \underbrace{h_\mu(\mathbf{x})}_{=\sum_{\lambda\in\mathrm{Par}} K_{\lambda,\mu} s_\lambda(\mathbf{x})} m_\mu(\mathbf{y}) \qquad (\text{here, we renamed the summation index } \lambda \text{ as } \mu) \\
&= \underbrace{\sum_{\mu\in\mathrm{Par}} \sum_{\lambda\in\mathrm{Par}}}_{=\sum_{\lambda\in\mathrm{Par}} \sum_{\mu\in\mathrm{Par}}} K_{\lambda,\mu} s_\lambda(\mathbf{x}) m_\mu(\mathbf{y}) \\
&= \sum_{\lambda\in\mathrm{Par}} \sum_{\mu\in\mathrm{Par}} K_{\lambda,\mu} s_\lambda(\mathbf{x}) m_\mu(\mathbf{y}) = \sum_{\lambda\in\mathrm{Par}} s_\lambda(\mathbf{x}) \underbrace{\sum_{\mu\in\mathrm{Par}} K_{\lambda,\mu} m_\mu(\mathbf{y})}_{\substack{=s_\lambda(\mathbf{y}) \\ (\text{by } (13.72.3))}} \\
&= \sum_{\lambda\in\mathrm{Par}} s_\lambda(\mathbf{x}) s_\lambda(\mathbf{y}) .
\end{aligned}$$

This proves Theorem 2.5.1. Thus, Exercise 2.7.10(b) is solved.

(c) We start out just as in the proof of Proposition 2.2.10: We fix an $n \in \mathbb{N}$, and we restrict our attention to a given homogeneous component Λ_n and partitions of n. Regard the set Par_n as a poset with smaller-or-equal relation $\triangleright$. We will check that the family $(h_\lambda)_{\lambda\in\mathrm{Par}_n}$ expands unitriangularly[618] in the basis $(s_\lambda)_{\lambda\in\mathrm{Par}_n}$.

If λ and μ are two partitions satisfying $|\lambda| \neq |\mu|$, then

$$K_{\lambda,\mu} = 0 \tag{13.72.4}$$

(because $K_{\lambda,\mu}$ counts the column-strict tableaux T of shape λ having $\mathrm{cont}(T) = \mu$; but no such tableaux exist when $|\lambda| \neq |\mu|$).

In Exercise 2.7.10(a), we have shown that $h_\mu = \sum_{\lambda\in\mathrm{Par}} K_{\lambda,\mu} s_\lambda$ for every $\mu \in \mathrm{Par}$. Thus, for every $\mu \in \mathrm{Par}_n$, we have

$$\begin{aligned}
h_\mu = \sum_{\lambda\in\mathrm{Par}} K_{\lambda,\mu} s_\lambda &= \underbrace{\sum_{\substack{\lambda\in\mathrm{Par}; \\ |\lambda|=n}}}_{=\sum_{\lambda\in\mathrm{Par}_n}} K_{\lambda,\mu} s_\lambda + \sum_{\substack{\lambda\in\mathrm{Par}; \\ |\lambda|\neq n}} \underbrace{K_{\lambda,\mu}}_{\substack{=0 \\ (\text{by } (13.72.4), \text{ since } |\lambda|\neq n=|\mu|)}} s_\lambda \\
&= \sum_{\lambda\in\mathrm{Par}_n} K_{\lambda,\mu} s_\lambda + \underbrace{\sum_{\substack{\lambda\in\mathrm{Par}; \\ |\lambda|\neq n}} 0 s_\lambda}_{=0} = \sum_{\lambda\in\mathrm{Par}_n} K_{\lambda,\mu} s_\lambda .
\end{aligned} \tag{13.72.5}$$

In the proof of Proposition 2.2.10, we showed that any two partitions λ and μ in Par_n satisfy $K_{\lambda,\mu} = 0$ unless $\lambda \triangleright \mu$. Hence, the $\mathrm{Par}_n \times \mathrm{Par}_n$-matrix $(K_{\lambda,\mu})_{(\mu,\lambda)\in\mathrm{Par}_n\times\mathrm{Par}_n}$ is triangular[619]. This matrix is furthermore unitriangular (since $K_{\lambda,\lambda} = 1$ for every partition λ, as shown in the proof of Proposition 2.2.10), and therefore invertibly triangular. But the family $(h_\lambda)_{\lambda\in\mathrm{Par}_n}$ expands in the family $(s_\lambda)_{\lambda\in\mathrm{Par}_n}$ through this matrix $(K_{\lambda,\mu})_{(\mu,\lambda)\in\mathrm{Par}_n\times\mathrm{Par}_n}$ (because of (13.72.5)). Therefore, the family $(h_\lambda)_{\lambda\in\mathrm{Par}_n}$ expands invertibly triangularly in the family $(s_\lambda)_{\lambda\in\mathrm{Par}_n}$ (since our matrix is invertibly triangular). Hence, Corollary 11.1.19(e) (applied to Λ_n, Par_n, $(h_\lambda)_{\lambda\in\mathrm{Par}_n}$ and $(s_\lambda)_{\lambda\in\mathrm{Par}_n}$ instead of M, S, $(e_s)_{s\in S}$ and $(f_s)_{s\in S}$) shows that the family $(h_\lambda)_{\lambda\in\mathrm{Par}_n}$ is a basis of the $\mathbf{k}$-module Λ_n if and only if the family $(s_\lambda)_{\lambda\in\mathrm{Par}_n}$ is a basis of the $\mathbf{k}$-module Λ_n. Therefore, the family $(h_\lambda)_{\lambda\in\mathrm{Par}_n}$ is a basis of the $\mathbf{k}$-module Λ_n (since the family $(s_\lambda)_{\lambda\in\mathrm{Par}_n}$ is a basis of the

[618] See Definition 11.1.16(c) for what this means.

[619] See Definition 11.1.7 for the meaning of this.

k-module Λ_n). Since this has been proven for all $n \in \mathbb{N}$, we can combine this to conclude that the family $(h_\lambda)_{\lambda \in \mathrm{Par}}$ is a graded basis of the graded **k**-module Λ. This solves Exercise 2.7.10(c).

Remark. Just as in the Remark after the solution of Exercise 2.5.20(b), we can use our solution of Exercise 2.7.10(c) to further prove that $(e_\lambda)_{\lambda \in \mathrm{Par}}$ is a **k**-basis of Λ (in a different way than we have done in the proof of Proposition 2.2.10).

13.73. **Solution to Exercise 2.7.11.** *Solution to Exercise 2.7.11.* Let us first observe that

$$\mathfrak{z}(h_n) = e_n \qquad \text{for every positive integer } n. \tag{13.73.1}$$

(In fact, $h_n = s_{(n)}$, and so the map $\mathfrak{z}$ maps h_n to $s_{(n)^t} = s_{(1^n)} = e_n$.)

(a) Fix $n \in \mathbb{N}$. We need to prove that

$$\mathfrak{z}(f h_n) = \mathfrak{z}(f) \cdot \mathfrak{z}(h_n) \qquad \text{for every } f \in \Lambda. \tag{13.73.2}$$

Since this equality is linear in f, it is clearly enough to prove it when f is of the form s_λ for some $\lambda \in \mathrm{Par}$ (because $(s_\lambda)_{\lambda \in \mathrm{Par}}$ is a **k**-basis of Λ). In other words, it is enough to show that every $\lambda \in \mathrm{Par}$ satisfies

$$\mathfrak{z}(s_\lambda h_n) = \mathfrak{z}(s_\lambda) \cdot \mathfrak{z}(h_n). \tag{13.73.3}$$

Thus, we will focus on proving (13.73.3) now.

We WLOG assume that $n \neq 0$, since otherwise (13.73.3) is obvious. Let $\lambda \in \mathrm{Par}$. The equality (2.7.1) yields

$$s_\lambda h_n = \sum_{\substack{\lambda^+ \,:\, \lambda^+/\lambda \text{ is a} \\ \text{horizontal } n\text{-strip}}} s_{\lambda^+}.$$

Applying the map $\mathfrak{z}$ to both sides of this equality, we obtain

$$\mathfrak{z}(s_\lambda h_n) = \mathfrak{z}\left(\sum_{\substack{\lambda^+ \,:\, \lambda^+/\lambda \text{ is a} \\ \text{horizontal } n\text{-strip}}} s_{\lambda^+} \right) = \sum_{\substack{\lambda^+ \,:\, \lambda^+/\lambda \text{ is a} \\ \text{horizontal } n\text{-strip}}} \underbrace{\mathfrak{z}(s_{\lambda^+})}_{= s_{(\lambda^+)^t}} = \sum_{\substack{\lambda^+ \,:\, \lambda^+/\lambda \text{ is a} \\ \text{horizontal } n\text{-strip}}} s_{(\lambda^+)^t}. \tag{13.73.4}$$

However, for any given partition λ^+, the assertion that λ^+/λ be a horizontal n-strip is equivalent to the assertion that $(\lambda^+)^t/\lambda^t$ be a vertical n-strip[620]. Hence, we can replace the summation sign "$\sum\limits_{\substack{\lambda^+ \,:\, \lambda^+/\lambda \text{ is a} \\ \text{horizontal } n\text{-strip}}}$" in (13.73.4) by a "$\sum\limits_{\substack{\lambda^+ \,:\, (\lambda^+)^t/\lambda^t \text{ is a} \\ \text{vertical } n\text{-strip}}}$" sign. Thus, we obtain

$$\sum_{\substack{\lambda^+ \,:\, \lambda^+/\lambda \text{ is a} \\ \text{horizontal } n\text{-strip}}} s_{(\lambda^+)^t} = \sum_{\substack{\lambda^+ \,:\, (\lambda^+)^t/\lambda^t \text{ is a} \\ \text{vertical } n\text{-strip}}} s_{(\lambda^+)^t} = \sum_{\substack{\lambda^+ \,:\, \lambda^+/\lambda^t \text{ is a} \\ \text{vertical } n\text{-strip}}} s_{\lambda^+}$$

$$\left(\begin{array}{c} \text{here, we substituted } \lambda^+ \text{ for } (\lambda^+)^t \text{ in the sum, because} \\ \text{the map } \mathrm{Par} \to \mathrm{Par},\ \mu \mapsto \mu^t \text{ is a bijection} \end{array} \right).$$

Thus, (13.73.4) becomes

$$\mathfrak{z}(s_\lambda h_n) = \sum_{\substack{\lambda^+ \,:\, \lambda^+/\lambda \text{ is a} \\ \text{horizontal } n\text{-strip}}} s_{(\lambda^+)^t} = \sum_{\substack{\lambda^+ \,:\, \lambda^+/\lambda^t \text{ is a} \\ \text{vertical } n\text{-strip}}} s_{\lambda^+}. \tag{13.73.5}$$

On the other hand, the equality (2.7.2) yields

$$s_\lambda e_n = \sum_{\substack{\lambda^+ \,:\, \lambda^+/\lambda \text{ is a} \\ \text{vertical } n\text{-strip}}} s_{\lambda^+}.$$

[620] This is because the notion of a vertical n-strip is obtained from the notion of a horizontal n-strip by interchanging the roles of rows and columns, and the operation which sends a partition μ to its transpose partition μ^t interchanges the roles of rows and columns as well.

Applying this to λ^t instead of λ, we obtain

$$s_{\lambda^t} e_n = \sum_{\substack{\lambda^+ \ : \ \lambda^+/\lambda^t \text{ is a} \\ \text{vertical } n\text{-strip}}} s_{\lambda^+}.$$

Now,

$$\underbrace{\mathfrak{Z}\left(s_\lambda\right)}_{=s_{\lambda^t}} \cdot \underbrace{\mathfrak{Z}\left(h_n\right)}_{=e_n} = s_{\lambda^t} e_n = \sum_{\substack{\lambda^+ \ : \ \lambda^+/\lambda^t \text{ is a} \\ \text{vertical } n\text{-strip}}} s_{\lambda^+}.$$

Compared with (13.73.5), this yields $\mathfrak{Z}\left(s_\lambda h_n\right) = \mathfrak{Z}\left(s_\lambda\right) \cdot \mathfrak{Z}\left(h_n\right)$. This proves (13.73.3). This concludes the solution of Exercise 2.7.11(a).

(b) We shall show that

$$\mathfrak{Z}\left(h_\lambda\right) = \omega\left(h_\lambda\right) \qquad \text{for every partition } \lambda. \tag{13.73.6}$$

Proof of (13.73.6)*:* We will prove (13.73.6) by induction over $\ell\left(\lambda\right)$. The *induction base* is the case $\ell\left(\lambda\right) = 0$, which is utterly obvious. For the *induction step*, we fix a positive integer L and assume (as the induction hypothesis) that the equality (13.73.6) holds if $\ell\left(\lambda\right) = L - 1$. We need to prove that the equality (13.73.6) holds if $\ell\left(\lambda\right) = L$.

Let λ be a partition satisfying $\ell\left(\lambda\right) = L$. Write λ in the form $\lambda = \left(\lambda_1, \lambda_2, ..., \lambda_L\right)$ with all of λ_1, λ_2, ..., λ_L being positive integers. Let $\overline{\lambda}$ be the partition $\left(\lambda_1, \lambda_2, ..., \lambda_{L-1}\right)$ (this is obviously well-defined); then, $\ell\left(\overline{\lambda}\right) = L - 1$.

Recall that $e_n = \omega\left(h_n\right)$ for every $n \geq 1$ (by Proposition 2.4.3(b)).

Recall that $\lambda = \left(\lambda_1, \lambda_2, ..., \lambda_L\right)$. Hence, by the definition of h_λ, we have

$$h_\lambda = h_{\lambda_1} h_{\lambda_2} ... h_{\lambda_L} = \underbrace{h_{\lambda_1} h_{\lambda_2} ... h_{\lambda_{L-1}}}_{\substack{= h_{\overline{\lambda}} \\ \left(\text{since } \overline{\lambda} = \left(\lambda_1, \lambda_2, ..., \lambda_{L-1}\right)\right)}} h_{\lambda_L} = h_{\overline{\lambda}} h_{\lambda_L}.$$

Applying the map $\mathfrak{Z}$ to both sides of this equality, we obtain

$$\begin{aligned}
\mathfrak{Z}\left(h_\lambda\right) = \mathfrak{Z}\left(h_{\overline{\lambda}} h_{\lambda_L}\right) &= \underbrace{\mathfrak{Z}\left(h_{\overline{\lambda}}\right)}_{\substack{= \omega\left(h_{\overline{\lambda}}\right) \\ \text{(by the induction} \\ \text{hypothesis,} \\ \text{since } \ell\left(\overline{\lambda}\right) = L-1)}} \cdot \underbrace{\mathfrak{Z}\left(h_{\lambda_L}\right)}_{\substack{= e_{\lambda_L} \\ \text{(by (13.73.1))}}} \\
&\qquad \left(\text{by Exercise 2.7.11(a), applied to } h_{\overline{\lambda}} \text{ and } \lambda_L \text{ instead of } f \text{ and } n\right) \\
&= \omega\left(h_{\overline{\lambda}}\right) \cdot \underbrace{e_{\lambda_L}}_{\substack{= \omega\left(h_{\lambda_L}\right) \\ \text{(since } e_n = \omega\left(h_n\right) \\ \text{for every } n \geq 1)}} = \omega\left(h_{\overline{\lambda}}\right) \cdot \omega\left(h_{\lambda_L}\right) = \omega\left(\underbrace{h_{\overline{\lambda}} h_{\lambda_L}}_{= h_\lambda}\right) \qquad \text{(since } \omega \text{ is a } \mathbf{k}\text{-algebra morphism)} \\
&= \omega\left(h_\lambda\right).
\end{aligned}$$

In other words, the equality (13.73.6) holds if $\ell\left(\lambda\right) = L$. This completes the induction step. The induction proof of (13.73.6) is therefore complete.

The equality (13.73.6) shows that the $\mathbf{k}$-linear maps $\mathfrak{Z} : \Lambda \to \Lambda$ and $\omega : \Lambda \to \Lambda$ are equal to each other on every element of the basis $\left(h_\lambda\right)_{\lambda \in \mathrm{Par}}$ of the $\mathbf{k}$-module Λ. Hence, $\mathfrak{Z} = \omega$. This solves Exercise 2.7.11(b).

(c) Fix two partitions μ and ν. Then, (2.5.6) yields

$$s_\mu s_\nu = \sum_{\lambda \in \mathrm{Par}} c^\lambda_{\mu,\nu} s_\lambda. \tag{13.73.7}$$

Applying this to μ^t and ν^t instead of μ and ν, we obtain

$$s_{\mu^t} s_{\nu^t} = \sum_{\lambda \in \mathrm{Par}} c^\lambda_{\mu^t,\nu^t} s_\lambda. \tag{13.73.8}$$

Applying the map $\mathfrak{Z}$ to the identity (13.73.7), we obtain

$$\begin{aligned}
\mathfrak{Z}(s_\mu s_\nu) &= \mathfrak{Z}\left(\sum_{\lambda\in\operatorname{Par}} c^{\lambda}_{\mu,\nu} s_\lambda\right) = \sum_{\lambda\in\operatorname{Par}} c^{\lambda}_{\mu,\nu} \underbrace{\mathfrak{Z}(s_\lambda)}_{=s_{\lambda^t}} \\
&= \sum_{\lambda\in\operatorname{Par}} \underbrace{c^{\lambda}_{\mu,\nu}}_{\substack{=c^{(\lambda^t)^t}_{\mu,\nu} \\ (\text{since } \lambda=(\lambda^t)^t)}} s_{\lambda^t} = \sum_{\lambda\in\operatorname{Par}} c^{(\lambda^t)^t}_{\mu,\nu} s_{\lambda^t} = \sum_{\lambda\in\operatorname{Par}} c^{\lambda^t}_{\mu,\nu} s_\lambda \\
&\qquad \left(\begin{array}{c}\text{here, we substituted } \lambda \text{ for } \lambda^t \text{ in the sum, since} \\ \text{the map } \operatorname{Par}\to\operatorname{Par},\ \lambda\mapsto\lambda^t \text{ is a bijection}\end{array}\right).
\end{aligned} \tag{13.73.9}$$

But recall that ω is a **k**-algebra homomorphism. Since $\mathfrak{Z}=\omega$ (by Exercise 2.7.11(b)), this shows that $\mathfrak{Z}$ is a **k**-algebra homomorphism. Hence,

$$\mathfrak{Z}(s_\mu s_\nu) = \underbrace{\mathfrak{Z}(s_\mu)}_{=s_{\mu^t}} \cdot \underbrace{\mathfrak{Z}(s_\nu)}_{=s_{\nu^t}} = s_{\mu^t} s_{\nu^t} = \sum_{\lambda\in\operatorname{Par}} c^{\lambda}_{\mu^t,\nu^t} s_\lambda \qquad (\text{by } (13.73.8)).$$

Compared with (13.73.9), this yields $\sum_{\lambda\in\operatorname{Par}} c^{\lambda^t}_{\mu,\nu} s_\lambda = \sum_{\lambda\in\operatorname{Par}} c^{\lambda}_{\mu^t,\nu^t} s_\lambda$. Since $(s_\lambda)_{\lambda\in\operatorname{Par}}$ is a basis of the **k**-module Λ, we can compare coefficients in this equality, and obtain

$$c^{\lambda^t}_{\mu,\nu} = c^{\lambda}_{\mu^t,\nu^t} \qquad \text{for every } \lambda\in\operatorname{Par}. \tag{13.73.10}$$

We can substitute λ^t for λ in this result, and conclude that $c^{(\lambda^t)^t}_{\mu,\nu} = c^{\lambda^t}_{\mu^t,\nu^t}$ for every $\lambda\in\operatorname{Par}$. Since $(\lambda^t)^t=\lambda$, this rewrites as $c^{\lambda}_{\mu,\nu} = c^{\lambda^t}_{\mu^t,\nu^t}$. This solves Exercise 2.7.11(c).

(d) Let μ and λ be two partitions such that $\mu\subseteq\lambda$. In Remark 2.5.9, it has been shown that $s_{\lambda/\mu} = \sum_\nu c^{\lambda}_{\mu,\nu} s_\nu$, where the sum ranges over all partitions ν. In other words,

$$s_{\lambda/\mu} = \sum_{\nu\in\operatorname{Par}} c^{\lambda}_{\mu,\nu} s_\nu. \tag{13.73.11}$$

Applying this to λ^t and μ^t instead of λ and μ, we obtain

$$s_{\lambda^t/\mu^t} = \sum_{\nu\in\operatorname{Par}} c^{\lambda^t}_{\mu^t,\nu} s_\nu. \tag{13.73.12}$$

Applying the map $\mathfrak{Z}$ to both sides of the identity (13.73.11), we obtain

$$\begin{aligned}
\mathfrak{Z}(s_{\lambda/\mu}) &= \mathfrak{Z}\left(\sum_{\nu\in\operatorname{Par}} c^{\lambda}_{\mu,\nu} s_\nu\right) = \sum_{\nu\in\operatorname{Par}} \underbrace{c^{\lambda}_{\mu,\nu}}_{\substack{=c^{\lambda^t}_{\mu^t,\nu^t} \\ (\text{by Exercise 2.7.11(c)})}} \underbrace{\mathfrak{Z}(s_\nu)}_{=s_{\nu^t}} = \sum_{\nu\in\operatorname{Par}} c^{\lambda^t}_{\mu^t,\nu^t} s_{\nu^t} = \sum_{\nu\in\operatorname{Par}} c^{\lambda^t}_{\mu^t,\nu} s_\nu \\
&\qquad \left(\begin{array}{c}\text{here, we substituted } \nu \text{ for } \nu^t \text{ in the sum, since} \\ \text{the map } \operatorname{Par}\to\operatorname{Par},\ \nu\mapsto\nu^t \text{ is a bijection}\end{array}\right) \\
&= s_{\lambda^t/\mu^t} \qquad (\text{by } (13.73.12)).
\end{aligned}$$

Since $\mathfrak{Z}=\omega$ (by Exercise 2.7.11(b)), this rewrites as $\omega(s_{\lambda/\mu}) = s_{\lambda^t/\mu^t}$. This proves the first identity of (2.4.15).

It is clear that the skew Schur function $s_{\lambda/\mu}$ is homogeneous of degree $|\lambda/\mu|$. In other words, $s_{\lambda/\mu}\in\Lambda_{|\lambda/\mu|}$. Hence, (2.4.11) (applied to $f=s_{\lambda/\mu}$ and $n=|\lambda/\mu|$) yields

$$S(s_{\lambda/\mu}) = (-1)^{|\lambda/\mu|} \underbrace{\omega(s_{\lambda/\mu})}_{=s_{\lambda^t/\mu^t}} = (-1)^{|\lambda/\mu|} s_{\lambda^t/\mu^t}.$$

This proves the second identity of (2.4.15). The proof of (2.4.15) is thus complete.

13.74. **Solution to Exercise 2.7.12.** *Solution to Exercise 2.7.12.* (a)

First solution to Exercise 2.7.12(a): Let us first show that $\prod_{i,j=1}^{\infty}(1+x_iy_j) = \sum_{\lambda\in\mathrm{Par}} e_\lambda(\mathbf{x})\, m_\lambda(\mathbf{y})$.

In fact, in the proof of Proposition 2.5.15, we showed that $\prod_{j=1}^{\infty}\sum_{n\geq 0} h_n(\mathbf{x})\, y_j^n = \sum_{\lambda\in\mathrm{Par}} h_\lambda(\mathbf{x})\, m_\lambda(\mathbf{y})$. The same argument (with all appearances of the letter "h" replaced by the letter "e") shows that

$$\prod_{j=1}^{\infty}\sum_{n\geq 0} e_n(\mathbf{x})\, y_j^n = \sum_{\lambda\in\mathrm{Par}} e_\lambda(\mathbf{x})\, m_\lambda(\mathbf{y}). \tag{13.74.1}$$

But (2.2.19) yields

$$\prod_{i=1}^{\infty}(1+x_it) = \sum_{n\geq 0} e_n(\mathbf{x})\, t^n \tag{13.74.2}$$

in the ring $(\mathbf{k}[[\mathbf{x}]])[[t]]$. For every $j\in\{1,2,3,\ldots\}$, we have

$$\prod_{i=1}^{\infty}(1+x_iy_j) = \sum_{n\geq 0} e_n(\mathbf{x})\, y_j^n \tag{13.74.3}$$

in the ring $(\mathbf{k}[[\mathbf{x}]])[[\mathbf{y}]]$ (in fact, this results from (13.74.2) by substituting y_j for t). Thus, (13.74.3) holds in $\mathbf{k}[[\mathbf{x},\mathbf{y}]]$ (since $(\mathbf{k}[[\mathbf{x}]])[[\mathbf{y}]] = \mathbf{k}[[\mathbf{x},\mathbf{y}]]$ as rings). Now,

$$\underbrace{\prod_{i,j=1}^{\infty}}_{=\prod_{j=1}^{\infty}\prod_{i=1}^{\infty}}(1+x_iy_j) = \prod_{j=1}^{\infty}\underbrace{\prod_{i=1}^{\infty}(1+x_iy_j)}_{\substack{=\sum_{n\geq 0}e_n(\mathbf{x})y_j^n \\ \text{(by (13.74.3))}}} = \prod_{j=1}^{\infty}\sum_{n\geq 0} e_n(\mathbf{x})\, y_j^n$$

$$= \sum_{\lambda\in\mathrm{Par}} e_\lambda(\mathbf{x})\, m_\lambda(\mathbf{y}) \qquad (\text{by (13.74.1)}). \tag{13.74.4}$$

Next, we shall show that $\sum_{\lambda\in\mathrm{Par}} s_\lambda(\mathbf{x})\, s_{\lambda^t}(\mathbf{y}) = \sum_{\lambda\in\mathrm{Par}} e_\lambda(\mathbf{x})\, m_\lambda(\mathbf{y})$.

In fact, consider the $\mathbf{k}$-algebra $(\mathbf{k}[[\mathbf{x}]])[[\mathbf{y}]] = \mathbf{k}[[\mathbf{x},\mathbf{y}]]$. This $\mathbf{k}$-algebra $(\mathbf{k}[[\mathbf{x}]])[[\mathbf{y}]]$ is a $\mathbf{k}[[\mathbf{y}]]$-algebra, and contains $\Lambda[[\mathbf{y}]]$ as a $\mathbf{k}[[\mathbf{y}]]$-subalgebra.

The equality (2.5.1) yields

$$\sum_{\lambda\in\mathrm{Par}} s_\lambda(\mathbf{x})\, s_\lambda(\mathbf{y}) = \prod_{i,j=1}^{\infty}(1-x_iy_j)^{-1} = \sum_{\lambda\in\mathrm{Par}} h_\lambda(\mathbf{x})\, m_\lambda(\mathbf{y}) \qquad (\text{by (2.5.11)}). \tag{13.74.5}$$

This is an equality in $\Lambda[[\mathbf{y}]]$ (because $s_\lambda(\mathbf{x})$ and $h_\lambda(\mathbf{x})$ belong to Λ for every $\lambda\in\mathrm{Par}$).

Recall that $\omega:\Lambda\to\Lambda$ is a $\mathbf{k}$-algebra homomorphism. It thus induces a $\mathbf{k}[[\mathbf{y}]]$-algebra homomorphism $\omega[[\mathbf{y}]]:\Lambda[[\mathbf{y}]]\to\Lambda[[\mathbf{y}]]$ which sends every $q\in\Lambda$ to $\omega(q)$, and is continuous with respect to the usual topology[621] on $\Lambda[[\mathbf{y}]]$. Applying this homomorphism $\omega[[\mathbf{y}]]$ to both sides of (13.74.5), we obtain

$$\sum_{\lambda\in\mathrm{Par}} \omega(s_\lambda(\mathbf{x}))\, s_\lambda(\mathbf{y}) = \sum_{\lambda\in\mathrm{Par}} \omega(h_\lambda(\mathbf{x}))\, m_\lambda(\mathbf{y}) \tag{13.74.6}$$

(because $\omega[[\mathbf{y}]]$, being a $\mathbf{k}[[\mathbf{y}]]$-algebra homomorphism, leaves the $s_\lambda(\mathbf{y})$ and $m_\lambda(\mathbf{y})$ terms unchanged, while the $s_\lambda(\mathbf{x})$ and $h_\lambda(\mathbf{x})$ terms are elements of Λ and thus are transformed as by ω).

But we know that

$$\omega(s_\lambda) = s_{\lambda^t} \qquad \text{for every partition } \lambda. \tag{13.74.7}$$

(This follows from the first equality in (2.4.15) by setting $\mu=\varnothing$.) Using (2.4.12) and (13.74.7), we can rewrite (13.74.6) as

$$\sum_{\lambda\in\mathrm{Par}} s_{\lambda^t}(\mathbf{x})\, s_\lambda(\mathbf{y}) = \sum_{\lambda\in\mathrm{Par}} e_\lambda(\mathbf{x})\, m_\lambda(\mathbf{y}).$$

[621] The usual topology on a power series ring $Z[[\mathbf{y}]]$ (where Z is a commutative ring) is the direct-product topology obtained by viewing the set $Z[[\mathbf{y}]]$ as a direct product of many copies of Z (this is done by identifying every power series with the family of its coefficients), each of which is endowed with the discrete topology. In this topology, a sequence (or, more generally, a net) $(f_i)_i$ of power series converges to a power series f if and only if for every monomial $\mathfrak{m}$, the sequence (resp. net) (the coefficient of $\mathfrak{m}$ in $f_i)_i$ converges to (the coefficient of $\mathfrak{m}$ in f) with respect to the discrete topology. (The notion of a *net* is a generalization of the notion of a sequence; it is useful in topology. See [219] for an introduction to it.)

Hence,

$$\sum_{\lambda\in\operatorname{Par}} e_\lambda(\mathbf{x})\, m_\lambda(\mathbf{y}) = \sum_{\lambda\in\operatorname{Par}} s_{\lambda^t}(\mathbf{x})\, s_\lambda(\mathbf{y}) = \sum_{\lambda\in\operatorname{Par}} \underbrace{s_{(\lambda^t)^t}(\mathbf{x})}_{\substack{=s_\lambda(\mathbf{x}) \\ (\text{since } (\lambda^t)^t=\lambda)}} s_{\lambda^t}(\mathbf{y})$$

$$\left(\begin{array}{c}\text{here, we substituted } \lambda^t \text{ for } \lambda \text{ in the sum, since the map} \\ \operatorname{Par}\to\operatorname{Par},\ \lambda\mapsto\lambda^t \text{ is a bijection}\end{array}\right)$$

$$= \sum_{\lambda\in\operatorname{Par}} s_\lambda(\mathbf{x})\, s_{\lambda^t}(\mathbf{y}). \tag{13.74.8}$$

Combined with (13.74.4), this yields

$$\prod_{i,j=1}^{\infty}(1+x_iy_j) = \sum_{\lambda\in\operatorname{Par}} e_\lambda(\mathbf{x})\, m_\lambda(\mathbf{y}) = \sum_{\lambda\in\operatorname{Par}} s_\lambda(\mathbf{x})\, s_{\lambda^t}(\mathbf{y}). \tag{13.74.9}$$

This solves Exercise 2.7.12(a).

Second solution to Exercise 2.7.12(a): We can prove (13.74.4) as in the First solution to Exercise 2.7.12(a). Thus, in order to solve Exercise 2.7.12(a), it remains to verify

$$\prod_{i,j=1}^{\infty}(1+x_iy_j) = \sum_{\lambda\in\operatorname{Par}} s_\lambda(\mathbf{x})\, s_{\lambda^t}(\mathbf{y}). \tag{13.74.10}$$

Instead of proving (13.74.10), we will verify the identity

$$\prod_{i,j=1}^{\infty}(1+tx_iy_j) = \sum_{\lambda\in\operatorname{Par}} t^{|\lambda|} s_\lambda(\mathbf{x})\, s_{\lambda^t}(\mathbf{y}) \tag{13.74.11}$$

in the ring $R(\mathbf{x},\mathbf{y})[[t]]$. This identity will clearly yield (13.74.10) (by substituting $t=1$), and thus conclude the solution of Exercise 2.7.12(a).

We shall prove (13.74.11) in a similar way to how we proved (2.5.2), but using a variation on the RSK correspondence. This is not in itself a new idea; indeed, this is how the equivalent identity (13.74.10) is proven in [206, Theorem 7.14.3], in [111, §7] and in [186, Theorem 4.8.6]. However, instead of using the dual RSK algorithm (also known as the RSK* algorithm; see [206, §7.14], [111, §5] and [186, Theorem 4.8.5] for it) like these proofs do, we introduce a variation of the RSK algorithm that relies on the same RS-insertion operation but changes the order in which the biletters are processed. (We will reprove (13.74.11) using the dual RSK algorithm in the Third solution further below.)

For a given partition λ, let us define a *row-strict tableau* of shape λ to be an assignment T of entries in $\{1,2,3,\ldots\}$ to the cells of the Ferrers diagram for λ which is strictly increasing left-to-right in rows, and weakly increasing top-to-bottom in columns. It is clear that if λ is a partition, then the row-strict tableaux of shape λ are in 1-to-1 correspondence with the column-strict tableaux of shape λ^t, and the correspondence is given by transposing the tableau (i.e., taking whatever entry was assigned to a cell (i,j) in the input tableau, and reassigning it to the cell (j,i) in the output tableau). Hence, for every partition λ, we have

$$\sum_{\substack{Q \text{ is a row-strict} \\ \text{tableau of shape } \lambda}} \mathbf{x}^{\operatorname{cont}(Q)} = \sum_{\substack{T \text{ is a column-strict} \\ \text{tableau of shape } \lambda^t}} \mathbf{x}^{\operatorname{cont}(T)}$$

$$= s_{\lambda^t} \qquad (\text{since this is how } s_{\lambda^t} \text{ is defined}) \tag{13.74.12}$$

(where $\mathbf{x}^{\operatorname{cont}(Q)}$ is defined for a row-strict tableau Q in the same way as it is defined for a column-strict tableau Q). Substituting $\mathbf{y}$ for $\mathbf{x}$ in this equality, we obtain

$$\sum_{\substack{Q \text{ is a row-strict} \\ \text{tableau of shape } \lambda}} \mathbf{y}^{\operatorname{cont}(Q)} = s_{\lambda^t}(\mathbf{y}). \tag{13.74.13}$$

We also have

$$\sum_{\substack{P \text{ is a column-strict} \\ \text{tableau of shape } \lambda}} \mathbf{x}^{\operatorname{cont}(P)} = s_\lambda(\mathbf{x}) \tag{13.74.14}$$

(because this is how $s_\lambda = s_\lambda(\mathbf{x})$ is defined) for every partition λ.

A *tableau-cotableau pair* will mean a pair (P, Q) such that P is a column-strict tableau, Q is a row-strict tableau, and P and Q both have shape λ for one and the same partition λ. Multiplying the identities (13.74.14) and (13.74.13) and multiplying the result with $t^{|\lambda|}$, we obtain

$$t^{|\lambda|} \sum_{\substack{(P,Q) \text{ is a pair with} \\ P \text{ being a column-strict tableau} \\ \text{of shape } \lambda, \text{ and } Q \text{ being} \\ \text{a row-strict tableau of shape } \lambda}} \mathbf{x}^{\operatorname{cont}(P)} \mathbf{y}^{\operatorname{cont}(Q)} = t^{|\lambda|} s_\lambda(\mathbf{x}) s_{\lambda^t}(\mathbf{y})$$

for every partition λ. Summing this equality over all partitions λ, we obtain

$$\sum t^{|\lambda|} \mathbf{x}^{\operatorname{cont}(P)} \mathbf{y}^{\operatorname{cont}(Q)} = \sum_{\lambda \in \operatorname{Par}} t^{|\lambda|} s_\lambda(\mathbf{x}) s_{\lambda^t}(\mathbf{y}), \tag{13.74.15}$$

where the sum on the left hand side is over all tableau-cotableau pairs (P, Q) and where λ denotes the common shape of P and Q.

We shall use all notations introduced in the proof of Theorem 2.5.1. Define the *antilexicographic order* $\leq_{alex}$ to be the total order on the set of all biletters which is given by

$$\binom{i_1}{j_1} \leq_{alex} \binom{i_2}{j_2} \iff (\text{we have } i_1 \leq i_2, \text{ and if } i_1 = i_2, \text{ then } j_1 \geq j_2).$$

We denote by $<_{alex}$ the (strict) smaller relation of this order. A *strict cobiword* will mean an array $\binom{\mathbf{i}}{\mathbf{j}} = \binom{i_1 \ldots i_\ell}{j_1 \ldots j_\ell}$ in which the biletters satisfy $\binom{i_1}{j_1} <_{alex} \cdots <_{alex} \binom{i_\ell}{j_\ell}$ (that is, the biletters are distinct and ordered). Strict cobiwords are clearly in 1-to-1 correspondence with sets (not multisets!) of biletters. Now, the left hand side of (13.74.11) is $\prod_{i,j=1}^{\infty}(1 + t x_i y_j) = \prod_{i,j=1}^{\infty}(1 + t x_j y_i)$ (here, we substituted (j, i) for the index (i, j) in the product), and thus can be rewritten as the sum of $t^\ell (x_{j_1} y_{i_1})(x_{j_2} y_{i_2}) \cdots (x_{j_\ell} y_{i_\ell})$ over all sets $\left\{\binom{i_1}{j_1}, \ldots, \binom{i_\ell}{j_\ell}\right\}$ of biletters. Thus, the left hand side of (13.74.11) is the sum $\sum t^\ell \mathbf{x}^{\operatorname{cont}(\mathbf{j})} \mathbf{y}^{\operatorname{cont}(\mathbf{i})}$ over all strict cobiwords $\binom{\mathbf{i}}{\mathbf{j}}$, where ℓ stands for the number of biletters in the strict cobiword. Meanwhile, we know that the right hand side of (13.74.11) is the sum $\sum t^{|\lambda|} \mathbf{x}^{\operatorname{cont}(P)} \mathbf{y}^{\operatorname{cont}(Q)}$ over all tableau-cotableau pairs (P, Q) (because of (13.74.15)). Thus, in order to prove (13.74.11), we only need to construct a bijection between the strict cobiwords $\binom{\mathbf{i}}{\mathbf{j}}$ and the tableau-cotableau pairs (P, Q), which has the property that

$$\operatorname{cont}(\mathbf{i}) = \operatorname{cont}(Q);$$
$$\operatorname{cont}(\mathbf{j}) = \operatorname{cont}(P).$$

[622]

This bijection is the *coRSK algorithm*, which we shall now define.[623]

Let $\binom{\mathbf{i}}{\mathbf{j}}$ be a strict cobiword. Starting with the pair $(P_0, Q_0) = (\varnothing, \varnothing)$ and $m = 0$, the algorithm applies the following steps (see Example 13.74.1 below):

- If i_{m+1} does not exist (that is, m is the length of $\mathbf{i}$), stop.
- Apply RS-insertion to the column-strict tableau P_m and the letter j_{m+1} (the bottom letter of $\binom{i_{m+1}}{j_{m+1}}$). Let P_{m+1} be the resulting column-strict tableau, and let c_{m+1} be the resulting corner cell.
- Create Q_{m+1} from Q_m by adding the top letter i_{m+1} of $\binom{i_{m+1}}{j_{m+1}}$ to Q_m in the cell c_{m+1} (which, as we recall, is the extra corner cell of P_{m+1} not present in P_m).
- Set m to $m + 1$.

After all of the biletters have been thus processed, the result of the coRSK algorithm is $(P_\ell, Q_\ell) =: (P, Q)$.

Example 13.74.1. The term in the expansion of the left side of (13.74.10) corresponding to

$$(x_4 y_1)(x_2 y_1)(x_1 y_2)(x_3 y_4)(x_1 y_4)(x_2 y_5)$$

[622] Such a bijection will then automatically satisfy $|\lambda| = \ell$, where λ is the (common) shape of P and Q, and where ℓ is the length of the strict cobiword $\binom{\mathbf{i}}{\mathbf{j}}$. This is because $|\lambda| = |\operatorname{cont}(Q)|$ and $\ell = |\operatorname{cont}(\mathbf{i})|$.

[623] This algorithm appears in Fulton's [73, §A.4] as construction (1d).

is the strict cobiword $\binom{\mathbf{i}}{\mathbf{j}} = \binom{112445}{421312}$, and the coRSK algorithm applied to this cobiword proceeds as follows:

$$\begin{array}{rcl|rcl}
P_0 & = & \varnothing & Q_0 & = & \varnothing \\
P_1 & = & 4 & Q_1 & = & 1 \\
P_2 & = & \begin{array}{l} 2 \\ 4 \end{array} & Q_2 & = & \begin{array}{l} 1 \\ 1 \end{array} \\
P_3 & = & \begin{array}{l} 1 \\ 2 \\ 4 \end{array} & Q_3 & = & \begin{array}{l} 1 \\ 1 \\ 3 \end{array} \\
P_4 & = & \begin{array}{ll} 1 & 3 \\ 2 & \\ 4 & \end{array} & Q_4 & = & \begin{array}{ll} 1 & 4 \\ 1 & \\ 2 & \end{array} \\
P_5 & = & \begin{array}{ll} 1 & 1 \\ 2 & 3 \\ 4 & \end{array} & Q_5 & = & \begin{array}{ll} 1 & 4 \\ 1 & 4 \\ 2 & \end{array} \\
P := P_6 & = & \begin{array}{lll} 1 & 1 & 2 \\ 2 & 3 & \\ 4 & & \end{array} & Q := Q_6 & = & \begin{array}{lll} 1 & 4 & 5 \\ 1 & 4 & \\ 2 & & \end{array}
\end{array}$$

It is clear that P_m remains a column-strict tableau of some Ferrers shape throughout the execution of the coRSK algorithm, and that Q_m remains a filling of the same shape as P_m which is (at least) weakly increasing left-to-right along rows and weakly increasing top-to-bottom in columns. But we can also see that Q_m is strictly increasing left-to-right along rows[624], so that Q_m is a row-strict tableau. Thus, the result (P, Q) of the coRSK algorithm is a tableau-cotableau pair.

To see that the coRSK algorithm is a bijection, we show how to recover $\binom{\mathbf{i}}{\mathbf{j}}$ from (P, Q). This is done by reverse bumping in the same way as for the usual RSK algorithm, with the only difference that now Q_m is obtained by removing the bottommost (rather than the rightmost) occurrence of the letter i_{m+1} from Q_{m+1}.[625]

Finally, to see that the coRSK map is surjective, one needs to show that the reverse bumping procedure can be applied to any tableau-cotableau pair (P, Q), and will result in a strict cobiword $\binom{\mathbf{i}}{\mathbf{j}}$. We leave this verification to the reader.[626]

[624] Indeed, this follows from the observation that when one has a string of equal letters $i_m = i_{m+1} = \cdots = i_{m+r}$ on top of the strict cobiword, then the bottom letters bumped in are $j_m > j_{m+1} > \cdots > j_{m+r}$, and therefore (as a consequence of the last claim of part (b) of the row bumping lemma) the new cells form a *vertical strip*, that is, no two of these cells lie in the same row. Actually, more can be said: Each of these new cells (except for the first one) is in a row further down than the previous one. We will use this stronger fact further below.

[625] It necessarily has to be the bottommost occurrence, since (according to the previous footnote) the cell into which i_{m+1} was filled at the step from Q_m to Q_{m+1} lies further down than any existing cell of Q_m containing the letter i_{m+1}.

[626] It is easy to see that repeatedly applying reverse bumping to (P, Q) will result in a sequence $\binom{i_\ell}{j_\ell}, \binom{i_{\ell-1}}{j_{\ell-1}}, \ldots, \binom{i_1}{j_1}$ of biletters such that applying the coRSK algorithm to $\binom{i_1 \cdots i_\ell}{j_1 \cdots j_\ell}$ gives back (P, Q). The question is why we have $\binom{i_1}{j_1} <_{alex} \cdots <_{alex} \binom{i_\ell}{j_\ell}$. Since the chain of inequalities $i_1 \leq i_2 \leq \cdots \leq i_\ell$ is clear from the choice of entry to reverse-bump, it only remains to show that for every string $i_m = i_{m+1} = \cdots = i_{m+r}$ of equal top letters, the corresponding bottom letters strictly decrease (that is, $j_m > j_{m+1} > \cdots > j_{m+r}$). One way to see this is the following:

Assume the contrary; i.e., assume that the bottom letters corresponding to some string $i_m = i_{m+1} = \cdots = i_{m+r}$ of equal top letters do not strictly decrease. Thus, $j_{m+p} \leq j_{m+p+1}$ for some $p \in \{0, 1, \ldots, r-1\}$. Consider this p.

Let us consider the cells containing the equal letters $i_m = i_{m+1} = \cdots = i_{m+r}$ in the tableau Q_{m+r}. Label these cells as $c_m, c_{m+1}, \ldots, c_{m+r}$ from top to bottom (noticing that no two of them lie in the same row, since Q_{m+r} is row-strict). By the definition of reverse bumping, the first entry to be reverse bumped from P_{m+r} is the entry in position c_{m+r} (since this is the bottommost occurrence of the letter i_{m+r} in Q_{m+r}); then, the next entry to be reverse bumped is the one in position c_{m+r-1}, So the coRSK map is a bijection having the required properties. As we have said, this proves (13.74.11). This solves Exercise 2.7.12(a) again.

[*Remark:* By combining the two solutions of Exercise 2.7.12(a) given above, one can obtain a new proof of the equality (13.74.7) (a proof which does not use (2.4.15)). Indeed, using (2.4.12), we can rewrite (13.74.6) as

$$\sum_{\lambda\in\operatorname{Par}} \omega\left(s_\lambda\left(\mathbf{x}\right)\right) s_\lambda\left(\mathbf{y}\right) = \sum_{\lambda\in\operatorname{Par}} e_\lambda\left(\mathbf{x}\right) m_\lambda\left(\mathbf{y}\right).$$

Compared with (13.74.4), this yields

$$\begin{aligned} &\sum_{\lambda\in\operatorname{Par}} \omega\left(s_\lambda\left(\mathbf{x}\right)\right) s_\lambda\left(\mathbf{y}\right) \\ &= \prod_{i,j=1}^{\infty}\left(1+x_iy_j\right) = \prod_{i,j=1}^{\infty}\left(1+x_jy_i\right) \qquad \text{(here, we substituted } (j,i) \text{ for } (i,j) \text{ in the sum)} \\ &= \prod_{i,j=1}^{\infty}\left(1+y_ix_j\right) = \sum_{\lambda\in\operatorname{Par}} s_\lambda\left(\mathbf{y}\right) s_{\lambda^t}\left(\mathbf{x}\right) \qquad \text{(by (13.74.10), with } \mathbf{x} \text{ and } \mathbf{y} \text{ substituted for } \mathbf{y} \text{ and } \mathbf{x}\text{)} \\ &= \sum_{\lambda\in\operatorname{Par}} s_{\lambda^t}\left(\mathbf{x}\right) s_\lambda\left(\mathbf{y}\right). \end{aligned}$$

Since the $s_\lambda\left(\mathbf{y}\right)$ for $\lambda\in\operatorname{Par}$ are linearly independent over $\mathbf{k}\left[\left[\mathbf{x}\right]\right]$, we can compare coefficients before $s_\lambda\left(\mathbf{y}\right)$ in this equality. As a result, we obtain $\omega\left(s_\lambda\left(\mathbf{x}\right)\right) = s_{\lambda^t}\left(\mathbf{x}\right)$ for every $\lambda\in\operatorname{Par}$. Thus, (13.74.7) is proven again.]

Third solution to Exercise 2.7.12(a): In order to solve Exercise 2.7.12(a), it suffices to verify the identity (13.74.11). (This can be proven by the same argument as in the Second solution to Exercise 2.7.12(a).)

We shall now verify the identity (13.74.11) using the so-called *dual RSK algorithm* (also known as *RSK* algorithm*)[627]. This will be fairly similar to the Second solution to Exercise 2.7.12(a) given above, but not identical to it; in particular, the dual RSK algorithm that we will introduce below will (unlike the coRSK algorithm from the Second solution) not rely on the same row bumping operation as the usual RSK algorithm, but on a somewhat modified version of it.

For a given partition λ, let us define a *row-strict tableau* of shape λ to be an assignment T of entries in $\{1,2,3,\ldots\}$ to the cells of the Ferrers diagram for λ which is strictly increasing left-to-right in rows, and weakly increasing top-to-bottom in columns. The equality (13.74.12) holds[628]; thus, we have

$$\begin{aligned} \sum_{\substack{P \text{ is a row-strict} \\ \text{tableau of shape } \lambda}} \mathbf{x}^{\operatorname{cont}(P)} &= \sum_{\substack{Q \text{ is a row-strict} \\ \text{tableau of shape } \lambda}} \mathbf{x}^{\operatorname{cont}(Q)} = s_{\lambda^t} \qquad \text{(by (13.74.12))} \\ &= s_{\lambda^t}\left(\mathbf{x}\right). \end{aligned} \tag{13.74.16}$$

We also have

$$\sum_{\substack{Q \text{ is a column-strict} \\ \text{tableau of shape } \lambda}} \mathbf{x}^{\operatorname{cont}(Q)} = \sum_{\substack{P \text{ is a column-strict} \\ \text{tableau of shape } \lambda}} \mathbf{x}^{\operatorname{cont}(P)} = s_\lambda\left(\mathbf{x}\right)$$

(because this is how $s_\lambda = s_\lambda\left(\mathbf{x}\right)$ is defined) for every partition λ. Substituting $\mathbf{y}$ for $\mathbf{x}$ in this equality, we obtain

$$\sum_{\substack{Q \text{ is a column-strict} \\ \text{tableau of shape } \lambda}} \mathbf{y}^{\operatorname{cont}(Q)} = s_\lambda\left(\mathbf{y}\right). \tag{13.74.17}$$

etc., moving further and further up. Thus, for each $q\in\{0,1,\ldots,r\}$, the tableau P_{m+q-1} is obtained from P_{m+q} by reverse bumping the entry in position c_{m+q}. Hence, conversely, the tableau P_{m+q} is obtained from P_{m+q-1} by RS-inserting the entry j_{m+q}, which creates the corner cell c_{m+q}.

But recall that $j_{m+p}\leq j_{m+p+1}$. Hence, part (a) of the row bumping lemma (applied to P_{m+p-1}, j_{m+p}, j_{m+p+1}, P_{m+p}, c_{m+p}, P_{m+p+1} and c_{m+p+1} instead of P, j, j', P', c, P'' and c') shows that the cell c_{m+p+1} is in the same row as the cell c_{m+p} or in a row further up. But this contradicts the fact that the cell c_{m+p+1} is in a row further down than the cell c_{m+p} (since we have labeled our cells as $c_m, c_{m+1},\ldots,c_{m+r}$ from top to bottom, and no two of them lied in the same row). This contradiction completes our proof.

[627] We will define this algorithm further below. It also frequently appears in literature: see, e.g., [206, §7.14], [111, §5], [186, Theorem 4.8.5] and (with different conventions) [73, §A.4.3, Prop. 3].

[628] Indeed, it can be proven as in the Second solution to Exercise 2.7.12(a).

A *cotableau-tableau pair* will mean a pair (P, Q) such that P is a row-strict tableau, Q is a column-strict tableau, and P and Q both have shape λ for one and the same partition λ. Multiplying the identities (13.74.16) and (13.74.17) and multiplying the result with $t^{|\lambda|}$, we obtain

$$t^{|\lambda|} \sum_{\substack{(P,Q) \text{ is a pair with} \\ P \text{ being a row-strict tableau} \\ \text{of shape } \lambda, \text{ and } Q \text{ being} \\ \text{a column-strict tableau of shape } \lambda}} \mathbf{x}^{\operatorname{cont}(P)} \mathbf{y}^{\operatorname{cont}(Q)} = t^{|\lambda|} s_{\lambda^t} (\mathbf{x}) \, s_\lambda (\mathbf{y})$$

for every partition λ. Summing this equality over all partitions λ, we obtain

$$\sum t^{|\lambda|} \mathbf{x}^{\operatorname{cont}(P)} \mathbf{y}^{\operatorname{cont}(Q)} = \sum_{\lambda \in \operatorname{Par}} t^{|\lambda|} s_{\lambda^t} (\mathbf{x}) \, s_\lambda (\mathbf{y}) ,$$

where the sum on the left hand side is over all cotableau-tableau pairs (P, Q) and where λ denotes the common shape of P and Q. This becomes

$$\begin{aligned} \sum t^{|\lambda|} \mathbf{x}^{\operatorname{cont}(P)} \mathbf{y}^{\operatorname{cont}(Q)} &= \sum_{\lambda \in \operatorname{Par}} t^{|\lambda|} s_{\lambda^t} (\mathbf{x}) \, s_\lambda (\mathbf{y}) = \sum_{\lambda \in \operatorname{Par}} \underbrace{t^{\left|\lambda^t\right|}}_{\substack{= t^{|\lambda|} \\ (\text{since } \left|\lambda^t\right| = |\lambda|)}} \underbrace{s_{\left(\lambda^t\right)^t} (\mathbf{x})}_{\substack{= s_\lambda(\mathbf{x}) \\ (\text{since } \left(\lambda^t\right)^t = \lambda)}} s_{\lambda^t} (\mathbf{y}) \\ &\qquad \left(\begin{array}{c} \text{here, we substituted } \lambda^t \text{ for } \lambda \text{ in the sum, since the map} \\ \operatorname{Par} \to \operatorname{Par}, \ \lambda \mapsto \lambda^t \text{ is a bijection} \end{array} \right) \\ &= \sum_{\lambda \in \operatorname{Par}} t^{|\lambda|} s_\lambda (\mathbf{x}) \, s_{\lambda^t} (\mathbf{y}) . \end{aligned} \tag{13.74.18}$$

We shall use all notations introduced in the proof of Theorem 2.5.1. A *strict biword* will mean an array $\binom{\mathbf{i}}{\mathbf{j}} = \binom{i_1 \ldots i_\ell}{j_1 \ldots j_\ell}$ in which the biletters satisfy $\binom{i_1}{j_1} <_{lex} \cdots <_{lex} \binom{i_\ell}{j_\ell}$ (that is, the biletters are distinct and ordered with respect to the lexicographic order). Strict biwords are clearly in 1-to-1 correspondence with sets (not multisets!) of biletters. Now, the left hand side of (13.74.11) is $\prod_{i,j=1}^{\infty} (1 + t x_i y_j) = \prod_{i,j=1}^{\infty} (1 + t x_j y_i)$ (here, we substituted (j, i) for the index (i, j) in the product), and thus can be rewritten as the sum of $t^\ell (x_{j_1} y_{i_1}) (x_{j_2} y_{i_2}) \cdots (x_{j_\ell} y_{i_\ell})$ over all sets $\left\{ \binom{i_1}{j_1}, \ldots, \binom{i_\ell}{j_\ell} \right\}$ of biletters. Thus, the left hand side of (13.74.11) is the sum $\sum t^\ell \mathbf{x}^{\operatorname{cont}(\mathbf{j})} \mathbf{y}^{\operatorname{cont}(\mathbf{i})}$ over all strict biwords $\binom{\mathbf{i}}{\mathbf{j}}$, where ℓ stands for the number of biletters in the biword. Meanwhile, we know that the right hand side of (13.74.11) is the sum $\sum t^{|\lambda|} \mathbf{x}^{\operatorname{cont}(P)} \mathbf{y}^{\operatorname{cont}(Q)}$ over all cotableau-tableau pairs (P, Q) (because of (13.74.18)). Thus, in order to prove (13.74.11), we only need to construct a bijection between the strict biwords $\binom{\mathbf{i}}{\mathbf{j}}$ and the cotableau-tableau pairs (P, Q), which has the property that

$$\begin{aligned} \operatorname{cont}(\mathbf{i}) &= \operatorname{cont}(Q) ; \\ \operatorname{cont}(\mathbf{j}) &= \operatorname{cont}(P) . \end{aligned}$$

[629]

This bijection is the *dual RSK algorithm*, which we shall define below.

First, we shall introduce a simpler operation which we call *dual RS-insertion* (and which is similar to RS-insertion, but not identical with it[630]).

Dual RS-insertion takes as input a row-strict tableau P and a letter j, and returns a new row-strict tableau P' along with a corner cell c of P', which is constructed as follows: Start out by setting $P' = P$. The letter j tries to insert itself into the first row of P' by either bumping out the leftmost letter in the first row **larger or equal to** j, or else placing itself at the right end of the row if no such letter (larger or equal to j) exists. If a letter was bumped from the first row, this letter follows the same rules to insert itself into the second row, and so on. This series of bumps must eventually come to an end. At the end of the bumping, the tableau P' created has an extra corner cell not present in P. If we call this corner cell c, then P' (in its final form) and c are what the dual RS-insertion operation returns. One says that P' is the

[629] Such a bijection will then automatically satisfy $|\lambda| = \ell$, where λ is the (common) shape of P and Q, and where ℓ is the length of the strict biword $\binom{\mathbf{i}}{\mathbf{j}}$. This is because $|\lambda| = |\operatorname{cont}(Q)|$ and $\ell = |\operatorname{cont}(\mathbf{i})|$.

[630] We will leave many of its properties unproven, because their proofs are analogous (or at least very similar) to the proofs of the corresponding properties of RS-insertion, and thus can be easily reconstructed by the reader.

result of *dually inserting* j into the tableau P. It is straightforward to see that this resulting filling P' is a row-strict tableau[631].

Example 13.74.2. To give an example of this operation, let us dually insert the letter $j = 3$ into the row-strict tableau $\begin{array}{llll} 1 & 2 & 3 & 5 \\ 1 & 2 & 4 & \\ 3 & 4 & 7 & \\ 3 & 6 & & \\ 4 & & & \end{array}$ (we are showing all intermediate states of P'; the underlined letter is always the one that is going to be bumped out at the next step):

$$\begin{array}{llll} 1 & 2 & \underline{3} & 5 \\ 1 & 2 & 4 & \\ 3 & 4 & 7 & \\ 3 & 6 & & \\ 4 & & & \end{array} \xmapsto[\substack{\text{insert 3;}\\ \text{bump out 3}}]{} \begin{array}{llll} 1 & 2 & 3 & 5 \\ 1 & 2 & \underline{4} & \\ 3 & 4 & 7 & \\ 3 & 6 & & \\ 4 & & & \end{array} \xmapsto[\substack{\text{insert 3;}\\ \text{bump out 4}}]{} \begin{array}{llll} 1 & 2 & 3 & 5 \\ 1 & 2 & 3 & \\ 3 & \underline{4} & 7 & \\ 3 & 6 & & \\ 4 & & & \end{array}$$

$$\xmapsto[\substack{\text{insert 4;}\\ \text{bump out 4}}]{} \begin{array}{llll} 1 & 2 & 3 & 5 \\ 1 & 2 & 3 & \\ 3 & 4 & 7 & \\ 3 & \underline{6} & & \\ 4 & & & \end{array} \xmapsto[\substack{\text{insert 4;}\\ \text{bump out 6}}]{} \begin{array}{llll} 1 & 2 & 3 & 5 \\ 1 & 2 & 3 & \\ 3 & 4 & 7 & \\ 3 & 4 & & \\ 4 & & & \end{array} \xmapsto[\substack{\text{insert 6;}\\ \text{done}}]{} \begin{array}{llll} 1 & 2 & 3 & 5 \\ 1 & 2 & 3 & \\ 3 & 4 & 7 & \\ 3 & 4 & & \\ 4 & 6 & & \end{array} .$$

The last tableau in this sequence is the row-strict tableau that is returned. The corner cell that is returned is the second cell of the fifth row (the one containing 6).

Dual RS-insertion will be used as a step in the dual RSK algorithm; the construction will rely on a simple fact known as the *dual row bumping lemma*. Let us first define the notion of a *dual bumping path* (or *dual bumping route*): If P is a row-strict tableau, and j is a letter, then some letters are inserted into some cells when dual RS-insertion is applied to P and j. The sequence of these cells (in the order in which they see letters inserted into them)[632] is called the *dual bumping path* for P and j. This dual bumping path always ends with the corner cell c which is returned by dual RS-insertion. As an example, when $j = 3$ is dually inserted into the tableau P shown below, the result P' is shown with all entries on the dual bumping path underlined:

$$P = \begin{array}{llll} 1 & 2 & 3 & 4 \\ 1 & 2 & 4 & 6 \\ 2 & 4 & 5 & 6 \\ 2 & 4 & & \\ 3 & & & \end{array} \quad \xmapsto[j=3]{\text{dually insert}} \quad P' = \begin{array}{llll} 1 & 2 & \underline{3} & 4 \\ 1 & 2 & \underline{3} & 6 \\ 2 & \underline{4} & 5 & 6 \\ 2 & \underline{4} & & \\ 3 & \underline{4} & & \end{array}$$

A first simple observation about dual bumping paths is that dual bumping paths trend weakly left (just as bumping paths for regular RS-insertion do). A subtler property of bumping paths is the following *dual row bumping lemma*[633]:

> **Dual row bumping lemma:** Let P be a row-strict tableau, and let j and j' be two letters. Applying dual RS-insertion to the tableau P and the letter j yields a new row-strict tableau P' and a corner cell c. Applying dual RS-insertion to the tableau P' and the letter j' yields a new row-strict tableau P'' and a corner cell c'.
>
> (a) Assume that $j < j'$. Then, the dual bumping path for P' and j' stays strictly to the right, within each row, of the dual bumping path for P and j. The cell c' (in which the dual bumping path for P' and j' ends) is in the same row as the cell c (in which the dual bumping path for P and j ends) or in a row further up; it is also in a column further right than c.

[631] Indeed, the reader can check that P' remains a row-strict tableau throughout the algorithm that defines dual RS-insertion.

[632] In particular, this includes those cells whose entries did not change under the insertion (because the entry inserted was the same as the entry they contained before the insertion).

[633] This lemma is equivalent to the "column bumping lemma" in Fulton [73, p. 187].

(b) Assume instead that $j \geq j'$. Then, the dual bumping path for P' and j' stays weakly to the left, within each row, of the dual bumping path for P and j. The cell c' (in which the dual bumping path for P' and j' ends) is in a row further down than the cell c (in which the dual bumping path for P and j ends); it is also in the same column as c or in a column further left.

This lemma can be easily proven by induction over the row (similarly to the usual row bumping lemma).

We will now define the dual RSK algorithm. Let $\binom{\mathbf{i}}{\mathbf{j}}$ be a strict biword. Starting with the pair $(P_0, Q_0) = (\varnothing, \varnothing)$ and $m = 0$, the algorithm applies the following steps (see Example 13.74.3 below):

- If i_{m+1} does not exist (that is, m is the length of $\mathbf{i}$), stop.
- Apply dual RS-insertion to the row-strict tableau P_m and the letter j_{m+1} (the bottom letter of $\binom{i_{m+1}}{j_{m+1}}$). Let P_{m+1} be the resulting row-strict tableau, and let c_{m+1} be the resulting corner cell.
- Create Q_{m+1} from Q_m by adding the top letter i_{m+1} of $\binom{i_{m+1}}{j_{m+1}}$ to Q_m in the cell c_{m+1} (which, as we recall, is the extra corner cell of P_{m+1} not present in P_m).
- Set m to $m + 1$.

After all of the biletters have been thus processed, the result of the dual RSK algorithm is $(P_\ell, Q_\ell) =: (P, Q)$.

Example 13.74.3. The term in the expansion of the left side of (13.74.10) corresponding to

$$(x_2y_1)(x_4y_1)(x_1y_2)(x_1y_4)(x_3y_4)(x_2y_5)$$

is the strict biword $\binom{\mathbf{i}}{\mathbf{j}} = \binom{112445}{241132}$, and the dual RSK algorithm applied to this biword proceeds as follows:

$$\begin{array}{rcl|rcl}
P_0 & = & \varnothing & Q_0 & = & \varnothing \\[1em]
P_1 & = & 2 & Q_1 & = & 1 \\[1em]
P_2 & = & 2 \quad 4 & Q_2 & = & 1 \quad 1 \\[1em]
P_3 & = & \begin{array}{ll} 1 & 4 \\ 2 & \end{array} & Q_3 & = & \begin{array}{ll} 1 & 1 \\ 2 & \end{array} \\[1em]
P_4 & = & \begin{array}{ll} 1 & 4 \\ 1 & \\ 2 & \end{array} & Q_4 & = & \begin{array}{ll} 1 & 1 \\ 2 & \\ 4 & \end{array} \\[1em]
P_5 & = & \begin{array}{ll} 1 & 3 \\ 1 & 4 \\ 2 & \end{array} & Q_5 & = & \begin{array}{ll} 1 & 1 \\ 2 & 4 \\ 4 & \end{array} \\[1em]
P := P_6 & = & \begin{array}{ll} 1 & 2 \\ 1 & 3 \\ 2 & 4 \end{array} & Q := Q_6 & = & \begin{array}{ll} 1 & 1 \\ 2 & 4 \\ 4 & 5 \end{array}
\end{array}$$

It is clear that P_m remains a row-strict tableau of some Ferrers shape throughout the execution of the dual RSK algorithm, and that Q_m remains a filling of the same shape as P_m which is (at least) weakly increasing left-to-right along rows and weakly increasing top-to-bottom in columns. But we can also see that Q_m is strictly increasing top-to-bottom along columns[634], so that Q_m is a column-strict tableau. Thus, the result (P, Q) of the dual RSK algorithm is a cotableau-tableau pair.

[634] Indeed, this follows from the observation that when one has a string of equal letters $i_m = i_{m+1} = \cdots = i_{m+r}$ on top of the strict biword, then the bottom letters bumped in are $j_m < j_{m+1} < \cdots < j_{m+r}$, and therefore (as a consequence of the second-to-last claim of part (a) of the dual row bumping lemma) the new cells form a *horizontal strip*, that is, no two of these cells lie in the same column. Actually, more can be said: Each of these new cells (except for the first one) is in a column further right than the previous one. We will use this stronger fact further below.

To see that the dual RSK algorithm is a bijection, we show how to recover $\binom{\mathbf{i}}{\mathbf{j}}$ from (P, Q). This is done by *dually reverse bumping* from (P_{m+1}, Q_{m+1}) to recover both the biletter $\binom{i_{m+1}}{j_{m+1}}$ and the tableaux (P_m, Q_m), as follows. Firstly, i_{m+1} is the maximum entry of Q_{m+1}, and Q_m is obtained by removing the rightmost occurrence of this letter i_{m+1} from Q_{m+1}. [635] To produce P_m and j_{m+1}, find the position of the rightmost occurrence of i_{m+1} in Q_{m+1}, and start *dually reverse bumping* in P_{m+1} from the entry in this same position, where dually reverse bumping an entry means inserting it into one row higher by having it bump out the rightmost entry which is smaller or equal to it.[636] The entry bumped out of the first row is j_{m+1}, and the resulting tableau is P_m.

Finally, to see that the dual RSK map is surjective, one needs to show that the dually reverse bumping procedure can be applied to any cotableau-tableau pair (P, Q), and will result in a strict biword $\binom{\mathbf{i}}{\mathbf{j}}$. We leave this verification to the reader.[637]

So the dual RSK map is a bijection having the required properties. As we have said, this proves (13.74.11). This solves Exercise 2.7.12(a) again.

(b) Let us consider the map $\omega\left[\left[\mathbf{y}\right]\right] : \Lambda\left[\left[\mathbf{y}\right]\right] \to \Lambda\left[\left[\mathbf{y}\right]\right]$ defined as in the solution of Exercise 2.7.12(a).

Now, the equality (2.5.1) yields

$$\sum_{\lambda\in\operatorname{Par}} s_\lambda\left(\mathbf{x}\right) s_\lambda\left(\mathbf{y}\right) = \prod_{i,j=1}^{\infty} \left(1 - x_i y_j\right)^{-1} = \sum_{\lambda\in\operatorname{Par}} z_\lambda^{-1} p_\lambda\left(\mathbf{x}\right) p_\lambda\left(\mathbf{y}\right) \qquad \left(\text{by } (2.5.11)\right). \tag{13.74.19}$$

This is an equality in $\Lambda\left[\left[\mathbf{y}\right]\right]$ (because $s_\lambda\left(\mathbf{x}\right)$ and $p_\lambda\left(\mathbf{x}\right)$ belong to Λ for every $\lambda \in \operatorname{Par}$). Hence, we can apply the map $\omega\left[\left[\mathbf{y}\right]\right] : \Lambda\left[\left[\mathbf{y}\right]\right] \to \Lambda\left[\left[\mathbf{y}\right]\right]$ to both sides of this equality. As a result, we obtain

$$\sum_{\lambda\in\operatorname{Par}} \omega\left(s_\lambda\left(\mathbf{x}\right)\right) s_\lambda\left(\mathbf{y}\right) = \sum_{\lambda\in\operatorname{Par}} z_\lambda^{-1} \omega\left(p_\lambda\left(\mathbf{x}\right)\right) p_\lambda\left(\mathbf{y}\right)$$

(because $\omega\left[\left[\mathbf{y}\right]\right]$, being a $\mathbf{k}\left[\left[\mathbf{y}\right]\right]$-algebra homomorphism, leaves the $s_\lambda\left(\mathbf{y}\right)$ and $p_\lambda\left(\mathbf{y}\right)$ terms unchanged, while the $s_\lambda\left(\mathbf{x}\right)$ and $p_\lambda\left(\mathbf{x}\right)$ terms are elements of Λ and thus are transformed as by ω). Due to (2.4.14) and

[635] It necessarily has to be the rightmost occurrence, since (according to the previous footnote) the cell into which i_{m+1} was filled at the step from Q_m to Q_{m+1} lies further right than any existing cell of Q_m containing the letter i_{m+1}.

[636] Let us give a few more details on this "dually reverse bumping" procedure. Dually reverse bumping (also known as *dual RS-deletion* or *reverse dual RS-insertion*) is an operation which takes a row-strict tableau P' and a corner cell c of P', and constructs a row-strict tableau P and a letter j such that dual RS-insertion for P and j yields P' and c. It starts by setting $P = P'$, and removing the entry in the cell c from P. This removed entry is then denoted by k, and is inserted into the row of P above c, bumping out the rightmost entry which is smaller or equal to k. The letter which is bumped out – say, ℓ –, in turn, is inserted into the row above it, bumping out the rightmost entry which is smaller or equal to ℓ. This procedure continues in the same way until an entry is bumped out of the first row (which will eventually happen). The dually reverse bumping operation returns the resulting tableau P and the entry which is bumped out of the first row.

It is straightforward to check that the dually reverse bumping operation is well-defined (i.e., P does stay a row-strict tableau throughout the procedure) and is the inverse of the dual RS-insertion operation. (In fact, these two operations undo each other step by step.)

[637] It is easy to see that repeatedly applying dually reverse bumping to (P, Q) will result in a sequence $\binom{i_\ell}{j_\ell}, \binom{i_{\ell-1}}{j_{\ell-1}}, \ldots, \binom{i_1}{j_1}$ of biletters such that applying the dual RSK algorithm to $\binom{i_1 \cdots i_\ell}{j_1 \cdots j_\ell}$ gives back (P, Q). The question is why we have $\binom{i_1}{j_1} <_{lex} \cdots <_{lex} \binom{i_\ell}{j_\ell}$. Since the chain of inequalities $i_1 \leq i_2 \leq \cdots \leq i_\ell$ is clear from the choice of entry to dually reverse-bump, it only remains to show that for every string $i_m = i_{m+1} = \cdots = i_{m+r}$ of equal top letters, the corresponding bottom letters strictly increase (that is, $j_m < j_{m+1} < \cdots < j_{m+r}$). One way to see this is the following:

Assume the contrary; i.e., assume that the bottom letters corresponding to some string $i_m = i_{m+1} = \cdots = i_{m+r}$ of equal top letters do not strictly increase. Thus, $j_{m+p} \geq j_{m+p+1}$ for some $p \in \{0, 1, \ldots, r-1\}$. Consider this p.

Let us consider the cells containing the equal letters $i_m = i_{m+1} = \cdots = i_{m+r}$ in the tableau Q_{m+r}. Label these cells as $c_m, c_{m+1}, \ldots, c_{m+r}$ from left to right (noticing that no two of them lie in the same column, since Q_{m+r} is column-strict). By the definition of dually reverse bumping, the first entry to be dually reverse bumped from P_{m+r} is the entry in position c_{m+r} (since this is the rightmost occurrence of the letter i_{m+r} in Q_{m+r}); then, the next entry to be dually reverse bumped is the one in position c_{m+r-1}, etc., moving further and further left. Thus, for each $q \in \{0, 1, \ldots, r\}$, the tableau P_{m+q-1} is obtained from P_{m+q} by dually reverse bumping the entry in position c_{m+q}. Hence, conversely, the tableau P_{m+q} is obtained from P_{m+q-1} by dually RS-inserting the entry j_{m+q}, which creates the corner cell c_{m+q}.

But recall that $j_{m+p} \geq j_{m+p+1}$. Hence, part (b) of the dual row bumping lemma (applied to P_{m+p-1}, j_{m+p}, j_{m+p+1}, P_{m+p}, c_{m+p}, P_{m+p+1} and c_{m+p+1} instead of P, j, j', P', c, P'' and c') shows that the cell c_{m+p+1} is in the same column as the cell c_{m+p} or in a column further left. But this contradicts the fact that the cell c_{m+p+1} is in a column further right than the cell c_{m+p} (since we have labeled our cells as $c_m, c_{m+1}, \ldots, c_{m+r}$ from left to right, and no two of them lied in the same column). This contradiction completes our proof.

(13.74.7), this rewrites as

$$\sum_{\lambda\in\operatorname{Par}} s_{\lambda^t}\left(\mathbf{x}\right) s_\lambda\left(\mathbf{y}\right) = \sum_{\lambda\in\operatorname{Par}} z_\lambda^{-1}\left(-1\right)^{\left|\lambda\right|-\ell\left(\lambda\right)} p_\lambda\left(\mathbf{x}\right) p_\lambda\left(\mathbf{y}\right).$$

Since $\sum_{\lambda\in\operatorname{Par}} s_{\lambda^t}\left(\mathbf{x}\right) s_\lambda\left(\mathbf{y}\right) = \sum_{\lambda\in\operatorname{Par}} s_\lambda\left(\mathbf{x}\right) s_{\lambda^t}\left(\mathbf{y}\right)$ (this can be proven just as in (13.74.8)), this rewrites as

$$\sum_{\lambda\in\operatorname{Par}} s_\lambda\left(\mathbf{x}\right) s_{\lambda^t}\left(\mathbf{y}\right) = \sum_{\lambda\in\operatorname{Par}} z_\lambda^{-1}\left(-1\right)^{\left|\lambda\right|-\ell\left(\lambda\right)} p_\lambda\left(\mathbf{x}\right) p_\lambda\left(\mathbf{y}\right) = \sum_{\lambda\in\operatorname{Par}} \left(-1\right)^{\left|\lambda\right|-\ell\left(\lambda\right)} z_\lambda^{-1} p_\lambda\left(\mathbf{x}\right) p_\lambda\left(\mathbf{y}\right).$$

Hence,

$$\sum_{\lambda\in\operatorname{Par}} \left(-1\right)^{\left|\lambda\right|-\ell\left(\lambda\right)} z_\lambda^{-1} p_\lambda\left(\mathbf{x}\right) p_\lambda\left(\mathbf{y}\right) = \sum_{\lambda\in\operatorname{Par}} s_\lambda\left(\mathbf{x}\right) s_{\lambda^t}\left(\mathbf{y}\right) = \prod_{i,j=1}^{\infty}\left(1+x_iy_j\right)$$

(by (13.74.9)). This solves Exercise 2.7.12(b).

13.75. **Solution to Exercise 2.7.13.** *Solution to Exercise 2.7.13. Proof of Theorem 2.4.6.*

Let $n\in\mathbb{N}$. Let μ be a partition having at most n parts. Exercise 2.5.11(a) yields

$$\sum_{\lambda\in\operatorname{Par}} s_\lambda\left(\mathbf{x}\right) s_{\lambda/\mu}\left(\mathbf{y}\right) = s_\mu\left(\mathbf{x}\right)\cdot\prod_{i,j=1}^{\infty}\left(1-x_iy_j\right)^{-1}$$

in the ring $\mathbf{k}\left[\left[\mathbf{x},\mathbf{y}\right]\right] = \mathbf{k}\left[\left[x_1,x_2,x_3,...,y_1,y_2,y_3,...\right]\right]$. Switching the roles of the variables $\mathbf{x}$ and $\mathbf{y}$ in this equality, we obtain

$$\sum_{\lambda\in\operatorname{Par}} s_\lambda\left(\mathbf{y}\right) s_{\lambda/\mu}\left(\mathbf{x}\right) = s_\mu\left(\mathbf{y}\right)\cdot\prod_{i,j=1}^{\infty}\left(1-y_ix_j\right)^{-1}.$$

We can now substitute $\left(y_1,y_2,...,y_n,0,0,0,...\right)$ for $\left(y_1,y_2,y_3,...\right)$ on both sides of this, and obtain the equality

$$\sum_{\lambda\in\operatorname{Par}} s_\lambda\left(y_1,y_2,...,y_n\right) s_{\lambda/\mu}\left(\mathbf{x}\right) = s_\mu\left(y_1,y_2,...,y_n\right)\cdot\prod_{i=1}^{n}\prod_{j=1}^{\infty}\left(1-y_ix_j\right)^{-1}$$

in the subring $\left(\mathbf{k}\left[\left[\mathbf{x}\right]\right]\right)\left[\left[y_1,y_2,...,y_n\right]\right]$ of $\mathbf{k}\left[\left[\mathbf{x},\mathbf{y}\right]\right]$ (notice that the $\prod_{i,j=1}^{\infty}\left(1-y_ix_j\right)^{-1}$ on the right hand side became $\prod_{i=1}^{n}\prod_{j=1}^{\infty}\left(1-y_ix_j\right)^{-1}$ because all factors $1-y_ix_j$ with $i>n$ got sent to $1-0x_j=1$). In the sum on the left hand side of this equality, all addends corresponding to partitions λ having more than n parts are 0 (because Exercise 2.3.8(b) yields that all such λ satisfy $s_\lambda\left(x_1,x_2,\ldots,x_n\right)=0$, hence $s_\lambda\left(y_1,y_2,...,y_n\right)=0$). Thus, we can remove all these addends, and the equality thus becomes

$$\sum_{\substack{\lambda\in\operatorname{Par};\\ \lambda\text{ has at most }n\text{ parts}}} s_\lambda\left(y_1,y_2,...,y_n\right) s_{\lambda/\mu}\left(\mathbf{x}\right) = s_\mu\left(y_1,y_2,...,y_n\right)\cdot\prod_{i=1}^{n}\prod_{j=1}^{\infty}\left(1-y_ix_j\right)^{-1}. \tag{13.75.1}$$

Let ρ be the n-tuple $\left(n-1,n-2,...,2,1,0\right)\in\mathbb{N}^n$. We can regard ρ as a weak composition by padding it with zeroes at the end (i.e., identifying ρ with the weak composition $\left(n-1,n-2,...,2,1,0,0,0,0,...\right)$).

For every n-tuple $\alpha=\left(\alpha_1,\alpha_2,\ldots,\alpha_n\right)\in\mathbb{N}^n$, we define the alternant $a_\alpha\in\mathbf{k}\left[x_1,x_2,\ldots,x_n\right]$ as in Definition 2.6.2. (But other than this, we are not adapting the notations of Section 2.6.)

Now, Corollary 2.6.7 states that $s_\lambda\left(x_1,x_2,...,x_n\right) = \dfrac{a_{\lambda+\rho}}{a_\rho}$ in $\mathbf{k}\left[x_1,x_2,...,x_n\right]$ whenever λ is a partition having at most n parts. Applied to the variables $y_1,y_2,...,y_n$ instead of $x_1,x_2,...,x_n$, this yields that $s_\lambda\left(y_1,y_2,...,y_n\right) = \dfrac{a_{\lambda+\rho}\left(y_1,y_2,...,y_n\right)}{a_\rho\left(y_1,y_2,...,y_n\right)}$ whenever λ is a partition having at most n parts. This equality, applied to $\lambda=\mu$, yields $s_\mu\left(y_1,y_2,...,y_n\right) = \dfrac{a_{\mu+\rho}\left(y_1,y_2,...,y_n\right)}{a_\rho\left(y_1,y_2,...,y_n\right)}$. Substituting the last two equalities into (13.75.1), we obtain

$$\sum_{\substack{\lambda\in\operatorname{Par};\\ \lambda\text{ has at most }n\text{ parts}}} \frac{a_{\lambda+\rho}\left(y_1,y_2,...,y_n\right)}{a_\rho\left(y_1,y_2,...,y_n\right)} s_{\lambda/\mu}\left(\mathbf{x}\right) = \frac{a_{\mu+\rho}\left(y_1,y_2,...,y_n\right)}{a_\rho\left(y_1,y_2,...,y_n\right)}\cdot\prod_{i=1}^{n}\prod_{j=1}^{\infty}\left(1-y_ix_j\right)^{-1}.$$

Multiplied with $a_\rho(y_1, y_2, ..., y_n)$, this becomes

$$\sum_{\substack{\lambda \in \operatorname{Par}; \\ \lambda \text{ has at most } n \text{ parts}}} a_{\lambda+\rho}(y_1, y_2, ..., y_n) s_{\lambda/\mu}(\mathbf{x}) = a_{\mu+\rho}(y_1, y_2, ..., y_n) \cdot \prod_{i=1}^{n} \prod_{j=1}^{\infty} (1 - y_i x_j)^{-1}.$$

Renaming λ as ν on the left hand side of this equality, we obtain

$$\sum_{\substack{\nu \in \operatorname{Par}; \\ \nu \text{ has at most } n \text{ parts}}} a_{\nu+\rho}(y_1, y_2, ..., y_n) s_{\nu/\mu}(\mathbf{x}) = a_{\mu+\rho}(y_1, y_2, ..., y_n) \cdot \prod_{i=1}^{n} \prod_{j=1}^{\infty} (1 - y_i x_j)^{-1}. \tag{13.75.2}$$

Now, let λ be a partition having at most n parts. We want to find the coefficient of $y_1^{\lambda_1+n-1} y_2^{\lambda_2+n-2} ... y_n^{\lambda_n+n-n}$ on the left and the right hand sides of (13.75.2). Here, we regard (13.75.2) as an equality in the ring $(\mathbf{k}[[\mathbf{x}]])[[y_1, y_2, ..., y_n]]$, so that we consider the variables $x_1, x_2, x_3, ...$ as constants, and thus (for example) the coefficient of y_1 in $(1 + x_1)(1 + y_1)$ is $1 + x_1$ rather than 1.

We first notice that

$$\prod_{i=1}^{n} \prod_{j=1}^{\infty} (1 - y_i x_j)^{-1} = \sum_{(q_1, q_2, ..., q_n) \in \mathbb{N}^n} \left(\prod_{j=1}^{n} h_{q_j}(\mathbf{x}) \right) \cdot \left(\prod_{j=1}^{n} y_j^{q_j} \right) \tag{13.75.3}$$

[638].

Let us recall a basic fact from linear algebra, namely the explicit formula for the determinant of a matrix as a sum over permutations: Any matrix $(\alpha_{i,j})_{i,j=1,2,...,\ell}$ over a commutative ring satisfies

$$\det\left((\alpha_{i,j})_{i,j=1,2,...,\ell} \right) = \sum_{\sigma \in \mathfrak{S}_\ell} (-1)^\sigma \prod_{i=1}^{\ell} \alpha_{i,\sigma(i)}. \tag{13.75.5}$$

Applying this to $\ell = n$ and $\alpha_{i,j} = y_i^{(\mu+\rho)_j}$, we obtain

$$\det\left(\left(y_i^{(\mu+\rho)_j} \right)_{i,j=1,2,...,n} \right) = \sum_{\sigma \in \mathfrak{S}_n} (-1)^\sigma \prod_{i=1}^{n} y_i^{(\mu+\rho)_{\sigma(i)}} = \sum_{\sigma \in \mathfrak{S}_n} (-1)^\sigma \prod_{j=1}^{n} y_j^{(\mu+\rho)_{\sigma(j)}} \tag{13.75.6}$$

[638] *Proof:* The equality (2.2.18) (with the indices i and n renamed as j and q) yields

$$\prod_{j=1}^{\infty} (1 - x_j t)^{-1} = 1 + h_1(\mathbf{x}) t + h_2(\mathbf{x}) t^2 + ... = \sum_{q \geq 0} h_q(\mathbf{x}) t^q \qquad \text{in } (\mathbf{k}[[\mathbf{x}]])[[t]]. \tag{13.75.4}$$

For every $i \in \{1, 2, ..., n\}$, we have

$$\prod_{j=1}^{\infty} (1 - y_i x_j)^{-1} = \sum_{q \geq 0} h_q(\mathbf{x}) y_i^q$$

(this follows by substituting y_i for t on both sides of (13.75.4)). Thus,

$$\begin{aligned} \prod_{i=1}^{n} \underbrace{\prod_{j=1}^{\infty} (1 - y_i x_j)^{-1}}_{= \sum_{q \geq 0} h_q(\mathbf{x}) y_i^q} &= \prod_{i=1}^{n} \sum_{q \geq 0} h_q(\mathbf{x}) y_i^q = \prod_{j=1}^{n} \sum_{q \geq 0} h_q(\mathbf{x}) y_j^q \\ &= \sum_{(q_1, q_2, ..., q_n) \in \mathbb{N}^n} \underbrace{\prod_{j=1}^{n} \left(h_{q_j}(\mathbf{x}) \cdot y_j^{q_j} \right)}_{= \left(\prod_{j=1}^{n} h_{q_j}(\mathbf{x}) \right) \cdot \left(\prod_{j=1}^{n} y_j^{q_j} \right)} \qquad \text{(by the product rule)} \\ &= \sum_{(q_1, q_2, ..., q_n) \in \mathbb{N}^n} \left(\prod_{j=1}^{n} h_{q_j}(\mathbf{x}) \right) \cdot \left(\prod_{j=1}^{n} y_j^{q_j} \right), \end{aligned}$$

qed.

(here, we renamed the index i as j in the product). Now, the right hand side of (13.75.2) becomes

$$\begin{aligned}
&\underbrace{a_{\mu+\rho}\left(y_1,y_2,\ldots,y_n\right)}_{\substack{=\det\left(\left(y_i^{(\mu+\rho)_j}\right)_{i,j=1,2,\ldots,n}\right) \\ \text{(by the definition of the alternant } a_{\mu+\rho}\text{)}}} \cdot \underbrace{\prod_{i=1}^{n}\prod_{j=1}^{\infty}\left(1-y_ix_j\right)^{-1}}_{\substack{=\sum_{(q_1,q_2,\ldots,q_n)\in\mathbb{N}^n}\left(\prod_{j=1}^{n}h_{q_j}(\mathbf{x})\right)\cdot\left(\prod_{j=1}^{n}y_j^{q_j}\right) \\ \text{(by (13.75.3))}}} \\
&= \underbrace{\det\left(\left(y_i^{(\mu+\rho)_j}\right)_{i,j=1,2,\ldots,n}\right)}_{\substack{=\sum_{\sigma\in\mathfrak{S}_n}(-1)^\sigma\prod_{j=1}^{n}y_j^{(\mu+\rho)_{\sigma(j)}} \\ \text{(by (13.75.6))}}} \cdot \sum_{(q_1,q_2,\ldots,q_n)\in\mathbb{N}^n}\left(\prod_{j=1}^{n}h_{q_j}(\mathbf{x})\right)\cdot\left(\prod_{j=1}^{n}y_j^{q_j}\right) \\
&= \left(\sum_{\sigma\in\mathfrak{S}_n}(-1)^\sigma\prod_{j=1}^{n}y_j^{(\mu+\rho)_{\sigma(j)}}\right) \cdot \sum_{(q_1,q_2,\ldots,q_n)\in\mathbb{N}^n}\left(\prod_{j=1}^{n}h_{q_j}(\mathbf{x})\right)\cdot\left(\prod_{j=1}^{n}y_j^{q_j}\right) \\
&= \sum_{\sigma\in\mathfrak{S}_n}(-1)^\sigma\sum_{(q_1,q_2,\ldots,q_n)\in\mathbb{N}^n}\left(\prod_{j=1}^{n}h_{q_j}(\mathbf{x})\right)\cdot\underbrace{\left(\prod_{j=1}^{n}y_j^{q_j}\right)\left(\prod_{j=1}^{n}y_j^{(\mu+\rho)_{\sigma(j)}}\right)}_{=\prod_{j=1}^{n}\left(y_j^{q_j}y_j^{(\mu+\rho)_{\sigma(j)}}\right)=\prod_{j=1}^{n}y_j^{q_j+(\mu+\rho)_{\sigma(j)}}} \\
&= \sum_{\sigma\in\mathfrak{S}_n}(-1)^\sigma\sum_{(q_1,q_2,\ldots,q_n)\in\mathbb{N}^n}\left(\prod_{j=1}^{n}h_{q_j}(\mathbf{x})\right)\cdot\left(\prod_{j=1}^{n}y_j^{q_j+(\mu+\rho)_{\sigma(j)}}\right).
\end{aligned}$$

Hence, the coefficient of $y_1^{\lambda_1+n-1}y_2^{\lambda_2+n-2}\ldots y_n^{\lambda_n+n-n}$ on the right hand side of (13.75.2) equals

$$\sum_{\sigma\in\mathfrak{S}_n}(-1)^\sigma\sum_{\substack{(q_1,q_2,\ldots,q_n)\in\mathbb{N}^n; \\ q_j+(\mu+\rho)_{\sigma(j)}=\lambda_j+n-j\text{ for every }j\in\{1,2,\ldots,n\}}}\prod_{j=1}^{n}h_{q_j}(\mathbf{x}). \tag{13.75.7}$$

However, for every $\sigma\in\mathfrak{S}_n$, it is easy to see that

$$\sum_{\substack{(q_1,q_2,\ldots,q_n)\in\mathbb{N}^n; \\ q_j+(\mu+\rho)_{\sigma(j)}=\lambda_j+n-j\text{ for every }j\in\{1,2,\ldots,n\}}}\prod_{j=1}^{n}h_{q_j}(\mathbf{x})=\prod_{j=1}^{n}h_{\lambda_j+n-j-(\mu+\rho)_{\sigma(j)}}(\mathbf{x}). \tag{13.75.8}$$

[639] Hence, the coefficient of $y_1^{\lambda_1+n-1} y_2^{\lambda_2+n-2} ... y_n^{\lambda_n+n-n}$ on the right hand side of (13.75.2) equals

$$\begin{aligned}
& \sum_{\sigma\in\mathfrak{S}_n} (-1)^\sigma \underbrace{\sum_{\substack{(q_1,q_2,...,q_n)\in\mathbb{N}^n;\\ q_j+(\mu+\rho)_{\sigma(j)}=\lambda_j+n-j \text{ for every } j\in\{1,2,...,n\}}} \prod_{j=1}^n h_{q_j}(\mathbf{x})}_{\substack{=\prod_{j=1}^n h_{\lambda_j+n-j-(\mu+\rho)_{\sigma(j)}}(\mathbf{x}) \\ \text{(by (13.75.8))}}} \qquad \text{(by (13.75.7))} \\
&= \sum_{\sigma\in\mathfrak{S}_n} (-1)^\sigma \prod_{j=1}^n \underbrace{h_{\lambda_j+n-j-(\mu+\rho)_{\sigma(j)}}}_{\substack{=h_{\lambda_j-\mu_{\sigma(j)}-j+\sigma(j)} \\ \text{(since it is easy to see that} \\ \lambda_j+n-j-(\mu+\rho)_{\sigma(j)}=\lambda_j-\mu_{\sigma(j)}-j+\sigma(j))}} (\mathbf{x}) \\
&= \sum_{\sigma\in\mathfrak{S}_n} (-1)^\sigma \prod_{j=1}^n h_{\lambda_j-\mu_{\sigma(j)}-j+\sigma(j)}(\mathbf{x}) \\
&= \sum_{\sigma\in\mathfrak{S}_n} (-1)^\sigma \prod_{i=1}^n h_{\lambda_i-\mu_{\sigma(i)}-i+\sigma(i)}(\mathbf{x}) = \det\left(\left(h_{\lambda_i-\mu_j-i+j}(\mathbf{x})\right)_{i,j=1,2,...,n}\right) \qquad (13.75.11)
\end{aligned}$$

(because applying (13.75.5) to $\ell = n$ and $\alpha_{i,j} = h_{\lambda_i-\mu_j-i+j}(\mathbf{x})$ yields $\det\left(\left(h_{\lambda_i-\mu_j-i+j}(\mathbf{x})\right)_{i,j=1,2,...,n}\right) = \sum_{\sigma\in\mathfrak{S}_n} (-1)^\sigma \prod_{i=1}^n h_{\lambda_i-\mu_{\sigma(i)}-i+\sigma(i)}(\mathbf{x})$).

On the other hand, let us study the left hand side of (13.75.2). Every partition ν having at most n parts satisfies

$$\begin{aligned}
a_{\nu+\rho}(y_1,y_2,...,y_n) &= \det\left(\left(y_i^{(\nu+\rho)_j}\right)_{i,j=1,2,...,n}\right) \qquad \text{(by the definition of the alternant } a_{\nu+\rho}\text{)} \\
&= \sum_{\sigma\in\mathfrak{S}_n} (-1)^\sigma \prod_{i=1}^n y_i^{(\nu+\rho)_{\sigma(i)}} \qquad \left(\text{by (13.75.5), applied to } \ell = n \text{ and } \alpha_{i,j} = y_i^{(\nu+\rho)_j}\right).
\end{aligned}$$

[639] *Proof.* Let $\sigma \in \mathfrak{S}_n$. It is clear that the sum

$$\sum_{\substack{(q_1,q_2,...,q_n)\in\mathbb{N}^n;\\ q_j+(\mu+\rho)_{\sigma(j)}=\lambda_j+n-j \text{ for every } j\in\{1,2,...,n\}}} \prod_{j=1}^n h_{q_j}(\mathbf{x}) \qquad (13.75.9)$$

has at most one addend: namely, the one corresponding to the n-tuple $(q_1,q_2,...,q_n) \in \mathbb{Z}^n$ defined by

$$q_j = \lambda_j + n - j - (\mu+\rho)_{\sigma(j)} \quad \text{for every } j \in \{1,2,...,n\}. \qquad (13.75.10)$$

If this n-tuple $(q_1,q_2,...,q_n)$ belongs to $\mathbb{N}^n$, then the sum (13.75.9) does have this summand. Thus, if the n-tuple $(q_1,q_2,...,q_n)$ belongs to $\mathbb{N}^n$, we have

$$\sum_{\substack{(q_1,q_2,...,q_n)\in\mathbb{N}^n;\\ q_j+(\mu+\rho)_{\sigma(j)}=\lambda_j+n-j \text{ for every } j\in\{1,2,...,n\}}} \prod_{j=1}^n h_{q_j}(\mathbf{x}) = \prod_{j=1}^n h_{\lambda_j+n-j-(\mu+\rho)_{\sigma(j)}}(\mathbf{x}).$$

Thus, if the n-tuple $(q_1,q_2,...,q_n)$ defined by (13.75.10) belongs to $\mathbb{N}^n$, the equality (13.75.8) is proven. It remains to consider the case when the n-tuple $(q_1,q_2,...,q_n)$ defined by (13.75.10) does not belong to $\mathbb{N}^n$. In this case, the sum (13.75.9) is empty (because the only addend it can have corresponds to the n-tuple $(q_1,q_2,...,q_n)$ defined by (13.75.10), but this n-tuple does not belong to $\mathbb{N}^n$ and therefore does not appear in the sum), hence equals 0. But since the n-tuple $(q_1,q_2,...,q_n)$ defined by (13.75.10) does not belong to $\mathbb{N}^n$, there must be a $j \in \{1,2,...,n\}$ satisfying $\lambda_j + n - j - (\mu+\rho)_{\sigma(j)} \notin \mathbb{N}$. For this j, we have $h_{\lambda_j+n-j-(\mu+\rho)_{\sigma(j)}} = 0$. Hence, the product $\prod_{j=1}^n h_{\lambda_j+n-j-(\mu+\rho)_{\sigma(j)}}(\mathbf{x})$ has a factor equal to 0; consequently, the product $\prod_{j=1}^n h_{\lambda_j+n-j-(\mu+\rho)_{\sigma(j)}}(\mathbf{x})$ must be 0. Thus, both the sum (13.75.9) and the product $\prod_{j=1}^n h_{\lambda_j+n-j-(\mu+\rho)_{\sigma(j)}}(\mathbf{x})$ are 0, and thus (13.75.8) is proven.

Hence, the left hand side of (13.75.2) is

$$\begin{aligned}
&\sum_{\substack{\nu\in\operatorname{Par};\\ \nu\text{ has at most }n\text{ parts}}} \underbrace{a_{\nu+\rho}\left(y_1,y_2,...,y_n\right)}_{=\sum_{\sigma\in\mathfrak{S}_n}(-1)^\sigma\prod_{i=1}^n y_i^{(\nu+\rho)_{\sigma(i)}}} s_{\nu/\mu}\left(\mathbf{x}\right)\\
&=\sum_{\substack{\nu\in\operatorname{Par};\\ \nu\text{ has at most }n\text{ parts}}} \left(\sum_{\sigma\in\mathfrak{S}_n}(-1)^\sigma\prod_{i=1}^n y_i^{(\nu+\rho)_{\sigma(i)}}\right) s_{\nu/\mu}\left(\mathbf{x}\right)\\
&=\sum_{\sigma\in\mathfrak{S}_n}(-1)^\sigma \sum_{\substack{\nu\in\operatorname{Par};\\ \nu\text{ has at most }n\text{ parts}}} \left(\prod_{i=1}^n y_i^{(\nu+\rho)_{\sigma(i)}}\right) s_{\nu/\mu}\left(\mathbf{x}\right).
\end{aligned}$$

Thus, the coefficient of $y_1^{\lambda_1+n-1}y_2^{\lambda_2+n-2}...y_n^{\lambda_n+n-n}$ on the left hand side of (13.75.2) equals

$$\sum_{\sigma\in\mathfrak{S}_n}(-1)^\sigma \sum_{\substack{\nu\in\operatorname{Par};\\ \nu\text{ has at most }n\text{ parts};\\ (\nu+\rho)_{\sigma(j)}=\lambda_j+n-j\text{ for every }j\in\{1,2,...,n\}}} s_{\nu/\mu}\left(\mathbf{x}\right). \tag{13.75.12}$$

Now let us simplify this. First, we claim that every permutation $\sigma\in\mathfrak{S}_n$ satisfying $\sigma\neq\operatorname{id}$ satisfies

$$\sum_{\substack{\nu\in\operatorname{Par};\\ \nu\text{ has at most }n\text{ parts};\\ (\nu+\rho)_{\sigma(j)}=\lambda_j+n-j\text{ for every }j\in\{1,2,...,n\}}} s_{\nu/\mu}\left(\mathbf{x}\right)=0. \tag{13.75.13}$$

[640] Thus, in the outer sum in (13.75.12), all addends which correspond to permutations $\sigma\in\mathfrak{S}_n$ satisfying $\sigma\neq\operatorname{id}$ are 0. We can therefore remove all these addends, leaving only the addend corresponding to $\sigma=\operatorname{id}$. Thus, the sum simplifies as follows:

$$\begin{aligned}
&\sum_{\sigma\in\mathfrak{S}_n}(-1)^\sigma \sum_{\substack{\nu\in\operatorname{Par};\\ \nu\text{ has at most }n\text{ parts};\\ (\nu+\rho)_{\sigma(j)}=\lambda_j+n-j\text{ for every }j\in\{1,2,...,n\}}} s_{\nu/\mu}\left(\mathbf{x}\right)=\underbrace{(-1)^{\operatorname{id}}}_{=1}\underbrace{\sum_{\substack{\nu\in\operatorname{Par};\\ \nu\text{ has at most }n\text{ parts};\\ (\nu+\rho)_{\operatorname{id}(j)}=\lambda_j+n-j\text{ for every }j\in\{1,2,...,n\}}}}_{\substack{=\sum_{\substack{\nu\in\operatorname{Par};\\ \nu\text{ has at most }n\text{ parts};\\ \nu_j+n-j=\lambda_j+n-j\text{ for every }j\in\{1,2,...,n\}}}\\ (\text{since }(\nu+\rho)_{\operatorname{id}(j)}=(\nu+\rho)_j=\nu_j+n-j\text{ for all }j)}} s_{\nu/\mu}\left(\mathbf{x}\right)\\
&=\sum_{\substack{\nu\in\operatorname{Par};\\ \nu\text{ has at most }n\text{ parts};\\ \nu_j+n-j=\lambda_j+n-j\text{ for every }j\in\{1,2,...,n\}}} s_{\nu/\mu}\left(\mathbf{x}\right)=s_{\lambda/\mu}\left(\mathbf{x}\right)
\end{aligned}$$

(because there is only one $\nu\in\operatorname{Par}$ such that ν has at most n parts and satisfies $\nu_j+n-j=\lambda_j+n-j$ for every $j\in\{1,2,...,n\}$; namely, this ν is λ). Hence, the coefficient of $y_1^{\lambda_1+n-1}y_2^{\lambda_2+n-2}...y_n^{\lambda_n+n-n}$ on the left

[640] *Proof of (13.75.13):* Let $\sigma\in\mathfrak{S}_n$ be a permutation satisfying $\sigma\neq\operatorname{id}$. We need to prove (13.75.13). Of course, it is enough to show that the sum on the left hand side of (13.75.13) is empty, i.e., that there exists no $\nu\in\operatorname{Par}$ such that ν has at most n parts and satisfies $(\nu+\rho)_{\sigma(j)}=\lambda_j+n-j$ for every $j\in\{1,2,...,n\}$. Assume the contrary. Then, there exists a $\nu\in\operatorname{Par}$ such that ν has at most n parts and satisfies $(\nu+\rho)_{\sigma(j)}=\lambda_j+n-j$ for every $j\in\{1,2,...,n\}$. Consider this ν. Since ν is a partition, $\nu+\rho$ is a strict partition, i.e., we have $(\nu+\rho)_1>(\nu+\rho)_2>...>(\nu+\rho)_n$. In other words, the entries of the n-tuple $\nu+\rho$ are in strictly decreasing order. Since $\sigma\neq\operatorname{id}$, the permutation σ must mess up the order of the entries of $\nu+\rho$; thus, we **cannot have** $(\nu+\rho)_{\sigma(1)}>(\nu+\rho)_{\sigma(2)}>...>(\nu+\rho)_{\sigma(n)}$. Since $(\nu+\rho)_{\sigma(j)}=\lambda_j+n-j$ for every $j\in\{1,2,...,n\}$, this rewrites as follows: We **cannot** have $\lambda_1+n-1>\lambda_2+n-2>...>\lambda_n+n-n$.

But since λ is a partition, we have $\lambda_1\geq\lambda_2\geq...\geq\lambda_n$, thus $\lambda_1+n-1>\lambda_2+n-2>...>\lambda_n+n-n$. This contradicts the fact that we cannot have $\lambda_1+n-1>\lambda_2+n-2>...>\lambda_n+n-n$. This contradiction finishes the proof.

hand side of (13.75.2) equals

$$\sum_{\sigma\in\mathfrak{S}_n}(-1)^\sigma \sum_{\substack{\nu\in\operatorname{Par};\\ \nu\text{ has at most }n\text{ parts;}\\ (\nu+\rho)_{\sigma(j)}=\lambda_j+n-j\text{ for every }j\in\{1,2,\ldots,n\}}} s_{\nu/\mu}(\mathbf{x}) \qquad (\text{by (13.75.12)})$$

$$= s_{\lambda/\mu}(\mathbf{x}). \tag{13.75.14}$$

But the coefficients of $y_1^{\lambda_1+n-1}y_2^{\lambda_2+n-2}...y_n^{\lambda_n+n-n}$ on the left hand side of (13.75.2) and on the right hand side of (13.75.2) must be equal. Since the former coefficient is $s_{\lambda/\mu}(\mathbf{x})$ (by (13.75.14)), and the latter coefficient is $\det\left(\left(h_{\lambda_i-\mu_j-i+j}(\mathbf{x})\right)_{i,j=1,2,\ldots,n}\right)$ (by (13.75.11)), this shows that

$$s_{\lambda/\mu}(\mathbf{x}) = \det\left(\left(h_{\lambda_i-\mu_j-i+j}(\mathbf{x})\right)_{i,j=1,2,\ldots,n}\right).$$

In other words, $s_{\lambda/\mu} = \det\left(\left(h_{\lambda_i-\mu_j-i+j}\right)_{i,j=1,2,\ldots,n}\right)$.

Now, forget that we fixed n, λ and μ. We thus have proven that if $n\in\mathbb{N}$, and if λ and μ are two partitions having at most n parts (each), then $s_{\lambda/\mu} = \det\left(\left(h_{\lambda_i-\mu_j-i+j}\right)_{i,j=1,2,\ldots,n}\right)$. Renaming n as ℓ in this claim, we obtain: If $\ell\in\mathbb{N}$, and if λ and μ are two partitions having at most ℓ parts (each), then

$$s_{\lambda/\mu} = \det\left(\left(h_{\lambda_i-\mu_j-i+j}\right)_{i,j=1,2,\ldots,\ell}\right). \tag{13.75.15}$$

This proves (2.4.16).

Now it remains to prove (2.4.17). Let $\ell\in\mathbb{N}$, and let λ and μ be two partitions having at most ℓ parts (each).

Let us first notice that every $m\in\mathbb{Z}$ satisfies

$$\omega(h_m) = e_m. \tag{13.75.16}$$

(Indeed, this follows from Proposition 2.4.3(b), applied to m instead of n.) But using (2.4.15), it is easy to see that

$$\omega\left(s_{\lambda/\mu}\right) = s_{\lambda^t/\mu^t}. \tag{13.75.17}$$

[641] Thus,

$$\begin{aligned} s_{\lambda^t/\mu^t} &= \omega\left(s_{\lambda/\mu}\right) = \omega\left(\det\left(\left(h_{\lambda_i-\mu_j-i+j}\right)_{i,j=1,2,\ldots,\ell}\right)\right) \qquad (\text{by (13.75.15)})\\ &= \det\left(\left(\underbrace{\omega\left(h_{\lambda_i-\mu_j-i+j}\right)}_{\substack{=e_{\lambda_i-\mu_j-i+j}\\ (\text{by (13.75.16)})}}\right)_{i,j=1,2,\ldots,\ell}\right)\\ &\qquad \left(\begin{array}{c}\text{since }\omega\text{ is a }\mathbf{k}\text{-algebra homomorphism, and thus}\\ \text{commutes with taking determinants}\end{array}\right)\\ &= \det\left(\left(e_{\lambda_i-\mu_j-i+j}\right)_{i,j=1,2,\ldots,\ell}\right). \end{aligned}$$

This proves (2.4.17). Therefore, the proof of Theorem 2.4.6 is complete.

---

[641] In fact, (13.75.17) follows immediately from (2.4.15) in the case when $\mu\subseteq\lambda$; but otherwise it follows from $s_{\lambda/\mu}=0$ and $s_{\lambda^t/\mu^t}=0$.

13.76. **Solution to Exercise 2.7.14.** *Solution to Exercise 2.7.14.* (a) The second identity of (2.4.15) shows that

$$S\left(s_{\lambda/\mu}\right) = (-1)^{|\lambda/\mu|} s_{\lambda^t/\mu^t} \tag{13.76.1}$$

whenever λ and μ are partitions satisfying $\mu \subseteq \lambda$. Hence,

$$S\left(s_{\lambda}\right) = (-1)^{|\lambda|} s_{\lambda^t} \qquad \text{for any partition } \lambda. \tag{13.76.2}$$

[*Proof of (13.76.2):* Let λ be any partition. Then, the empty partition $\varnothing$ clearly satisfies $\varnothing \subseteq \lambda$. Hence, (13.76.1) (applied to $\mu = \varnothing$) yields $S\left(s_{\lambda/\varnothing}\right) = (-1)^{|\lambda/\varnothing|} s_{\lambda^t/\varnothing^t}$. In view of $s_{\lambda/\varnothing} = s_\lambda$ and $\varnothing^t = \varnothing$, this rewrites as $S\left(s_{\lambda}\right) = (-1)^{|\lambda/\varnothing|} s_{\lambda^t/\varnothing}$. In view of $|\lambda/\varnothing| = |\lambda|$ and $s_{\lambda^t/\varnothing} = s_{\lambda^t}$, this rewrites as $S\left(s_{\lambda}\right) = (-1)^{|\lambda|} s_{\lambda^t}$. This proves (13.76.2).]

On the other hand, the definition of the Hall inner product $(\cdot,\cdot)$ shows that

$$\left(s_\lambda, s_\nu\right) = \delta_{\lambda,\nu} \qquad \text{for any two partitions } \lambda \text{ and } \nu. \tag{13.76.3}$$

Now, let $f \in \Lambda$ and $g \in \Lambda$.

Proposition 2.2.10 shows that the family $\left(s_\lambda\right)_{\lambda\in\operatorname{Par}}$ is a basis of the $\mathbf{k}$-module Λ. Hence, f can be written in the form $f = \sum_{\lambda\in\operatorname{Par}} a_\lambda s_\lambda$ for some family $\left(a_\lambda\right)_{\lambda\in\operatorname{Par}} \in \mathbf{k}^{\operatorname{Par}}$ of elements of $\mathbf{k}$ such that all but finitely many $\lambda \in \operatorname{Par}$ satisfy $a_\lambda = 0$. Consider this family $\left(a_\lambda\right)_{\lambda\in\operatorname{Par}}$.

Proposition 2.2.10 shows that the family $\left(s_\lambda\right)_{\lambda\in\operatorname{Par}}$ is a basis of the $\mathbf{k}$-module Λ. Hence, g can be written in the form $g = \sum_{\lambda\in\operatorname{Par}} b_\lambda s_\lambda$ for some family $\left(b_\lambda\right)_{\lambda\in\operatorname{Par}} \in \mathbf{k}^{\operatorname{Par}}$ of elements of $\mathbf{k}$ such that all but finitely many $\lambda \in \operatorname{Par}$ satisfy $b_\lambda = 0$. Consider this family $\left(b_\lambda\right)_{\lambda\in\operatorname{Par}}$.

The map $\operatorname{Par} \to \operatorname{Par},\ \lambda \mapsto \lambda^t$ is a bijection. (Indeed, this map is inverse to itself, since each partition λ satisfies $\left(\lambda^t\right)^t = \lambda$).

Applying the map S to both sides of the equality $f = \sum_{\lambda\in\operatorname{Par}} a_\lambda s_\lambda$, we obtain

$$\begin{aligned} S(f) &= S\left(\sum_{\lambda\in\operatorname{Par}} a_\lambda s_\lambda\right) = \sum_{\lambda\in\operatorname{Par}} a_\lambda \underbrace{S\left(s_\lambda\right)}_{\substack{=(-1)^{|\lambda|} s_{\lambda^t} \\ \text{(by (13.76.2))}}} \qquad \text{(since the map } S \text{ is } \mathbf{k}\text{-linear)} \\ &= \sum_{\lambda\in\operatorname{Par}} a_\lambda (-1)^{|\lambda|} s_{\lambda^t} = \sum_{\lambda\in\operatorname{Par}} a_{\lambda^t} \underbrace{(-1)^{\left|\lambda^t\right|}}_{\substack{=(-1)^{|\lambda|} \\ \text{(since } \left|\lambda^t\right| = |\lambda|)}} \underbrace{s_{\left(\lambda^t\right)^t}}_{\substack{=s_\lambda \\ \text{(since } \left(\lambda^t\right)^t = \lambda)}} \\ &\qquad \left(\begin{array}{c}\text{here, we have substituted } \lambda^t \text{ for } \lambda \text{ in the sum, since} \\ \text{the map } \operatorname{Par} \to \operatorname{Par},\ \lambda \mapsto \lambda^t \text{ is a bijection}\end{array}\right) \\ &= \sum_{\lambda\in\operatorname{Par}} a_{\lambda^t} (-1)^{|\lambda|} s_\lambda. \end{aligned}$$

The same argument (applied to g and b_λ instead of f and a_λ) yields

$$\begin{aligned} S(g) &= \sum_{\lambda\in\operatorname{Par}} b_{\lambda^t} (-1)^{|\lambda|} s_\lambda \qquad \left(\text{since } g = \sum_{\lambda\in\operatorname{Par}} b_\lambda s_\lambda\right) \\ &= \sum_{\mu\in\operatorname{Par}} b_{\mu^t} (-1)^{|\mu|} s_\mu \end{aligned}$$

(here, we have renamed the summation index λ as μ in the sum). Also,

$$g = \sum_{\lambda\in\operatorname{Par}} b_\lambda s_\lambda = \sum_{\mu\in\operatorname{Par}} b_\mu s_\mu$$

(here, we have renamed the summation index λ as μ in the sum).

Now,

$$\begin{aligned}
&\left(\underbrace{S\left(f\right)}_{=\sum_{\lambda\in\operatorname{Par}} a_{\lambda^t}(-1)^{|\lambda|} s_\lambda} , \underbrace{S\left(g\right)}_{=\sum_{\mu\in\operatorname{Par}} b_{\mu^t}(-1)^{|\mu|} s_\mu} \right) \\
&= \left(\sum_{\lambda\in\operatorname{Par}} a_{\lambda^t}\left(-1\right)^{|\lambda|} s_\lambda, \sum_{\mu\in\operatorname{Par}} b_{\mu^t}\left(-1\right)^{|\mu|} s_\mu \right) \\
&= \sum_{\lambda\in\operatorname{Par}} a_{\lambda^t}\left(-1\right)^{|\lambda|} \sum_{\mu\in\operatorname{Par}} b_{\mu^t}\left(-1\right)^{|\mu|} \underbrace{\left(s_\lambda, s_\mu\right)}_{\substack{=\delta_{\lambda,\mu} \\ \text{(by (13.76.3),} \\ \text{applied to } \nu=\mu)}} \\
&\qquad \left(\text{since the Hall inner product } \left(\cdot,\cdot\right) \text{ is } \mathbf{k}\text{-bilinear}\right) \\
&= \sum_{\lambda\in\operatorname{Par}} a_{\lambda^t}\left(-1\right)^{|\lambda|} \underbrace{\sum_{\mu\in\operatorname{Par}} b_{\mu^t}\left(-1\right)^{|\mu|} \delta_{\lambda,\mu}}_{\substack{=b_{\lambda^t}(-1)^{|\lambda|}\delta_{\lambda,\lambda}+\sum_{\substack{\mu\in\operatorname{Par};\\ \mu\neq\lambda}} b_{\mu^t}(-1)^{|\mu|}\delta_{\lambda,\mu} \\ \text{(here, we have split off the addend for } \mu=\lambda \text{ from the sum)}}} \\
&= \sum_{\lambda\in\operatorname{Par}} a_{\lambda^t}\left(-1\right)^{|\lambda|} \left(b_{\lambda^t}\left(-1\right)^{|\lambda|} \underbrace{\delta_{\lambda,\lambda}}_{=1} + \sum_{\substack{\mu\in\operatorname{Par};\\ \mu\neq\lambda}} b_{\mu^t}\left(-1\right)^{|\mu|} \underbrace{\delta_{\lambda,\mu}}_{\substack{=0 \\ \text{(since } \lambda\neq\mu \\ \text{(since } \mu\neq\lambda\text{))}}} \right) \\
&= \sum_{\lambda\in\operatorname{Par}} a_{\lambda^t}\left(-1\right)^{|\lambda|} \left(b_{\lambda^t}\left(-1\right)^{|\lambda|} + \underbrace{\sum_{\substack{\mu\in\operatorname{Par};\\ \mu\neq\lambda}} b_{\mu^t}\left(-1\right)^{|\mu|} 0}_{=0} \right) \\
&= \sum_{\lambda\in\operatorname{Par}} a_{\lambda^t} \underbrace{\left(-1\right)^{|\lambda|} b_{\lambda^t}}_{=b_{\lambda^t}(-1)^{|\lambda|}} \left(-1\right)^{|\lambda|} = \sum_{\lambda\in\operatorname{Par}} a_{\lambda^t} b_{\lambda^t} \underbrace{\left(-1\right)^{|\lambda|}\left(-1\right)^{|\lambda|}}_{\substack{=(-1)^{|\lambda|+|\lambda|}=1 \\ \text{(since } |\lambda|+|\lambda|=2|\lambda| \text{ is even)}}} = \sum_{\lambda\in\operatorname{Par}} a_{\lambda^t} b_{\lambda^t} = \sum_{\lambda\in\operatorname{Par}} a_\lambda b_\lambda
\end{aligned}$$

(here, we have substituted λ^t for λ in the sum, since the map $\operatorname{Par} \to \operatorname{Par},\ \lambda \mapsto \lambda^t$ is a bijection). Comparing this with

$$\begin{aligned}
&\left(\underbrace{f}_{=\sum_{\lambda\in\operatorname{Par}} a_\lambda s_\lambda}, \underbrace{g}_{=\sum_{\mu\in\operatorname{Par}} b_\mu s_\mu}\right) \\
&= \left(\sum_{\lambda\in\operatorname{Par}} a_\lambda s_\lambda, \sum_{\mu\in\operatorname{Par}} b_\mu s_\mu\right) = \sum_{\lambda\in\operatorname{Par}} a_\lambda \sum_{\mu\in\operatorname{Par}} b_\mu \underbrace{(s_\lambda, s_\mu)}_{\substack{=\delta_{\lambda,\mu} \\ \text{(by (13.76.3),} \\ \text{applied to } \nu=\mu)}} \\
&\qquad \text{(since the Hall inner product } (\cdot,\cdot) \text{ is } \mathbf{k}\text{-bilinear)} \\
&= \sum_{\lambda\in\operatorname{Par}} a_\lambda \underbrace{\sum_{\mu\in\operatorname{Par}} b_\mu \delta_{\lambda,\mu}}_{\substack{=b_\lambda\delta_{\lambda,\lambda} + \sum_{\substack{\mu\in\operatorname{Par}; \\ \mu\neq\lambda}} b_\mu\delta_{\lambda,\mu} \\ \text{(here, we have split off the addend for } \mu=\lambda \text{ from the sum)}}} \\
&= \sum_{\lambda\in\operatorname{Par}} a_\lambda \left(b_\lambda \underbrace{\delta_{\lambda,\lambda}}_{=1} + \sum_{\substack{\mu\in\operatorname{Par}; \\ \mu\neq\lambda}} b_\mu \underbrace{\delta_{\lambda,\mu}}_{\substack{=0 \\ \text{(since } \lambda\neq\mu \\ \text{(since } \mu\neq\lambda\text{))}}}\right) = \sum_{\lambda\in\operatorname{Par}} a_\lambda \left(b_\lambda + \underbrace{\sum_{\substack{\mu\in\operatorname{Par}; \\ \mu\neq\lambda}} b_\mu 0}_{=0}\right) = \sum_{\lambda\in\operatorname{Par}} a_\lambda b_\lambda,
\end{aligned}$$

we obtain $(S(f), S(g)) = (f, g)$. This solves Exercise 2.7.14(a).

(b) Let $n \in \mathbb{N}$ and $f \in \Lambda_n$. We know that the Hopf algebra Λ is graded; thus, its antipode S is a graded $\mathbf{k}$-linear map. Hence, $S(\Lambda_n) \subset \Lambda_n$. Now, from $f \in \Lambda_n$, we obtain $S(f) \in S(\Lambda_n) \subset \Lambda_n$. Hence, Exercise 2.5.13(b) (applied to $S(f)$ instead of f) yields $(h_n, S(f)) = (S(f))(1)$.

But Proposition 2.4.1(ii) yields $S(e_n) = (-1)^n h_n$. Hence,

$$\begin{aligned}
&\left(\underbrace{S(e_n)}_{=(-1)^n h_n}, S(f)\right) \\
&= ((-1)^n h_n, S(f)) = (-1)^n \cdot \underbrace{(h_n, S(f))}_{=(S(f))(1)} \qquad \text{(since the Hall inner product } (\cdot,\cdot) \text{ is } \mathbf{k}\text{-bilinear)} \\
&= (-1)^n \cdot (S(f))(1).
\end{aligned}$$

But Exercise 2.7.14(a) (applied to e_n and f instead of f and g) yields $(S(e_n), S(f)) = (e_n, f)$. Comparing these two equalities, we obtain $(e_n, f) = (-1)^n \cdot (S(f))(1)$. This solves Exercise 2.7.14(b).

13.77. **Solution to Exercise 2.8.4.**

Solution to Exercise 2.8.4. Let us first prove two lemmas:

Lemma 13.77.1. *Let $f \in \Lambda$ and $g \in \Lambda$. Assume that*

$$(s_\lambda, f) = (s_\lambda, g) \qquad \text{for each } \lambda \in \operatorname{Par}. \tag{13.77.1}$$

Then, $f = g$.

Proof of Lemma 13.77.1. The basis $(s_\lambda)_{\lambda\in\operatorname{Par}}$ of Λ is orthonormal with respect to the Hall inner product $(\cdot,\cdot)$ (by Definition 2.5.12). In other words, the two (identical) bases $(s_\lambda)_{\lambda\in\operatorname{Par}}$ and $(s_\lambda)_{\lambda\in\operatorname{Par}}$ of Λ are dual to each other with respect to the Hall inner product $(\cdot,\cdot)$.

Recall the following fundamental fact from linear algebra: If $\mathbf{k}$ is a commutative ring, if A is a $\mathbf{k}$-module, if $(\cdot,\cdot) : A \times A \to \mathbf{k}$ is a symmetric $\mathbf{k}$-bilinear form on A, and if $(u_\lambda)_{\lambda\in L}$ and $(v_\lambda)_{\lambda\in L}$ are two $\mathbf{k}$-bases of A

which are dual to each other with respect to the form $(\cdot, \cdot)$ (where L is some indexing set), then every $a \in A$ satisfies

$$a = \sum_{\lambda \in L} (u_\lambda, a) \, v_\lambda . \tag{13.77.2}$$

We can apply this fact to $A = \Lambda$, $L = \operatorname{Par}$, $(u_\lambda)_{\lambda \in L} = (s_\lambda)_{\lambda \in \operatorname{Par}}$ and $(v_\lambda)_{\lambda \in L} = (s_\lambda)_{\lambda \in \operatorname{Par}}$ (since the bases $(s_\lambda)_{\lambda \in \operatorname{Par}}$ and $(s_\lambda)_{\lambda \in \operatorname{Par}}$ of Λ are dual to each other with respect to the Hall inner product $(\cdot, \cdot)$). We thus conclude that every $a \in \Lambda$ satisfies

$$a = \sum_{\lambda \in \operatorname{Par}} (s_\lambda, a) \, s_\lambda . \tag{13.77.3}$$

Applying this to $a = f$, we obtain

$$f = \sum_{\lambda \in \operatorname{Par}} \underbrace{(s_\lambda, f)}_{\substack{=(s_\lambda, g) \\ \text{(by (13.77.1))}}} s_\lambda = \sum_{\lambda \in \operatorname{Par}} (s_\lambda, g) \, s_\lambda .$$

Comparing this with

$$g = \sum_{\lambda \in \operatorname{Par}} (s_\lambda, g) \, s_\lambda \qquad \text{(by (13.77.3), applied to } a = g\text{)} ,$$

we obtain $f = g$. This proves Lemma 13.77.1. □

Lemma 13.77.2. *Let $\gamma \in \operatorname{Par}$ and $k \in \mathbb{N}$.*

(a) *We have*

$$h_k^\perp s_\gamma = \sum_{\substack{\nu \in \operatorname{Par}; \\ \gamma/\nu \text{ is a} \\ \text{horizontal } k\text{-strip}}} s_\nu . \tag{13.77.4}$$

(b) *We have*

$$e_k^\perp s_\gamma = \sum_{\substack{\nu \in \operatorname{Par}; \\ \gamma/\nu \text{ is a} \\ \text{vertical } k\text{-strip}}} s_\nu . \tag{13.77.5}$$

Proof of Lemma 13.77.2. We know that every $f \in \Lambda$, $g \in \Lambda$ and $a \in \Lambda$ satisfy

$$\left(g, f^\perp (a)\right) = (fg, a) . \tag{13.77.6}$$

Indeed, this follows from Proposition 2.8.2(i) (applied to $A = \Lambda$), after we make the standard identification of Λ^o with Λ via the Hall inner product.

Recall that the basis $(s_\lambda)_{\lambda \in \operatorname{Par}}$ of Λ is orthonormal with respect to the Hall inner product $(\cdot, \cdot)$ (by Definition 2.5.12). In other words,

$$(s_\alpha, s_\beta) = \delta_{\alpha, \beta} \qquad \text{for every } (\alpha, \beta) \in \operatorname{Par} \times \operatorname{Par} . \tag{13.77.7}$$

(b) We claim that

$$\left(s_\lambda, e_k^\perp s_\gamma\right) = \left(s_\lambda, \sum_{\substack{\nu \in \operatorname{Par}; \\ \gamma/\nu \text{ is a} \\ \text{vertical } k\text{-strip}}} s_\nu \right) \tag{13.77.8}$$

for every $\lambda \in \operatorname{Par}$.

Proof of (13.77.8): Let $\lambda \in \operatorname{Par}$. We have

$$e_k s_\lambda = s_\lambda e_k = \sum_{\substack{\lambda^+ : \lambda^+/\lambda \text{ is a} \\ \text{vertical } k\text{-strip}}} s_{\lambda^+} \tag{13.77.9}$$

(by (2.7.2), applied to $n = k$). But (13.77.6) (applied to $f = e_k$, $a = s_\gamma$ and $g = s_\lambda$) yields

$$\begin{aligned}\left(s_\lambda, e_k^\perp s_\gamma\right) = \left(e_k s_\lambda, s_\gamma\right) &= \left(\sum_{\substack{\lambda^+ : \lambda^+/\lambda \text{ is a} \\ \text{vertical } k\text{-strip}}} s_{\lambda^+}, s_\gamma\right) && \text{(by (13.77.9))} \\ &= \sum_{\substack{\lambda^+ : \lambda^+/\lambda \text{ is a} \\ \text{vertical } k\text{-strip}}} \underbrace{\left(s_{\lambda^+}, s_\gamma\right)}_{\substack{=\delta_{\lambda^+,\gamma} \\ \text{(by (13.77.7), applied} \\ \text{to } (\alpha,\beta)=(\lambda^+,\gamma))}} && \text{(since the Hall inner product is } \mathbf{k}\text{-bilinear)} \\ &= \sum_{\substack{\lambda^+ : \lambda^+/\lambda \text{ is a} \\ \text{vertical } k\text{-strip}}} \delta_{\lambda^+,\gamma} = \begin{cases} 1, & \text{if } \gamma/\lambda \text{ is a vertical } k\text{-strip;} \\ 0, & \text{otherwise} \end{cases} . \end{aligned} \tag{13.77.10}$$

On the other hand, the Hall inner product is **k**-bilinear. Thus,

$$\begin{aligned}\left(s_\lambda, \sum_{\substack{\nu \in \mathrm{Par}; \\ \gamma/\nu \text{ is a} \\ \text{vertical } k\text{-strip}}} s_\nu\right) &= \sum_{\substack{\nu \in \mathrm{Par}; \\ \gamma/\nu \text{ is a} \\ \text{vertical } k\text{-strip}}} \underbrace{\left(s_\lambda, s_\nu\right)}_{\substack{=\delta_{\lambda,\nu} \\ \text{(by (13.77.7), applied} \\ \text{to } (\alpha,\beta)=(\lambda,\nu))}} \\ &= \sum_{\substack{\nu \in \mathrm{Par}; \\ \gamma/\nu \text{ is a} \\ \text{vertical } k\text{-strip}}} \delta_{\lambda,\nu} = \begin{cases} 1, & \text{if } \gamma/\lambda \text{ is a vertical } k\text{-strip;} \\ 0, & \text{otherwise} \end{cases} .\end{aligned}$$

Comparing this with (13.77.10), we obtain

$$\left(s_\lambda, e_k^\perp s_\gamma\right) = \left(s_\lambda, \sum_{\substack{\nu \in \mathrm{Par}; \\ \gamma/\nu \text{ is a} \\ \text{vertical } k\text{-strip}}} s_\nu\right).$$

Thus, (13.77.8) is proven.

Now, we have proven (13.77.8) for all $\lambda \in \mathrm{Par}$. Hence, Lemma 13.77.1 (applied to $f = e_k^\perp s_\gamma$ and $g = \sum_{\substack{\nu \in \mathrm{Par}; \\ \gamma/\nu \text{ is a} \\ \text{vertical } k\text{-strip}}} s_\nu$) yields $e_k^\perp s_\gamma = \sum_{\substack{\nu \in \mathrm{Par}; \\ \gamma/\nu \text{ is a} \\ \text{vertical } k\text{-strip}}} s_\nu$. This proves Lemma 13.77.2(b).

(a) We claim that

$$\left(s_\lambda, h_k^\perp s_\gamma\right) = \left(s_\lambda, \sum_{\substack{\nu \in \mathrm{Par}; \\ \gamma/\nu \text{ is a} \\ \text{horizontal } k\text{-strip}}} s_\nu\right) \tag{13.77.11}$$

for every $\lambda \in \mathrm{Par}$.

Proof of (13.77.11): Let $\lambda \in \mathrm{Par}$. Then,

$$h_k s_\lambda = s_\lambda h_k = \sum_{\substack{\lambda^+ : \lambda^+/\lambda \text{ is a} \\ \text{horizontal } k\text{-strip}}} s_{\lambda^+} \tag{13.77.12}$$

(by (2.7.1), applied to $n = k$). But (13.77.6) (applied to $f = h_k$, $a = s_\gamma$ and $g = s_\lambda$) yields

$$\begin{aligned}
\left(s_\lambda, h_k^\perp s_\gamma\right) = \left(h_k s_\lambda, s_\gamma\right) &= \left(\sum_{\substack{\lambda^+ : \lambda^+/\lambda \text{ is a} \\ \text{horizontal } k\text{-strip}}} s_{\lambda^+}, s_\gamma \right) \qquad \text{(by (13.77.12))} \\
&= \sum_{\substack{\lambda^+ : \lambda^+/\lambda \text{ is a} \\ \text{horizontal } k\text{-strip}}} \underbrace{\left(s_{\lambda^+}, s_\gamma\right)}_{\substack{=\delta_{\lambda^+,\gamma} \\ \text{(by (13.77.7), applied} \\ \text{to } (\alpha,\beta)=\left(\lambda^+,\gamma\right))}} \qquad \text{(since the Hall inner product is } \mathbf{k}\text{-bilinear)} \\
&= \sum_{\substack{\lambda^+ : \lambda^+/\lambda \text{ is a} \\ \text{horizontal } k\text{-strip}}} \delta_{\lambda^+,\gamma} = \begin{cases} 1, & \text{if } \gamma/\lambda \text{ is a horizontal } k\text{-strip;} \\ 0, & \text{otherwise} \end{cases} . \qquad (13.77.13)
\end{aligned}$$

On the other hand, the Hall inner product is **k**-bilinear. Thus,

$$\begin{aligned}
\left(s_\lambda, \sum_{\substack{\nu\in\operatorname{Par}; \\ \gamma/\nu \text{ is a} \\ \text{horizontal } k\text{-strip}}} s_\nu \right) &= \sum_{\substack{\nu\in\operatorname{Par}; \\ \gamma/\nu \text{ is a} \\ \text{horizontal } k\text{-strip}}} \underbrace{\left(s_\lambda, s_\nu\right)}_{\substack{=\delta_{\lambda,\nu} \\ \text{(by (13.77.7), applied} \\ \text{to } (\alpha,\beta)=(\lambda,\nu))}} \\
&= \sum_{\substack{\nu\in\operatorname{Par}; \\ \gamma/\nu \text{ is a} \\ \text{horizontal } k\text{-strip}}} \delta_{\lambda,\nu} = \begin{cases} 1, & \text{if } \gamma/\lambda \text{ is a horizontal } k\text{-strip;} \\ 0, & \text{otherwise} \end{cases} .
\end{aligned}$$

Comparing this with (13.77.13), we obtain

$$\left(s_\lambda, h_k^\perp s_\gamma\right) = \left(s_\lambda, \sum_{\substack{\nu\in\operatorname{Par}; \\ \gamma/\nu \text{ is a} \\ \text{horizontal } k\text{-strip}}} s_\nu \right).$$

Thus, (13.77.11) is proven.

Now, we have proven (13.77.11) for all $\lambda \in \operatorname{Par}$. Hence, Lemma 13.77.1 (applied to $f = h_k^\perp s_\gamma$ and $g = \sum_{\substack{\nu\in\operatorname{Par}; \\ \gamma/\nu \text{ is a} \\ \text{horizontal } k\text{-strip}}} s_\nu$) yields $h_k^\perp (s_\gamma) = \sum_{\substack{\nu\in\operatorname{Par}; \\ \gamma/\nu \text{ is a} \\ \text{horizontal } k\text{-strip}}} s_\nu$. This proves Lemma 13.77.2(a). □

Now, proving Proposition 2.8.3 is just a matter of renaming symbols in Lemma 13.77.2:

Proof of Proposition 2.8.3. Let λ be a partition. Let $n \in \mathbb{N}$. Then, Lemma 13.77.2(a) (applied to $\gamma = \lambda$ and $k = n$) yields

$$h_n^\perp s_\lambda = \sum_{\substack{\nu\in\operatorname{Par}; \\ \lambda/\nu \text{ is a} \\ \text{horizontal } n\text{-strip}}} s_\nu = \sum_{\substack{\lambda^-\in\operatorname{Par}; \\ \lambda/\lambda^- \text{ is a} \\ \text{horizontal } n\text{-strip}}} s_{\lambda^-}$$

(here, we have renamed the summation index ν as λ^-). This is precisely the equality (2.8.3). Thus, (2.8.3) is proven.

Lemma 13.77.2(b) (applied to $\gamma = \lambda$ and $k = n$) yields

$$e_n^\perp s_\lambda = \sum_{\substack{\nu\in\operatorname{Par}; \\ \lambda/\nu \text{ is a} \\ \text{vertical } n\text{-strip}}} s_\nu = \sum_{\substack{\lambda^-\in\operatorname{Par}; \\ \lambda/\lambda^- \text{ is a} \\ \text{vertical } n\text{-strip}}} s_{\lambda^-}$$

(here, we have renamed the summation index ν as λ^-). This is precisely the equality (2.8.4). Thus, (2.8.4) is proven. This completes the proof of Proposition 2.8.3. □

Thus, Exercise 2.8.4 is solved.

13.78. **Solution to Exercise 2.8.6.** *Solution to Exercise 2.8.6.*

Let $(\Lambda_{\mathbb{Q}})_n$ denote the n-th graded component of the $\mathbb{Q}$-vector space $\Lambda_{\mathbb{Q}}$. We are going to work in $\Lambda_{\mathbb{Q}}$ in this solution, making use of the fact that both $(s_\lambda)_{\lambda\in\operatorname{Par}_n}$ and $(p_\lambda)_{\lambda\in\operatorname{Par}_n}$ are bases of the $\mathbb{Q}$-vector space $(\Lambda_{\mathbb{Q}})_n$ (due to Proposition 2.2.10). Of course, notions such as comultiplication, the antipode, skewing etc. are defined in $\Lambda_{\mathbb{Q}}$ just in the same way as they have been defined in Λ, and their properties are proven analogously.

Our solutions for both parts of the exercise rely on the fact that the trace of an endomorphism of a finite-dimensional vector space can be computed using any basis of the vector space. Specifically, we will be applying this fact to certain endomorphisms of $(\Lambda_{\mathbb{Q}})_n$, and as bases we will use $(p_\lambda)_{\lambda\in\operatorname{Par}_n}$ and $(s_\lambda)_{\lambda\in\operatorname{Par}_n}$.

(a) The antipode S of $\Lambda_{\mathbb{Q}}$ is a graded map, and thus it restricts to an endomorphism of the $\mathbb{Q}$-vector space $(\Lambda_{\mathbb{Q}})_n$. Denote this endomorphism by S_n. We want to compute $\operatorname{trace}(S_n)$. (This is well-defined since $(\Lambda_{\mathbb{Q}})_n$ is a finite-dimensional $\mathbb{Q}$-vector space.)

From (2.4.15), we see that $S\left(s_{\lambda/\mu}\right) = (-1)^{|\lambda/\mu|} s_{\lambda^t/\mu^t}$ for any partitions λ and μ satisfying $\mu \subseteq \lambda$. In particular, $S(s_\lambda) = (-1)^{|\lambda|} s_{\lambda^t}$ for any partition λ. For any partition $\lambda \in \operatorname{Par}_n$, we now have $S_n(s_\lambda) = S(s_\lambda) = (-1)^{|\lambda|} s_{\lambda^t} = (-1)^n s_{\lambda^t}$. Thus, the matrix which represents the endomorphism S_n of $(\Lambda_{\mathbb{Q}})_n$ with respect to the basis $(s_\lambda)_{\lambda\in\operatorname{Par}_n}$ of $(\Lambda_{\mathbb{Q}})_n$ is the matrix whose (λ,μ)-th entry (for any $\lambda\in\operatorname{Par}_n$ and $\mu\in\operatorname{Par}_n$) is $(-1)^n$ whenever $\mu=\lambda^t$, and otherwise 0. But the trace $\operatorname{trace}(S_n)$ of S_n (by its definition) is the sum of the diagonal entries of this matrix; hence, this trace equals

$$\operatorname{trace}(S_n) = \sum_{\substack{\lambda\in\operatorname{Par}_n;\\ \lambda=\lambda^t}} (-1)^n = (-1)^n \underbrace{\left(\text{number of all } \lambda\in\operatorname{Par}_n \text{ satisfying } \lambda=\lambda^t\right)}_{=c(n)} = (-1)^n c(n).$$

On the other hand, every partition λ satisfies $S(p_\lambda) = (-1)^{\ell(\lambda)} p_\lambda$ (this follows from Proposition 2.4.1(i), after recalling that S is an algebra morphism and that $p_{(\lambda_1,\lambda_2,\ldots,\lambda_\ell)} = p_{\lambda_1}p_{\lambda_2}\ldots p_{\lambda_\ell}$). Hence, the matrix which represents the endomorphism S_n of $(\Lambda_{\mathbb{Q}})_n$ with respect to the basis $(p_\lambda)_{\lambda\in\operatorname{Par}_n}$ of $(\Lambda_{\mathbb{Q}})_n$ is a diagonal matrix, whose λ-th diagonal entry (for any $\lambda\in\operatorname{Par}_n$) is $(-1)^{\ell(\lambda)}$. But the trace $\operatorname{trace}(S_n)$ of S_n (by its definition) is the sum of the diagonal entries of this matrix; hence, this trace equals

$$\operatorname{trace}(S_n) = \sum_{\lambda\in\operatorname{Par}_n} (-1)^{\ell(\lambda)} = \sum_{k=0}^{n} \underbrace{\sum_{\substack{\lambda\in\operatorname{Par}_n;\\ \ell(\lambda)=k}} (-1)^k}_{=(-1)^k(\text{number of all } \lambda\in\operatorname{Par}_n \text{ satisfying } \ell(\lambda)=k)}$$
$$\left(\text{since } \ell(\lambda)\in\{0,1,\ldots,n\} \text{ for every } \lambda\in\operatorname{Par}_n\right)$$
$$= \sum_{k=0}^{n} (-1)^k \underbrace{\left(\text{number of all } \lambda\in\operatorname{Par}_n \text{ satisfying } \ell(\lambda)=k\right)}_{=p(n,k)} = \sum_{k=0}^{n} (-1)^k p(n,k).$$

Compared with $\operatorname{trace}(S_n) = (-1)^n c(n)$, this yields $(-1)^n c(n) = \sum_{k=0}^{n} (-1)^k p(n,k)$. This solves part (a) of the exercise.

(b) Consider the map $s_1 s_1^\perp : \Lambda\to\Lambda$ which sends every $f\in\Lambda$ to $s_1 s_1^\perp f$. This map $s_1 s_1^\perp$ is graded (because the map $s_1^\perp : \Lambda\to\Lambda$ lowers the degree of any homogeneous element by 1, while the map s_1 raises it back by 1) and thus restricts to an endomorphism of the $\mathbb{Q}$-vector space $(\Lambda_{\mathbb{Q}})_n$. Denote this endomorphism by P_n. We want to compute $\operatorname{trace}(P_n)$. (This is well-defined since $(\Lambda_{\mathbb{Q}})_n$ is a finite-dimensional $\mathbb{Q}$-vector space.)

In the following, if $\mathcal{A}$ is any statement, then $[\mathcal{A}]$ will denote the truth value of $\mathcal{A}$ (that is, 1 if $\mathcal{A}$ holds, and 0 if it doesn't).

We know that the basis $(s_\lambda)_{\lambda\in\operatorname{Par}_n}$ of $(\Lambda_{\mathbb{Q}})_n$ is orthonormal with respect to the Hall inner product (restricted to $(\Lambda_{\mathbb{Q}})_n$). Hence, for any $\lambda\in\operatorname{Par}_n$, we have

$$\begin{aligned}
&\left(\text{the } s_\lambda\text{-coordinate of } P_n s_\lambda \text{ with respect to this basis}\right)\\
&=\left(\underbrace{P_n s_\lambda}_{=s_1 s_1^\perp s_\lambda}, s_\lambda\right) = \left(s_1 s_1^\perp s_\lambda, s_\lambda\right) = \left(s_1^\perp s_\lambda, s_1^\perp s_\lambda\right)\\
&=\sum_{\mu\in\operatorname{Par}_{n-1}} \left(\underbrace{\left(s_1^\perp s_\lambda, s_\mu\right)}_{=(s_\lambda, s_1 s_\mu)}\right)^2\\
&\qquad\left(\begin{array}{c}\text{since } s_1^\perp s_\lambda\in(\Lambda_{\mathbb{Q}})_{n-1}\text{, and since the basis } (s_\mu)_{\mu\in\operatorname{Par}_{n-1}}\\ \text{of } (\Lambda_{\mathbb{Q}})_{n-1} \text{ is orthonormal with respect to the Hall}\\ \text{inner product (restricted to } (\Lambda_{\mathbb{Q}})_{n-1}\text{)}\end{array}\right)\\
&=\sum_{\mu\in\operatorname{Par}_{n-1}} \left(s_\lambda, s_1 s_\mu\right)^2 . \qquad (13.78.1)
\end{aligned}$$

Now, every $\mu\in\operatorname{Par}_{n-1}$ satisfies

$$\underbrace{s_1}_{=h_1} s_\mu = h_1 s_\mu \overset{(2.7.1)}{=} \sum_{\substack{\mu^+ \,:\, \mu^+/\mu \text{ is a}\\ \text{horizontal 1-strip}}} s_{\mu^+} = \sum_{\mu^+ \,:\, |\mu^+/\mu|=1} s_{\mu^+} = \sum_{\substack{\mu^+\in\operatorname{Par}_n;\\ \mu\subseteq\mu^+}} s_{\mu^+}.$$

Hence, every $\lambda\in\operatorname{Par}_n$ and $\mu\in\operatorname{Par}_{n-1}$ satisfy

$$\begin{aligned}
\left(s_\lambda, s_1 s_\mu\right) &= \left(s_\lambda, \sum_{\substack{\mu^+\in\operatorname{Par}_n;\\ \mu\subseteq\mu^+}} s_{\mu^+}\right) = \sum_{\substack{\mu^+\in\operatorname{Par}_n;\\ \mu\subseteq\mu^+}} \underbrace{\left(s_\lambda, s_{\mu^+}\right)}_{\substack{=[\lambda=\mu^+]\\ \text{(since the Schur functions}\\ \text{are orthonormal with respect}\\ \text{to the Hall inner product)}}} = \sum_{\substack{\mu^+\in\operatorname{Par}_n;\\ \mu\subseteq\mu^+}} \left[\lambda=\mu^+\right]\\
&= [\mu\subseteq\lambda] .
\end{aligned}$$

Now, consider again the basis $(s_\lambda)_{\lambda\in\operatorname{Par}_n}$ of $(\Lambda_{\mathbb{Q}})_n$. For every $\lambda\in\operatorname{Par}_n$, we have

$$\begin{aligned}
&\left(\text{the } s_\lambda\text{-coordinate of } P_n s_\lambda \text{ with respect to this basis}\right)\\
&=\sum_{\mu\in\operatorname{Par}_{n-1}} \left(\underbrace{\left(s_\lambda, s_1 s_\mu\right)}_{=[\mu\subseteq\lambda]}\right)^2 \qquad \text{(by (13.78.1))}\\
&=\sum_{\mu\in\operatorname{Par}_{n-1}} \underbrace{[\mu\subseteq\lambda]^2}_{=[\mu\subseteq\lambda]} = \sum_{\mu\in\operatorname{Par}_{n-1}} [\mu\subseteq\lambda]\\
&=\left(\text{the number of all } \mu\in\operatorname{Par}_{n-1} \text{ such that } \mu\subseteq\lambda\right)\\
&=\left(\text{the number of all ways to remove a single cell from } \lambda \text{ to obtain a partition}\right)\\
&=\left(\text{the number of all corners of the Ferrers diagram of } \lambda\right)\\
&=C(\lambda) . \qquad (13.78.2)
\end{aligned}$$

Now, consider the matrix which represents the endomorphism P_n of $(\Lambda_{\mathbb{Q}})_n$ with respect to the basis $(s_\lambda)_{\lambda\in\operatorname{Par}_n}$ of $(\Lambda_{\mathbb{Q}})_n$. For any $\lambda\in\operatorname{Par}_n$ and $\mu\in\operatorname{Par}_n$, the (λ,μ)-th entry of this matrix is

$$\left(\text{the } s_\lambda\text{-coordinate of } P_n s_\mu \text{ with respect to this basis}\right).$$

The trace $\operatorname{trace}(P_n)$ of P_n (by its definition) is the sum of the diagonal entries of this matrix; hence, this trace equals

$$\operatorname{trace}(P_n) = \sum_{\lambda \in \operatorname{Par}_n} \underbrace{(\text{the } s_\lambda\text{-coordinate of } P_n s_\lambda \text{ with respect to this basis})}_{\substack{=C(\lambda) \\ (\text{by } (13.78.2))}} = \sum_{\lambda \in \operatorname{Par}_n} C(\lambda).$$

On the other hand, it is easy to see that

$$s_1^\perp (p_n) = [n = 1] \qquad \text{for every positive integer } n \tag{13.78.3}$$

(in fact, recall that p_n is primitive, so that $\Delta(p_n) = p_n \otimes 1 + 1 \otimes p_n$ and thus $s_1^\perp(p_n) = \underbrace{(s_1, p_n)}_{=[n=1]} 1 + \underbrace{(s_1, 1)}_{=0} p_n = [n = 1]$). But the map $s_1^\perp : \Lambda \to \Lambda$ is a derivation (by Proposition 2.8.2(iv), since s_1 is primitive). Hence, every partition $\lambda = (\lambda_1, \lambda_2, ..., \lambda_\ell)$ with $\ell = \ell(\lambda)$ satisfies

$$\begin{aligned} s_1^\perp (p_{\lambda_1} p_{\lambda_2} ... p_{\lambda_\ell}) &= \sum_{k=1}^{\ell} p_{\lambda_1} p_{\lambda_2} ... p_{\lambda_{k-1}} \underbrace{s_1^\perp (p_{\lambda_k})}_{\substack{=[\lambda_k=1] \\ (\text{by } (13.78.3))}} p_{\lambda_{k+1}} p_{\lambda_{k+2}} ... p_{\lambda_\ell} \\ &= \sum_{k=1}^{\ell} p_{\lambda_1} p_{\lambda_2} ... p_{\lambda_{k-1}} [\lambda_k = 1] p_{\lambda_{k+1}} p_{\lambda_{k+2}} ... p_{\lambda_\ell} \end{aligned} \tag{13.78.4}$$

and

$$\begin{aligned} P_n p_\lambda &= s_1 s_1^\perp \underbrace{p_\lambda}_{=p_{\lambda_1} p_{\lambda_2} ... p_{\lambda_\ell}} = \underbrace{s_1}_{=p_1} \underbrace{s_1^\perp (p_{\lambda_1} p_{\lambda_2} ... p_{\lambda_\ell})}_{\substack{=\sum_{k=1}^{\ell} p_{\lambda_1} p_{\lambda_2} ... p_{\lambda_{k-1}} [\lambda_k=1] p_{\lambda_{k+1}} p_{\lambda_{k+2}} ... p_{\lambda_\ell} \\ (\text{by } (13.78.4))}} \\ &= p_1 \sum_{k=1}^{\ell} p_{\lambda_1} p_{\lambda_2} ... p_{\lambda_{k-1}} [\lambda_k = 1] p_{\lambda_{k+1}} p_{\lambda_{k+2}} ... p_{\lambda_\ell} \\ &= \sum_{k=1}^{\ell} p_{\lambda_1} p_{\lambda_2} ... p_{\lambda_{k-1}} \underbrace{[\lambda_k = 1] p_1}_{\substack{=[\lambda_k=1] p_{\lambda_k} \\ (\text{because if } \lambda_k \neq 1, \text{ then both sides} \\ \text{of this are 0, and otherwise they} \\ \text{are clearly equal})}} p_{\lambda_{k+1}} p_{\lambda_{k+2}} ... p_{\lambda_\ell} \\ &= \sum_{k=1}^{\ell} p_{\lambda_1} p_{\lambda_2} ... p_{\lambda_{k-1}} [\lambda_k = 1] p_{\lambda_k} p_{\lambda_{k+1}} p_{\lambda_{k+2}} ... p_{\lambda_\ell} = \sum_{k=1}^{\ell} [\lambda_k = 1] p_{\lambda_1} p_{\lambda_2} ... p_{\lambda_\ell} \\ &= \underbrace{\left(\sum_{k=1}^{\ell} [\lambda_k = 1] \right)}_{\substack{=(\text{the number of } k \in \{1,2,...,\ell\} \text{ such that } \lambda_k=1) \\ =(\text{the number of parts of } \lambda \text{ equal to } 1) \\ =\mu_1(\lambda)}} \underbrace{p_{\lambda_1} p_{\lambda_2} ... p_{\lambda_\ell}}_{=p_\lambda} \\ &= \mu_1(\lambda) p_\lambda. \end{aligned}$$

Hence, the matrix which represents the endomorphism P_n of $(\Lambda_{\mathbb{Q}})_n$ with respect to the basis $(p_\lambda)_{\lambda \in \operatorname{Par}_n}$ of $(\Lambda_{\mathbb{Q}})_n$ is a diagonal matrix, whose λ-th diagonal entry (for any $\lambda \in \operatorname{Par}_n$) is $\mu_1(\lambda)$. But the trace $\operatorname{trace}(P_n)$ of P_n (by its definition) is the sum of the diagonal entries of this matrix; hence, this trace equals

$$\operatorname{trace}(P_n) = \sum_{\lambda \in \operatorname{Par}_n} \mu_1(\lambda).$$

Compared with $\operatorname{trace}(P_n) = \sum_{\lambda \in \operatorname{Par}_n} C(\lambda)$, this yields $\sum_{\lambda \in \operatorname{Par}_n} C(\lambda) = \sum_{\lambda \in \operatorname{Par}_n} \mu_1(\lambda)$. This solves part (b) of the exercise.

13.79. **Solution to Exercise 2.8.7.** *Solution to Exercise 2.8.7.* We assume WLOG that $\mathbf{k} = \mathbb{Z}$, since it is enough to prove what we want over $\mathbb{Z}$.

(a) This is exactly the equality (2.4.14), and has already been proved. Here is a slightly different way to prove it:

Use the primitivity of p_n and Proposition 1.4.17 to obtain $S(p_n) = -p_n$ for every $n \geq 1$.

But the antipode S of Λ is an algebra anti-endomorphism (by Proposition 1.4.10), therefore an algebra endomorphism (by Exercise 1.5.8(a), since Λ is commutative). Hence, from the fact that $S(p_n) = -p_n$ for every $n \geq 1$, we can deduce by multiplicativity that $S(p_\lambda) = (-1)^{\ell(\lambda)} p_\lambda$ for every partition λ. Now recall (2.4.11), and the claim of part (a) follows.

(b) For this part of the exercise, we shall work in $\Lambda_\mathbb{R}$ and prove that the endomorphism ω of $\Lambda_\mathbb{R}$ is an isometry. This will clearly yield the analogous statement over $\mathbb{Z}$.

Recall that $\left\{ \dfrac{p_\lambda}{\sqrt{z_\lambda}} \right\}$ is an orthonormal basis of $\Lambda_\mathbb{R}$ (by Corollary 2.5.17(c)). Hence, in order to prove that the endomorphism ω of $\Lambda_\mathbb{R}$ is an isometry, it is enough to show that

$$\left(\omega\left(\frac{p_\lambda}{\sqrt{z_\lambda}}\right), \omega\left(\frac{p_\mu}{\sqrt{z_\mu}}\right) \right) = \left(\frac{p_\lambda}{\sqrt{z_\lambda}}, \frac{p_\mu}{\sqrt{z_\mu}} \right)$$

for any two partitions λ and μ. But this follows from part (a) and the fact that the basis $\left\{ \dfrac{p_\lambda}{\sqrt{z_\lambda}} \right\}$ is orthonormal. Hence, we have shown that the endomorphism ω of $\Lambda_\mathbb{R}$ is an isometry. The same holds therefore for the endomorphism ω of Λ, and thus part (b) of the exercise is solved.

(c) Part (c) of the exercise is precisely the statement of Proposition 2.4.3(f), which was proved in the solution to Exercise 2.4.4. Thus, we need not prove it again.

(d) follows from (b) and (c): Indeed, ω is a coalgebra morphism (by part (c) of the exercise). Now, Definition 2.8.1 yields

$$(\omega(a))^\perp \omega(b) = \sum_{(\omega(b))} (\omega(a), (\omega(b))_1)(\omega(b))_2 = \sum_{(b)} (\omega(a), \omega(b_1))\,\omega(b_2)$$

(since ω is a coalgebra morphism and thus we have $\sum_{(\omega(b))} (\omega(b))_1 \otimes (\omega(b))_2 = \sum_{(b)} \omega(b_1) \otimes \omega(b_2)$). Since ω is an isometry, this further simplifies to

$$(\omega(a))^\perp \omega(b) = \sum_{(b)} (a, b_1)\,\omega(b_2) = \omega\left(\underbrace{\sum_{(b)} (a, b_1)\, b_2}_{=a^\perp b} \right) = \omega\left(a^\perp b\right),$$

whence part (d) of the exercise is solved.

(e) and (f) follow from the fact that $s_\mu^\perp(s_\lambda) = s_{\lambda/\mu}$ (applied, respectively, to $\mu = \left(1^\ell\right)$ and to $\mu = (\lambda_1)$), and the fact that a parallel translation of a skew Ferrers shape doesn't change the corresponding skew Schur function.

(g) It should be clear that $S\left(s_{\lambda/\mu}\right) = (-1)^{|\lambda/\mu|} s_{\lambda^t/\mu^t}$ follows from $\omega\left(s_{\lambda/\mu}\right) = s_{\lambda^t/\mu^t}$, and this, in turn, follows from (d) if one knows that $\omega(s_\lambda) = s_{\lambda^t}$ for every partition λ.

So it remains to prove that $\omega(s_\lambda) = s_{\lambda^t}$ for every partition λ. To do so, we use induction over $|\lambda|$, and fix λ. Part (b) yields that $(\omega(s_\lambda), \omega(s_\lambda)) = 1$. Since $\omega(s_\lambda) \in \Lambda$, this shows that $\omega(s_\lambda) = \pm s_\nu$ for some partition ν (since a length-1 integral vector cannot have more than one nonzero coordinate). Hence, $\omega(s_\nu) = \pm s_\lambda$ (since ω is an involution). Writing λ as $(\lambda_1, \lambda_2, \ldots, \lambda_\ell)$ with all of λ_1, λ_2, $\ldots$, λ_ℓ being positive, and writing ν as $(\nu_1, \nu_2, \nu_3, \ldots)$, we easily see (using (d) and (f)) that $\omega\left(e_{\nu_1}^\perp s_\lambda\right) = \pm h_{\nu_1}^\perp(s_\nu) = \pm s_{(\nu_2,\nu_3,\nu_4,\ldots)} \neq 0$, so that $e_{\nu_1}^\perp s_\lambda \neq 0$, and therefore $\ell \geq \nu_1$. On the other hand, (d) and (e) yield $\omega\left(h_\ell^\perp s_\nu\right) = \pm e_\ell^\perp(s_\lambda) = \pm s_{(\lambda_1-1,\lambda_2-1,\ldots,\lambda_\ell-1)} \neq 0$, thus $h_\ell^\perp s_\nu \neq 0$, so that $\ell \leq \nu_1$. Combined with $\ell \geq \nu_1$, this yields $\ell = \nu_1$. Hence, $h_\ell^\perp s_\nu = h_{\nu_1}^\perp s_{\nu_1} = s_{(\nu_2,\nu_3,\nu_4,\ldots)}$ (by (f)), so that

$$\omega\left(h_\ell^\perp s_\nu\right) = \omega\left(s_{(\nu_2,\nu_3,\nu_4,\ldots)}\right) = s_{(\nu_2,\nu_3,\nu_4,\ldots)^t}$$

by the induction assumption. Compared with

$$\omega\left(h_\ell^\perp s_\nu\right) = \pm e_\ell^\perp\left(s_\lambda\right) = \pm s_{(\lambda_1-1,\lambda_2-1,\ldots,\lambda_\ell-1)}$$

(a consequence of (d) and (e)), this yields $s_{(\nu_2,\nu_3,\nu_4,\ldots)^t} = \pm s_{(\lambda_1-1,\lambda_2-1,\ldots,\lambda_\ell-1)}$, from which it directly follows that $\pm = +$ and $(\nu_2,\nu_3,\nu_4,\ldots)^t = (\lambda_1-1,\lambda_2-1,\ldots,\lambda_\ell-1)$. This quickly yields $\nu = \lambda^t$, so that $\omega(s_\lambda) = \pm s_\nu$ becomes $\omega(s_\lambda) = s_{\lambda^t}$ (the $\pm$ in $\omega(s_\lambda) = \pm s_\nu$ is the same as in $s_{(\nu_2,\nu_3,\nu_4,\ldots)^t} = \pm s_{(\lambda_1-1,\lambda_2-1,\ldots,\lambda_\ell-1)}$), so we are done.

13.80. **Solution to Exercise 2.8.8.** *Solution to Exercise 2.8.8.* Proposition 2.3.6(i) yields $\Delta p_n = 1 \otimes p_n + p_n \otimes 1$. Comparing this with $\Delta p_n = \sum_{(p_n)} (p_n)_1 \otimes (p_n)_2$ (here, we are using the Sweedler notation), we obtain

$$\sum_{(p_n)} (p_n)_1 \otimes (p_n)_2 = 1 \otimes p_n + p_n \otimes 1. \tag{13.80.1}$$

(a) Proposition 2.4.1(i) yields $S(p_n) = -p_n$.

But $p_n \in \Lambda_n$. Hence, Exercise 2.7.14(b) (applied to $f = p_n$) yields

$$\begin{aligned}(e_n, p_n) &= (-1)^n \cdot \underbrace{(S(p_n))}_{=-p_n}(1) = (-1)^n \cdot \underbrace{(-p_n)(1)}_{=-p_n(1)} \\ &= (-1)^n \cdot (-p_n(1)) = \underbrace{-(-1)^n}_{\substack{=(-1)^{n+1}=(-1)^{n-1} \\ (\text{since } n+1\equiv n-1 \bmod 2)}} \cdot \underbrace{p_n(1)}_{\substack{=1^n \\ (\text{by the definition of } p_n)}} \\ &= (-1)^{n-1} \cdot \underbrace{1^n}_{=1} = (-1)^{n-1}.\end{aligned}$$

This solves Exercise 2.8.8(a).

(b) Let $m \in \mathbb{N}$ satisfy $m \neq n$. Then, $e_m \in \Lambda_m$ and $p_n \in \Lambda_n$. But m and n are two distinct nonnegative integers (since $m \neq n$). Hence, Exercise 2.5.13(a) (applied to m, n, e_m and p_n instead of n, m, f and g) yields $(e_m, p_n) = 0$. This solves Exercise 2.8.8(b).

(c) We have $n \neq 0$ (since n is positive). Thus, n and 0 are two distinct nonnegative integers. Hence, Exercise 2.5.13(a) (applied to $m = 0$, $f = e_n$ and $g = 1$) yields $(e_n, 1) = 0$ (since $e_n \in \Lambda_n$ and $1 \in \Lambda_0$).

By the definition of $e_n^\perp p_n$, we have

$$\begin{aligned}e_n^\perp p_n &= \sum_{(p_n)} (e_n, (p_n)_1) \cdot (p_n)_2 = \underbrace{(e_n, 1)}_{=0} \cdot p_n + \underbrace{(e_n, p_n)}_{\substack{=(-1)^{n-1} \\ (\text{by Exercise 2.8.8(a)})}} \cdot 1 \\ &\qquad \left(\text{since } \sum_{(p_n)} (p_n)_1 \otimes (p_n)_2 = 1 \otimes p_n + p_n \otimes 1 \text{ (by (13.80.1))}\right) \\ &= \underbrace{0 \cdot p_n}_{=0} + (-1)^{n-1} = (-1)^{n-1}.\end{aligned}$$

This solves Exercise 2.8.8(c).

(d) Let m be a positive integer satisfying $m \neq n$. Then, Exercise 2.8.8(b) yields $(e_m, p_n) = 0$.

We have $m \neq 0$ (since m is positive). Thus, m and 0 are two distinct nonnegative integers. Hence, Exercise 2.5.13(a) (applied to m, 0, e_m and 1 instead of n, m, f and g) yields $(e_m, 1) = 0$ (since $e_m \in \Lambda_m$ and $1 \in \Lambda_0$).

By the definition of $e_m^\perp p_n$, we have

$$\begin{aligned} e_m^\perp p_n &= \sum_{(p_n)} \left(e_m, (p_n)_1\right) \cdot (p_n)_2 = \underbrace{(e_m, 1)}_{=0} \cdot p_n + \underbrace{(e_m, p_n)}_{=0} \cdot 1 \\ &\qquad \left(\text{since } \sum_{(p_n)} (p_n)_1 \otimes (p_n)_2 = 1 \otimes p_n + p_n \otimes 1 \text{ (by (13.80.1))}\right) \\ &= \underbrace{0 \cdot p_n}_{=0} + \underbrace{0 \cdot 1}_{=0} = 0. \end{aligned}$$

This solves Exercise 2.8.8(d).

13.81. **Solution to Exercise 2.9.1.** *Solution to Exercise 2.9.1.* (a) The claim that the sum $\sum_{i \in \mathbb{N}} (-1)^i h_{m+i} e_i^\perp f$ is convergent is very easy to see: Any given $f \in \Lambda = \bigoplus_{n \in \mathbb{N}} \Lambda_n$ lives in a finite direct sum $\bigoplus_{n=0}^m \Lambda_n \subset \Lambda$; if we take $i \in \mathbb{N}$ higher than m, then $e_i^\perp f = 0$ for degree reasons[642] and therefore $(-1)^i h_{m+i} e_i^\perp f = 0$.

The claim that the map $\mathbf{B}_m$ is $\mathbf{k}$-linear is obvious. Exercise 2.9.1(a) is solved.

(c) Exercise 2.9.1(c) makes three claims. Let us first prove the first of them: the identity (2.9.1).

Proof of (2.9.1): Let $\lambda = (\lambda_1, \lambda_2, \lambda_3, ...)$ be a partition having at most n parts. Then,

$$\begin{aligned} s_\lambda = s_{\lambda / \varnothing} &\overset{(2.4.16)}{=} \det\left(\left(h_{\lambda_i - \varnothing_j - i + j}\right)_{i,j=1,2,\ldots,n} \right) = \det\left(\left(h_{\lambda_i - i + j}\right)_{i,j=1,2,\ldots,n} \right) \qquad (\text{since } \varnothing_j = 0) \\ &= \overline{s}_{(\lambda_1, \lambda_2, \ldots, \lambda_n)} \qquad \left(\text{by the definition of } \overline{s}_{(\lambda_1, \lambda_2, \ldots, \lambda_n)}\right). \end{aligned}$$

This proves (2.9.1).

Next, we will show that for every n-tuple $(\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{Z}^n$, the symmetric function $\overline{s}_{(\alpha_1, \alpha_2, \ldots, \alpha_n)}$ either is 0 or equals $\pm s_\nu$ for some partition ν having at most n parts.

Proof. Let $(\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{Z}^n$ be any n-tuple. The definition of $\overline{s}_{(\alpha_1, \alpha_2, \ldots, \alpha_n)}$ yields

$$\overline{s}_{(\alpha_1, \alpha_2, \ldots, \alpha_n)} = \det\left(\left(h_{\alpha_i - i + j}\right)_{i,j=1,2,\ldots,n} \right).$$

If the n integers $\alpha_1 - 1$, $\alpha_2 - 2$, ..., $\alpha_n - n$ are not distinct, then the matrix $\left(h_{\alpha_i - i + j}\right)_{i,j=1,2,\ldots,n}$ has two equal rows and thus its determinant is 0, so that $\overline{s}_{(\alpha_1, \alpha_2, \ldots, \alpha_n)} = \det\left(\left(h_{\alpha_i - i + j}\right)_{i,j=1,2,\ldots,n} \right) = 0$ in this case. So the claim is proven in this case, and this is a case we do not need to address anymore. Thus, assume WLOG that the n integers $\alpha_1 - 1$, $\alpha_2 - 2$, ..., $\alpha_n - n$ are distinct. Hence, there exists a (unique) permutation $\tau \in \mathfrak{S}_n$ satisfying $\alpha_{\tau(1)} - \tau(1) > \alpha_{\tau(2)} - \tau(2) > ... > \alpha_{\tau(n)} - \tau(n)$. Consider this τ.

Define an n-tuple $(\gamma_1, \gamma_2, ..., \gamma_n) \in \mathbb{Z}^n$ by setting

$$\gamma_i = \alpha_{\tau(i)} - \tau(i) + i \qquad \text{for every } i \in \{1, 2, ..., n\}.$$

Then, it is easy to see that $\gamma_1 \geq \gamma_2 \geq ... \geq \gamma_n$. [643] Moreover, every $i \in \{1, 2, ..., n\}$ satisfies $\gamma_i - i = \alpha_{\tau(i)} - \tau(i)$ (by the definition of γ_i), and so the matrix $\left(h_{\gamma_i - i + j}\right)_{i,j=1,2,\ldots,n}$ is obtained from the matrix $\left(h_{\alpha_i - i + j}\right)_{i,j=1,2,\ldots,n}$ by permuting its rows according to the permutation τ. Since permuting the rows of a matrix multiplies the determinant of the matrix by the sign of the permutation, we thus see

$$\det\left(\left(h_{\gamma_i - i + j}\right)_{i,j=1,2,\ldots,n} \right) = (-1)^\tau \underbrace{\det\left(\left(h_{\alpha_i - i + j}\right)_{i,j=1,2,\ldots,n} \right)}_{= \overline{s}_{(\alpha_1, \alpha_2, \ldots, \alpha_n)}} = (-1)^\tau \overline{s}_{(\alpha_1, \alpha_2, \ldots, \alpha_n)},$$

so that

$$\overline{s}_{(\alpha_1, \alpha_2, \ldots, \alpha_n)} = (-1)^\tau \det\left(\left(h_{\gamma_i - i + j}\right)_{i,j=1,2,\ldots,n} \right). \tag{13.81.1}$$

If $\gamma_n < 0$, then the matrix $\left(h_{\gamma_i - i + j}\right)_{i,j=1,2,\ldots,n}$ on the right hand side of this equality has its n-th row consist of zeroes only, which implies that the determinant of this matrix is 0, and thus (13.81.1) becomes

[642] In fact, the map $e_i^\perp$ lowers degree by i, and thus annihilates $\bigoplus_{n=0}^m \Lambda_n$ when $i > m$.

[643] To prove this, it is enough to show that $\gamma_i \geq \gamma_{i+1}$ for every $i \in \{1, 2, ..., n-1\}$. But this simplifies to $\alpha_{\tau(i)} - \tau(i) \geq \alpha_{\tau(i+1)} - \tau(i+1) + 1$ (upon adding i), and this follows from $\alpha_{\tau(i)} - \tau(i) > \alpha_{\tau(i+1)} - \tau(i+1)$.

$\overline{s}_{(\alpha_1,\alpha_2,\ldots,\alpha_n)} = 0$. Hence, we are done in the case $\gamma_n < 0$. So assume WLOG that we don't have $\gamma_n < 0$. Thus, $\gamma_n \geq 0$. Combined with $\gamma_1 \geq \gamma_2 \geq \ldots \geq \gamma_n$, this yields that $(\gamma_1, \gamma_2, \ldots, \gamma_n)$ is a partition. This partition satisfies

$$s_{(\gamma_1,\gamma_2,\ldots,\gamma_n)} = s_{(\gamma_1,\gamma_2,\ldots,\gamma_n)/\varnothing} \overset{(2.4.16)}{=} \det\left(\left(h_{\gamma_i-\varnothing_j-i+j}\right)_{i,j=1,2,\ldots,n}\right) = \det\left(\left(h_{\gamma_i-i+j}\right)_{i,j=1,2,\ldots,n}\right)$$

(since $\varnothing_j = 0$ for all j). Thus, (13.81.1) becomes

$$\overline{s}_{(\alpha_1,\alpha_2,\ldots,\alpha_n)} = (-1)^{\tau} \underbrace{\det\left(\left(h_{\gamma_i-i+j}\right)_{i,j=1,2,\ldots,n}\right)}_{=s_{(\gamma_1,\gamma_2,\ldots,\gamma_n)}} = (-1)^{\tau} s_{(\gamma_1,\gamma_2,\ldots,\gamma_n)},$$

and therefore the symmetric function $\overline{s}_{(\alpha_1,\alpha_2,\ldots,\alpha_n)}$ equals $\pm s_\nu$ for some partition ν having at most n parts (namely, the partition ν is $(\gamma_1, \gamma_2, \ldots, \gamma_n)$, and the $\pm$ sign is $(-1)^\tau$).

We thus have shown that for every n-tuple $(\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{Z}^n$, the symmetric function $\overline{s}_{(\alpha_1,\alpha_2,\ldots,\alpha_n)}$ either is 0 or equals $\pm s_\nu$ for some partition ν having at most n parts.

In order to complete the solution of Exercise 2.9.1(c), it now remains to prove (2.9.2).

Proof of (2.9.2): Let $(\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{Z}^n$ and $(\beta_1, \beta_2, \ldots, \beta_n) \in \mathbb{N}^n$ be two n-tuples. We need to prove that (2.9.2) holds.

The definition of $\overline{s}_{(\alpha_1,\alpha_2,\ldots,\alpha_n)}$ yields

$$\overline{s}_{(\alpha_1,\alpha_2,\ldots,\alpha_n)} = \det\left(\left(h_{\alpha_i-i+j}\right)_{i,j=1,2,\ldots,n}\right).$$

If the n integers $\alpha_1 - 1$, $\alpha_2 - 2$, ..., $\alpha_n - n$ are not distinct, then (2.9.2) holds (because in this case, the matrices $\left(h_{\alpha_i-i+j}\right)_{i,j=1,2,\ldots,n}$ and $\left(h_{\alpha_i-\beta_j-i+j}\right)_{i,j=1,2,\ldots,n}$ have two equal rows each, so their determinants vanish, rendering both sides of (2.9.2) equal to zero[644]). Thus, we WLOG assume that the n integers $\alpha_1 - 1$, $\alpha_2 - 2$, ..., $\alpha_n - n$ are distinct. Thus, there exists a (unique) permutation $\tau \in \mathfrak{S}_n$ satisfying $\alpha_{\tau(1)} - \tau(1) > \alpha_{\tau(2)} - \tau(2) > \ldots > \alpha_{\tau(n)} - \tau(n)$. Consider this τ, and define an n-tuple $(\gamma_1, \gamma_2, \ldots, \gamma_n) \in \mathbb{Z}^n$ by setting

$$\gamma_i = \alpha_{\tau(i)} - \tau(i) + i \qquad \text{for every } i \in \{1, 2, \ldots, n\}.$$

Then, it is easy to see that $\gamma_1 \geq \gamma_2 \geq \ldots \geq \gamma_n$. Moreover, the matrix $\left(h_{\gamma_i-i+j}\right)_{i,j=1,2,\ldots,n}$ is obtained from the matrix $\left(h_{\alpha_i-i+j}\right)_{i,j=1,2,\ldots,n}$ by permuting its rows according to the permutation τ (because every $i \in \{1, 2, \ldots, n\}$ satisfies $\gamma_i - i = \alpha_{\tau(i)} - \tau(i)$). This leads to (13.81.1) again (as in the previous proof), so that

$$\overline{s}_{(\alpha_1,\alpha_2,\ldots,\alpha_n)} = (-1)^{\tau} \underbrace{\det\left(\left(h_{\gamma_i-i+j}\right)_{i,j=1,2,\ldots,n}\right)}_{=\overline{s}_{(\gamma_1,\gamma_2,\ldots,\gamma_n)}} = (-1)^{\tau} \overline{s}_{(\gamma_1,\gamma_2,\ldots,\gamma_n)}.$$

This equality, along with

$$\det\left(\left(h_{\alpha_i-\beta_j-i+j}\right)_{i,j=1,2,\ldots,n}\right) = (-1)^{\tau} \det\left(\left(h_{\gamma_i-\beta_j-i+j}\right)_{i,j=1,2,\ldots,n}\right)$$

(which, again, follows from the fact that the matrix on the right hand side is obtained from the matrix on the left hand side by permuting the rows according to τ), shows that in order to prove (2.9.2), it is enough to show that

$$\overline{s}^{\perp}_{(\beta_1,\beta_2,\ldots,\beta_n)} \overline{s}_{(\gamma_1,\gamma_2,\ldots,\gamma_n)} = \det\left(\left(h_{\gamma_i-\beta_j-i+j}\right)_{i,j=1,2,\ldots,n}\right).$$

But this is the same equality as (2.9.2), except with $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ replaced by $(\gamma_1, \gamma_2, \ldots, \gamma_n)$. The advantage of $(\gamma_1, \gamma_2, \ldots, \gamma_n)$ over $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ is that we know that $\gamma_1 \geq \gamma_2 \geq \ldots \geq \gamma_n$ (while the analogous chain of inequalities $\alpha_1 \geq \alpha_2 \geq \ldots \geq \alpha_n$ does not necessarily hold). So, we can **WLOG assume that** $\alpha_1 \geq \alpha_2 \geq \ldots \geq \alpha_n$ (because otherwise, we can replace $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ by $(\gamma_1, \gamma_2, \ldots, \gamma_n)$). Assume this.

If $\alpha_n < 0$, then both matrices $\left(h_{\alpha_i-i+j}\right)_{i,j=1,2,\ldots,n}$ and $\left(h_{\alpha_i-\beta_j-i+j}\right)_{i,j=1,2,\ldots,n}$ have their n-th row consisting of only zeroes, and so their determinants both vanish, which shows that both sides of the equality (2.9.2) are zero[645]. So this is a case in which (2.9.2) trivially holds. We thus assume WLOG that we are

[644] The left hand side is affected because of $\overline{s}_{(\alpha_1,\alpha_2,\ldots,\alpha_n)} = \det\left(\left(h_{\alpha_i-i+j}\right)_{i,j=1,2,\ldots,n}\right)$.

[645] For the left hand side, this is because of $\overline{s}_{(\alpha_1,\alpha_2,\ldots,\alpha_n)} = \det\left(\left(h_{\alpha_i-i+j}\right)_{i,j=1,2,\ldots,n}\right)$.

not in this case; hence, we don't have $\alpha_n < 0$, so we have $\alpha_n \geq 0$. Combined with $\alpha_1 \geq \alpha_2 \geq ... \geq \alpha_n$, this yields that $(\alpha_1, \alpha_2, ..., \alpha_n)$ is a partition.

We have

$$\overline{s}_{(\beta_1,\beta_2,...,\beta_n)} = \det\left(\left(h_{\beta_i-i+j}\right)_{i,j=1,2,...,n}\right).$$

If the n integers $\beta_1 - 1$, $\beta_2 - 2$, ..., $\beta_n - n$ are not distinct, then (2.9.2) holds (because in this case, the matrix $\left(h_{\beta_i-i+j}\right)_{i,j=1,2,...,n}$ has two equal rows while the matrix $\left(h_{\alpha_i-\beta_j-i+j}\right)_{i,j=1,2,...,n}$ has two equal columns; thus, both of these matrices have determinant 0, so that both sides of (2.9.2) equal to zero[646]). Thus, we WLOG assume that the n integers $\beta_1 - 1$, $\beta_2 - 2$, ..., $\beta_n - n$ are distinct. Thus, there exists a (unique) permutation $\zeta \in \mathfrak{S}_n$ satisfying $\beta_{\zeta(1)} - \zeta(1) > \beta_{\zeta(2)} - \zeta(2) > ... > \beta_{\zeta(n)} - \zeta(n)$. Consider this ζ.

Define an n-tuple $(\delta_1, \delta_2, ..., \delta_n) \in \mathbb{Z}^n$ by setting

$$\delta_j = \beta_{\zeta(j)} - \zeta(j) + j \qquad \text{for every } j \in \{1, 2, ..., n\}.$$

Then, it is easy to see that $\delta_1 \geq \delta_2 \geq ... \geq \delta_n \geq 0$ [647]. Thus, $(\delta_1, \delta_2, ..., \delta_n)$ is a partition.

The matrix $\left(h_{\delta_i-i+j}\right)_{i,j=1,2,...,n}$ is obtained from the matrix $\left(h_{\beta_i-i+j}\right)_{i,j=1,2,...,n}$ by permuting its rows according to the permutation ζ (since the definition of δ_i yields $\delta_i = \beta_{\zeta(i)} - \zeta(i) + i$, thus $\delta_i - i = \beta_{\zeta(i)} - \zeta(i)$ for every $i \in \{1, 2, ..., n\}$). This leads to $\det\left(\left(h_{\delta_i-i+j}\right)_{i,j=1,2,...,n}\right) = (-1)^\zeta \underbrace{\det\left(\left(h_{\beta_i-i+j}\right)_{i,j=1,2,...,n}\right)}_{=\overline{s}_{(\beta_1,\beta_2,...,\beta_n)}} = (-1)^\zeta \overline{s}_{(\beta_1,\beta_2,...,\beta_n)}$ and thus

$$\overline{s}_{(\beta_1,\beta_2,...,\beta_n)} = (-1)^\zeta \underbrace{\det\left(\left(h_{\delta_i-i+j}\right)_{i,j=1,2,...,n}\right)}_{=\overline{s}_{(\delta_1,\delta_2,...,\delta_n)}} = (-1)^\zeta \overline{s}_{(\delta_1,\delta_2,...,\delta_n)}.$$

This equality, and the equality

$$\det\left(\left(h_{\alpha_i-\beta_j-i+j}\right)_{i,j=1,2,...,n}\right) = (-1)^\zeta \det\left(\left(h_{\alpha_i-\delta_j-i+j}\right)_{i,j=1,2,...,n}\right)$$

(this time because the matrix on the right hand side is obtained from that on the left hand side by permuting its **columns** according to ζ) show that in order to prove (2.9.2), it is enough to show that

$$\overline{s}^\perp_{(\delta_1,\delta_2,...,\delta_n)} \overline{s}_{(\alpha_1,\alpha_2,...,\alpha_n)} = \det\left(\left(h_{\alpha_i-\delta_j-i+j}\right)_{i,j=1,2,...,n}\right).$$

This is, of course, the same equality as (2.9.2), except with $(\beta_1, \beta_2, \ldots, \beta_n)$ replaced by $(\delta_1, \delta_2, ..., \delta_n)$. The advantage of $(\delta_1, \delta_2, ..., \delta_n)$ over $(\beta_1, \beta_2, \ldots, \beta_n)$ is that we know that $(\delta_1, \delta_2, ..., \delta_n)$ is a partition. So, we can **WLOG assume that** $(\beta_1, \beta_2, ..., \beta_n)$ is a partition (because otherwise, we can replace $(\beta_1, \beta_2, \ldots, \beta_n)$ by $(\delta_1, \delta_2, ..., \delta_n)$). Assume this.

Through a series of WLOG assumptions, we have now ensured that both $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ and $(\beta_1, \beta_2, \ldots, \beta_n)$ are partitions. Thus,

$$\begin{aligned} s_{(\alpha_1,\alpha_2,...,\alpha_n)} &= s_{(\alpha_1,\alpha_2,...,\alpha_n)/\varnothing} \overset{(2.4.16)}{=} \det\left(\left(h_{\alpha_i-\varnothing_j-i+j}\right)_{i,j=1,2,...,n}\right) \\ &= \det\left(\left(h_{\alpha_i-i+j}\right)_{i,j=1,2,...,n}\right) \qquad (\text{since } \varnothing_j = 0) \\ &= \overline{s}_{(\alpha_1,\alpha_2,...,\alpha_n)}, \end{aligned}$$

so that $\overline{s}_{(\alpha_1,\alpha_2,...,\alpha_n)} = s_{(\alpha_1,\alpha_2,...,\alpha_n)}$. Similarly $\overline{s}_{(\beta_1,\beta_2,...,\beta_n)} = s_{(\beta_1,\beta_2,...,\beta_n)}$. Hence,

$$\begin{aligned} \overline{s}^\perp_{(\beta_1,\beta_2,...,\beta_n)} \overline{s}_{(\alpha_1,\alpha_2,...,\alpha_n)} &= s^\perp_{(\beta_1,\beta_2,...,\beta_n)} s_{(\alpha_1,\alpha_2,...,\alpha_n)} = s_{(\alpha_1,\alpha_2,...,\alpha_n)/(\beta_1,\beta_2,...,\beta_n)} \\ &\overset{(2.4.16)}{=} \det\left(\left(h_{\alpha_i-\beta_j-i+j}\right)_{i,j=1,2,...,n}\right). \end{aligned}$$

[646] This time, the reason why this causes the left hand side to be zero is the identity $\overline{s}_{(\beta_1,\beta_2,...,\beta_n)} = \det\left(\left(h_{\beta_i-i+j}\right)_{i,j=1,2,...,n}\right)$.

[647] The $\delta_1 \geq \delta_2 \geq ... \geq \delta_n$ part here is proven like the similar inequalities $\gamma_1 \geq \gamma_2 \geq ... \geq \gamma_n$ shown above; but the $\delta_n \geq 0$ part might require explanation. It stems from the fact that $\delta_n = \beta_{\zeta(n)} - \underbrace{\zeta(n)}_{\leq n} + n \geq \beta_{\zeta(n)} - n + n = \beta_{\zeta(n)} \geq 0$ (because we have assumed $(\beta_1, \beta_2, ..., \beta_n) \in \mathbb{N}^n$).

This proves (2.9.2) and thus completes the proof of Exercise 2.9.1(c).

(d) In the following, we use the so-called *Iverson bracket notation*: For every assertion $\mathcal{A}$, we let $[\mathcal{A}]$ denote the integer $\begin{cases} 1, & \text{if } \mathcal{A} \text{ is true;} \\ 0, & \text{if } \mathcal{A} \text{ is false} \end{cases}$. (This integer is called the *truth value* of $\mathcal{A}$.)

Before we come to the solution of Exercise 2.9.1(d), let us recall a general fact about determinants:

- Every commutative ring A, every positive integer N and every matrix $(\beta_{i,j})_{i,j=1,2,\ldots,N} \in A^{N\times N}$ satisfy

$$\det\left((\beta_{i,j})_{i,j=1,2,\ldots,N}\right) = \sum_{k=1}^{N} (-1)^{k-1} \beta_{1,k} \det\left(\left(\beta_{i+1,j+[j\geq k]}\right)_{i,j=1,2,\ldots,N-1}\right). \tag{13.81.2}$$

(This is just one possible way to write the Laplace expansion formula for the expansion of the determinant of a matrix with respect to its first row.)

Let us now solve Exercise 2.9.1(d). Let $n \in \mathbb{N}$, let $m \in \mathbb{Z}$ and let $(\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{Z}^n$. We must prove (2.9.3).

Define an $(n+1)$-tuple $(\gamma_1, \gamma_2, ..., \gamma_{n+1}) \in \mathbb{Z}^{n+1}$ by $(\gamma_1, \gamma_2, ..., \gamma_{n+1}) = (m, \alpha_1, \alpha_2, ..., \alpha_n)$. Then, $\gamma_1 = m$, whereas every $i \in \{1, 2, ..., n\}$ satisfies $\gamma_{i+1} = \alpha_i$. Since $(m, \alpha_1, \alpha_2, ..., \alpha_n) = (\gamma_1, \gamma_2, ..., \gamma_{n+1})$, we have

$$\begin{aligned}
\overline{s}_{(m,\alpha_1,\alpha_2,...,\alpha_n)} &= \overline{s}_{(\gamma_1,\gamma_2,...,\gamma_{n+1})} = \det\left(\left(h_{\gamma_i - i + j}\right)_{i,j=1,2,...,n+1}\right) \qquad \left(\text{by the definition of } \overline{s}_{(\gamma_1,\gamma_2,...,\gamma_{n+1})}\right) \\
&= \sum_{k=1}^{n+1} (-1)^{k-1} h_{\gamma_1 - 1 + k} \det\left(\left(h_{\gamma_{i+1} - (i+1) + (j + [j\geq k])}\right)_{i,j=1,2,...,n}\right) \\
&\qquad \left(\text{by (13.81.2), applied to } A = \Lambda,\ N = n+1 \text{ and } \beta_{i,j} = h_{\gamma_i - i + j}\right) \\
&= \sum_{k=0}^{n} \underbrace{(-1)^{(k+1)-1}}_{=(-1)^k} \underbrace{h_{\gamma_1 - 1 + (k+1)}}_{\substack{=h_{m+k} \\ (\text{since } \gamma_1 - 1 + (k+1) = \gamma_1 + k = m + k \\ (\text{because } \gamma_1 = m))}} \det\left(\left(\underbrace{h_{\gamma_{i+1} - (i+1) + (j + [j\geq k+1])}}_{\substack{=h_{\alpha_i - (i+1) + (j + [j\geq k+1])} \\ (\text{since } \gamma_{i+1} = \alpha_i)}}\right)_{i,j=1,2,...,n}\right) \\
&\qquad (\text{here, we have substituted } k+1 \text{ for } k \text{ in the sum}) \\
&= \sum_{k=0}^{n} (-1)^k h_{m+k} \det\left(\left(h_{\alpha_i - (i+1) + (j + [j\geq k+1])}\right)_{i,j=1,2,...,n}\right). \qquad (13.81.3)
\end{aligned}$$

Now, let us check that

$$\det\left(\left(h_{\alpha_i - (i+1) + (j + [j\geq k+1])}\right)_{i,j=1,2,...,n}\right) = e_k^{\perp} \overline{s}_{(\alpha_1,\alpha_2,...,\alpha_n)} \qquad \text{for every } k \in \{0, 1, ..., n\}. \tag{13.81.4}$$

Proof of (13.81.4): Let $k \in \{0, 1, ..., n\}$. The partition (1^k) has length $k \leq n$, and thus can be identified with the n-tuple $\left(\underbrace{1, 1, ..., 1}_{k \text{ times}}, \underbrace{0, 0, ..., 0}_{n-k \text{ times}}\right) = (\beta_1, \beta_2, \ldots, \beta_n)$, where we set $\beta_j = \begin{cases} 1, & \text{if } j \leq k; \\ 0, & \text{if } j > k \end{cases} = [j \leq k] = 1 - [j \geq k+1]$ for every $j \in \{1, 2, ..., n\}$. Thus, $s_{(1^k)} = s_{(\beta_1,\beta_2,...,\beta_n)} = \overline{s}_{(\beta_1,\beta_2,...,\beta_n)}$ (by (2.9.1), applied to $\lambda = (\beta_1, \beta_2, \ldots, \beta_n)$). Hence, $e_k = s_{(1^k)} = \overline{s}_{(\beta_1,\beta_2,...,\beta_n)}$, so that

$$\begin{aligned}
e_k^{\perp} \overline{s}_{(\alpha_1,\alpha_2,...,\alpha_n)} &= \overline{s}^{\perp}_{(\beta_1,\beta_2,...,\beta_n)} \overline{s}_{(\alpha_1,\alpha_2,...,\alpha_n)} = \det\left(\left(h_{\alpha_i - \beta_j - i + j}\right)_{i,j=1,2,...,n}\right) \qquad (\text{by (2.9.2)}) \\
&= \det\left(\left(h_{\alpha_i - (1 - [j\geq k+1]) - i + j}\right)_{i,j=1,2,...,n}\right) \qquad \left(\text{since } \beta_j = 1 - [j \geq k+1] \text{ for every } j \in \{1, 2, ..., n\}\right) \\
&= \det\left(\left(h_{\alpha_i - (i+1) + (j + [j\geq k+1])}\right)_{i,j=1,2,...,n}\right)
\end{aligned}$$

(since $\alpha_i - (1 - [j \geq k+1]) - i + j = \alpha_i - (i+1) + (j + [j \geq k+1])$ for all $(i, j) \in \{1, 2, ..., n\}^2$). This proves (13.81.4).

Now, (13.81.3) becomes

$$\overline{s}_{(m,\alpha_1,\alpha_2,\ldots,\alpha_n)} = \sum_{k=0}^{n} (-1)^k h_{m+k} \underbrace{\det\left(\left(h_{\alpha_i-(i+1)+(j+[j\geq k+1])}\right)_{i,j=1,2,\ldots,n}\right)}_{\substack{=e_k^\perp \overline{s}_{(\alpha_1,\alpha_2,\ldots,\alpha_n)} \\ \text{(by (13.81.4))}}}$$

$$= \sum_{k=0}^{n} (-1)^k h_{m+k} e_k^\perp \overline{s}_{(\alpha_1,\alpha_2,\ldots,\alpha_n)} = \sum_{i=0}^{n} (-1)^i h_{m+i} e_i^\perp \overline{s}_{(\alpha_1,\alpha_2,\ldots,\alpha_n)} \tag{13.81.5}$$

(here, we renamed the summation index k as i).

But in order to solve Exercise 2.9.1(d), we need to prove a very similar yet different formula:

$$\overline{s}_{(m,\alpha_1,\alpha_2,\ldots,\alpha_n)} = \sum_{i\in\mathbb{N}} (-1)^i h_{m+i} e_i^\perp \overline{s}_{(\alpha_1,\alpha_2,\ldots,\alpha_n)}. \tag{13.81.6}$$

These two formulas differ in a minor detail: The sum on the right hand side of (13.81.6) runs over all $i \in \mathbb{N}$, whereas the sum on the right hand side of (13.81.5) only runs over $i \in \{0, 1, ..., n\}$. If we can show that these two sums are equal, then it will follow that the equality (13.81.6) that we are proving and the equality (13.81.5) that we have proven are equivalent, and so the former equality must hold, and Exercise 2.9.1(d) will be solved.

So we need to prove that the sum on the right hand side of (13.81.6) and the sum on the right hand side of (13.81.5) are equal. To achieve this, it clearly suffices to show that all addends in which these sums differ are 0. But these addends are the $(-1)^i h_{m+i} e_i^\perp \overline{s}_{(\alpha_1,\alpha_2,\ldots,\alpha_n)}$ for $i \in \mathbb{N}$ satisfying $i > n$. So we need to prove that $(-1)^i h_{m+i} e_i^\perp \overline{s}_{(\alpha_1,\alpha_2,\ldots,\alpha_n)} = 0$ for all $i \in \mathbb{N}$ satisfying $i > n$. Let us do this now.

Let $i \in \mathbb{N}$ be such that $i > n$. We need to show that $(-1)^i h_{m+i} e_i^\perp \overline{s}_{(\alpha_1,\alpha_2,\ldots,\alpha_n)} = 0$.

The second statement of Exercise 2.9.1(c) says that the symmetric function $\overline{s}_{(\alpha_1,\alpha_2,\ldots,\alpha_n)}$ either is 0 or equals $\pm s_\nu$ for some partition ν having at most n parts. If it is 0, then $(-1)^i h_{m+i} e_i^\perp \underbrace{\overline{s}_{(\alpha_1,\alpha_2,\ldots,\alpha_n)}}_{=0} = 0$ is obvious, and so we are done. Thus, we can WLOG assume that we are in the other case, i.e., the function $\overline{s}_{(\alpha_1,\alpha_2,\ldots,\alpha_n)}$ equals $\pm s_\nu$ for some partition ν having at most n parts. Assume that we are in this case, and consider this ν. Since ν has at most n parts, we have $\ell(\nu) \leq n$. Now, $\ell\left(\left(1^i\right)\right) = i > n \geq \ell(\nu)$, and therefore $\left(1^i\right) \not\subseteq \nu$. But $e_i = s_{(1^i)}$ and thus $e_i^\perp s_\nu = s_{(1^i)}^\perp s_\nu = s_{\nu/(1^i)} = 0$ (because $\left(1^i\right) \not\subseteq \nu$). Now,

$$(-1)^i h_{m+i} e_i^\perp \underbrace{\overline{s}_{(\alpha_1,\alpha_2,\ldots,\alpha_n)}}_{=\pm s_\nu} = \pm (-1)^i h_{m+i} \underbrace{e_i^\perp s_\nu}_{=0} = 0,$$

which concludes our proof. Exercise 2.9.1(d) is thus solved.

(b) Let $n = \ell(\lambda)$. Then, $\lambda = (\lambda_1, \lambda_2, ..., \lambda_n)$, and the equality (2.9.1) yields $s_\lambda = \overline{s}_{(\lambda_1,\lambda_2,\ldots,\lambda_n)}$.

Also, $\ell((m, \lambda_1, \lambda_2, \lambda_3, ...)) = n + 1$, and thus (2.9.1) (applied to $n + 1$ and $(m, \lambda_1, \lambda_2, \lambda_3, ...)$ instead of n and $\lambda = (\lambda_1, \lambda_2, \lambda_3, ...)$) yields

$$s_{(m,\lambda_1,\lambda_2,\lambda_3,\ldots)} = \overline{s}_{(m,\lambda_1,\lambda_2,\ldots,\lambda_n)}.$$

Compared with

$$\sum_{i\in\mathbb{N}} (-1)^i h_{m+i} e_i^\perp \underbrace{s_\lambda}_{=\overline{s}_{(\lambda_1,\lambda_2,\ldots,\lambda_n)}} = \sum_{i\in\mathbb{N}} (-1)^i h_{m+i} e_i^\perp \overline{s}_{(\lambda_1,\lambda_2,\ldots,\lambda_n)}$$
$$= \overline{s}_{(m,\lambda_1,\lambda_2,\ldots,\lambda_n)} \qquad \text{(by (2.9.3), applied to } \alpha_j = \lambda_j\text{)},$$

this yields $\sum_{i\in\mathbb{N}} (-1)^i h_{m+i} e_i^\perp s_\lambda = s_{(m,\lambda_1,\lambda_2,\lambda_3,\ldots)}$. This solves Exercise 2.9.1(b).

(e) This follows by induction over n from Exercise 2.9.1(d). The induction base (the $n = 0$ case) is trivial. For the induction step, we need to prove that

$$\overline{s}_{(\alpha_1,\alpha_2,\ldots,\alpha_{n+1})} = \left(\mathbf{B}_{\alpha_1} \circ \mathbf{B}_{\alpha_2} \circ \cdots \circ \mathbf{B}_{\alpha_{n+1}}\right)(1),$$

assuming that

$$\overline{s}_{(\alpha_2,\alpha_3,\ldots,\alpha_{n+1})} = \left(\mathbf{B}_{\alpha_2} \circ \mathbf{B}_{\alpha_3} \circ \cdots \circ \mathbf{B}_{\alpha_{n+1}}\right)(1). \tag{13.81.7}$$

But this is fairly straightforward: Applying the map $\mathbf{B}_{\alpha_1}$ to (13.81.7), we obtain

$$\mathbf{B}_{\alpha_1}\left(\overline{s}_{(\alpha_2,\alpha_3,\ldots,\alpha_{n+1})}\right)=\mathbf{B}_{\alpha_1}\left(\left(\mathbf{B}_{\alpha_2}\circ\mathbf{B}_{\alpha_3}\circ\cdots\circ\mathbf{B}_{\alpha_{n+1}}\right)(1)\right)=\left(\mathbf{B}_{\alpha_1}\circ\mathbf{B}_{\alpha_2}\circ\cdots\circ\mathbf{B}_{\alpha_{n+1}}\right)(1).$$

Compared with

$$\begin{aligned}\mathbf{B}_{\alpha_1}\left(\overline{s}_{(\alpha_2,\alpha_3,\ldots,\alpha_{n+1})}\right)&=\sum_{i\in\mathbb{N}}(-1)^i h_{\alpha_1+i}e_i^{\perp}\overline{s}_{(\alpha_2,\alpha_3,\ldots,\alpha_{n+1})} && \text{(by the definition of the map } \mathbf{B}_{\alpha_1}\text{)}\\ &=\overline{s}_{(\alpha_1,\alpha_2,\alpha_3,\ldots,\alpha_{n+1})} && \left(\begin{array}{c}\text{by Exercise 2.9.1(d), applied to } \alpha_1 \text{ and}\\ (\alpha_2,\alpha_3,\ldots,\alpha_{n+1}) \text{ instead of } m \text{ and } (\alpha_1,\alpha_2,\ldots,\alpha_n)\end{array}\right)\\ &=\overline{s}_{(\alpha_1,\alpha_2,\ldots,\alpha_{n+1})},\end{aligned}$$

this yields

$$\overline{s}_{(\alpha_1,\alpha_2,\ldots,\alpha_{n+1})}=\left(\mathbf{B}_{\alpha_1}\circ\mathbf{B}_{\alpha_2}\circ\cdots\circ\mathbf{B}_{\alpha_{n+1}}\right)(1),$$

which completes the induction step and thus the proof.

(f) Let $m\in\mathbb{Z}$. Let n be a positive integer. We have $e_0=1$ and thus $e_0^{\perp}=1^{\perp}=\operatorname{id}$ (by Proposition 2.8.2(iii), applied to $A=\Lambda$). Hence, $e_0^{\perp}p_n=\operatorname{id}(p_n)=p_n$.

If i is a positive integer satisfying $i\neq n$, then

$$e_i^{\perp}p_n=0 \tag{13.81.8}$$

(by Exercise 2.8.8(d), applied to $m=i$).

The definition of $\mathbf{B}_m$ yields

$$\begin{aligned}\mathbf{B}_m(p_n)&=\sum_{i\in\mathbb{N}}(-1)^i h_{m+i}e_i^{\perp}p_n\\ &=\underbrace{(-1)^0}_{=1}\underbrace{h_{m+0}}_{=h_m}\underbrace{e_0^{\perp}p_n}_{=p_n}+\sum_{i>0}(-1)^i h_{m+i}e_i^{\perp}p_n\\ &\qquad\text{(here, we have split off the addend for } i=0 \text{ from the sum)}\\ &=h_mp_n+\sum_{i>0}(-1)^i h_{m+i}e_i^{\perp}p_n.\end{aligned}$$

In view of

$$\begin{aligned}&\sum_{i>0}(-1)^i h_{m+i}e_i^{\perp}p_n\\ &=(-1)^n h_{m+n}\underbrace{e_n^{\perp}p_n}_{\substack{=(-1)^{n-1}\\ \text{(by Exercise 2.8.8(c))}}}+\sum_{\substack{i>0;\\ i\neq n}}(-1)^i h_{m+i}\underbrace{e_i^{\perp}p_n}_{\substack{=0\\ \text{(by (13.81.8))}}}\\ &\qquad\left(\begin{array}{c}\text{here, we have split off the addend for } i=n \text{ from the sum,}\\ \text{since } n \text{ is a positive integer}\end{array}\right)\\ &=(-1)^n h_{m+n}(-1)^{n-1}+\underbrace{\sum_{\substack{i>0;\\ i\neq n}}(-1)^i h_{m+i}0}_{=0}=(-1)^n h_{m+n}(-1)^{n-1}=\underbrace{(-1)^n(-1)^{n-1}}_{\substack{=(-1)^{n+(n-1)}=-1\\ \text{(since } n+(n-1)=2n-1\\ \text{is odd)}}}h_{m+n}=-h_{m+n},\end{aligned}$$

this becomes

$$\mathbf{B}_m(p_n)=h_mp_n+\underbrace{\sum_{i>0}(-1)^i h_{m+i}e_i^{\perp}p_n}_{=-h_{m+n}}=h_mp_n+(-h_{m+n})=h_mp_n-h_{m+n}.$$

This solves Exercise 2.9.1(f).

Remark: Our solution to Exercise 2.9.1(d) was modelled after the rough sketch of a solution to Exercise 2.9.1(b) given in [227, §4.20]; but it involved many technicalities which are not necessary if one is only interested in a solution to Exercise 2.9.1(b). (Specifically, if one only wants to solve Exercise 2.9.1(b), one can avoid the use of Exercise 2.9.1(c).)

We solved Exercise 2.9.1(e) using Exercise 2.9.1(d), but of course one can just as well turn this around and solve Exercise 2.9.1(d) using Exercise 2.9.1(e) if one has an independent solution to Exercise 2.9.1(e). Such an independent solution can be extracted from [17, Corollary 3.30] (using the realization that the immaculate creation operators $\mathbb{B}_m$ of [17] are lifts of our Bernstein operators $\mathbf{B}_m$ to NSym).

13.82. **Solution to Exercise 2.9.3.** *Solution to Exercise 2.9.3.* (a) We shall prove a more general result:

> *Claim 1:* Let A be **any** commutative ring. Let $f \in A[[t]]$ be a power series with constant term 1. Then, there is a unique family $(x_n)_{n\geq 1}$ of elements of A such that

$$f = \prod_{n=1}^{\infty} (1 - x_n t^n)^{-1}. \tag{13.82.1}$$

[*Proof of Claim 1:* This family is constructed recursively: If $x_1, x_2, \ldots, x_{k-1}$ have been determined (for some $k \geq 1$), then x_k is obtained by comparing coefficients before t^k in the equation (13.82.1) (the coefficient before t^k on the left hand side is a known constant, whereas the coefficient before t^k on the right hand side can be written in the form $x_k +$ (some polynomial in $x_1, x_2, ..., x_{k-1}$), and thus the equality of these coefficients gives a linear equation in x_k which can be uniquely solved for x_k). The family $(x_n)_{n\geq 1}$ thus constructed satisfies (13.82.1) (because the constant terms on both sides of (13.82.1) are 1, whereas for every $k \geq 1$, the coefficients before t^k on both sides of (13.82.1) are equal due to the construction of x_k). Moreover, it is the only family that satisfies (13.82.1) (since its construction was dictated by (13.82.1)). Thus, Claim 1 is proven.]

Exercise 2.9.3(a) follows by applying Claim 1 to $A = \Lambda$ and $f = H(t)$.

(b) Again, this generalizes: Let us say that a power series $f \in A[[t]]$ over a graded commutative ring A is *equigraded* if, for every $n \in \mathbb{N}$, the coefficient of f before t^n is homogeneous of degree n. Then, if $f \in A[[t]]$ is an equigraded power series with constant term 1, then the unique family $(x_n)_{n\geq 1}$ satisfying (13.82.1) has the property that x_n is homogeneous of degree n for every positive n. This is rather easy to see by induction.

(c) By the definition of the w_n, we have $H(t) = \prod_{n=1}^{\infty} (1 - w_n t^n)^{-1}$. Expanding and comparing coefficients yields precisely $\sum_{\lambda \in \mathrm{Par}_n} w_\lambda = h_n$. [648]

[648] Here are some more details of this argument. From (2.4.1), we have $H(t) = \sum_{n\geq 0} h_n(\mathbf{x}) t^n = \sum_{n\geq 0} h_n t^n$. Thus,

$$\begin{aligned}
\sum_{n\geq 0} h_n t^n = H(t) &= \prod_{n=1}^{\infty} \underbrace{(1 - w_n t^n)^{-1}}_{=\sum_{m\in\mathbb{N}} (w_n t^n)^m} = \prod_{n=1}^{\infty} \sum_{m\in\mathbb{N}} \underbrace{(w_n t^n)^m}_{=w_n^m t^{nm}} \\
&= \prod_{n=1}^{\infty} \sum_{m\in\mathbb{N}} w_n^m t^{nm} = \sum_{\substack{(m_1,m_2,m_3,\ldots) \\ \text{weak composition}}} \underbrace{\prod_{n=1}^{\infty} w_n^{m_n} t^{nm_n}}_{=\left(\prod_{n=1}^{\infty} w_n^{m_n}\right) t^{1m_1+2m_2+3m_3+\cdots}} \qquad \text{(by the product rule)} \\
&= \sum_{\substack{(m_1,m_2,m_3,\ldots) \\ \text{weak composition}}} \left(\prod_{n=1}^{\infty} w_n^{m_n} \right) t^{1m_1+2m_2+3m_3+\cdots}. \qquad (13.82.2)
\end{aligned}$$

But every partition λ can be uniquely written in the form $\lambda = (1^{m_1} 2^{m_2} 3^{m_3} \cdots)$ for some weak composition $(m_1, m_2, m_3, \ldots)$. Thus, we can substitute $(1^{m_1} 2^{m_2} 3^{m_3} \cdots)$ for λ in the sum $\sum_{\lambda\in\mathrm{Par}} w_\lambda t^{|\lambda|}$. As a result, we obtain

$$\begin{aligned}
\sum_{\lambda\in\mathrm{Par}} w_\lambda t^{|\lambda|} &= \sum_{\substack{(m_1,m_2,m_3,\ldots) \\ \text{weak composition}}} \underbrace{w_{(1^{m_1} 2^{m_2} 3^{m_3} \cdots)}}_{\substack{=w_1^{m_1} w_2^{m_2} w_3^{m_3} \cdots \\ \text{(by the definition of } w_{(1^{m_1} 2^{m_2} 3^{m_3} \cdots)})}} \underbrace{t^{|(1^{m_1} 2^{m_2} 3^{m_3} \cdots)|}}_{\substack{=t^{1m_1+2m_2+3m_3+\cdots} \\ \text{(since } |(1^{m_1} 2^{m_2} 3^{m_3} \cdots)| = 1m_1+2m_2+3m_3+\cdots)}} \\
&= \sum_{\substack{(m_1,m_2,m_3,\ldots) \\ \text{weak composition}}} \underbrace{\left(w_1^{m_1} w_2^{m_2} w_3^{m_3} \cdots\right)}_{=\prod_{n=1}^{\infty} w_n^{m_n}} t^{1m_1+2m_2+3m_3+\cdots} = \sum_{\substack{(m_1,m_2,m_3,\ldots) \\ \text{weak composition}}} \left(\prod_{n=1}^{\infty} w_n^{m_n} \right) t^{1m_1+2m_2+3m_3+\cdots}.
\end{aligned}$$

Compared with (13.82.2), this yields

$$\sum_{n\geq 0} h_n t^n = \sum_{\lambda\in\mathrm{Par}} w_\lambda t^{|\lambda|} = \sum_{n\geq 0} \underbrace{\sum_{\substack{\lambda\in\mathrm{Par}; \\ |\lambda|=n}}}_{=\sum_{\lambda\in\mathrm{Par}_n}} w_\lambda \underbrace{t^{|\lambda|}}_{\substack{=t^n \\ \text{(since } |\lambda|=n)}} = \sum_{n\geq 0} \sum_{\lambda\in\mathrm{Par}_n} w_\lambda t^n.$$

Comparing coefficients before t^n in this equality of power series, we conclude that $h_n = \sum_{\lambda\in\mathrm{Par}_n} w_\lambda$ for every $n \in \mathbb{N}$, qed.

(d) *First solution to Exercise 2.9.3(d) (sketched).* If $\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(\ell)}$ are finitely many partitions, then we let $\lambda^{(1)} \cup \lambda^{(2)} \cup \cdots \cup \lambda^{(\ell)}$ denote the partition obtained by sorting all the parts of $\lambda^{(1)}, \lambda^{(2)}, ..., \lambda^{(\ell)}$ in decreasing order. For instance, $(3,2,1) \cup (4,2) \cup (5,1) = (5,4,3,2,2,1,1)$. Clearly, if $\lambda^{(1)}, \lambda^{(2)}, ..., \lambda^{(\ell)}$ are finitely many partitions, then

$$w_{\lambda^{(1)}} w_{\lambda^{(2)}} \cdots w_{\lambda^{(\ell)}} = w_{\lambda^{(1)} \cup \lambda^{(2)} \cup \cdots \cup \lambda^{(\ell)}} \tag{13.82.3}$$

(due to the definition of w_λ for a partition λ).

From (c), we know that $h_n = \sum_{\lambda \in \operatorname{Par}_n} w_\lambda$ for every $n \in \mathbb{N}$. In other words,

$$h_n = \sum_{\lambda \vdash n} w_\lambda \qquad \text{for every } n \in \mathbb{N}, \tag{13.82.4}$$

where we are using the notation $\lambda \vdash n$ for $\lambda \in \operatorname{Par}_n$.

Every partition $\mu = (\mu_1, \mu_2, ..., \mu_\ell)$ with $\ell = \ell(\mu)$ satisfies

$$\begin{aligned} h_\mu = h_{\mu_1} h_{\mu_2} \cdots h_{\mu_\ell} &= \prod_{i=1}^{\ell} \underbrace{h_{\mu_i}}_{\substack{= \sum_{\lambda \vdash |\mu_i|} w_\lambda \\ \text{(by (13.82.4), applied to } n = \mu_i)}} \\ &= \prod_{i=1}^{\ell} \left(\sum_{\lambda \vdash |\mu_i|} w_\lambda \right) = \sum_{\substack{(\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(\ell)}) \in \operatorname{Par}^\ell; \\ \lambda^{(i)} \vdash \mu_i \text{ for every } i}} \underbrace{w_{\lambda^{(1)}} w_{\lambda^{(2)}} \cdots w_{\lambda^{(\ell)}}}_{\substack{= w_{\lambda^{(1)} \cup \lambda^{(2)} \cup \cdots \cup \lambda^{(\ell)}} \\ \text{(by (13.82.3))}}} \\ &= \sum_{\substack{(\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(\ell)}) \in \operatorname{Par}^\ell; \\ \lambda^{(i)} \vdash \mu_i \text{ for every } i}} w_{\lambda^{(1)} \cup \lambda^{(2)} \cup \cdots \cup \lambda^{(\ell)}}. \end{aligned} \tag{13.82.5}$$

Now, let us fix $n \in \mathbb{N}$. We shall prove that $(w_\lambda)_{\lambda \in \operatorname{Par}_n}$ is a basis of the $\mathbf{k}$-module Λ_n. First of all, we know that this family $(w_\lambda)_{\lambda \in \operatorname{Par}_n}$ is a family of elements of Λ_n (since for each $\lambda \in \operatorname{Par}_n$, the symmetric function w_λ is homogeneous of degree $|\lambda| = n$).

Now, we define a binary relation $\underset{\text{ref}}{\leq}$ on the set Par_n as follows: If $\lambda \in \operatorname{Par}_n$ and $\mu \in \operatorname{Par}_n$, then we let $\lambda \underset{\text{ref}}{\leq} \mu$ if and only if there exists a tuple $(\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(\ell)}) \in \operatorname{Par}^\ell$ satisfying $\ell = \ell(\mu)$, $(\lambda^{(i)} \vdash \mu_i \text{ for every } i)$ and $\lambda^{(1)} \cup \lambda^{(2)} \cup \cdots \cup \lambda^{(\ell)} = \lambda$. The intuitive meaning behind this is the following: We have $\lambda \underset{\text{ref}}{\leq} \mu$ if we can obtain the partition λ by splitting each part μ_i of μ into several smaller parts[649] (which are positive integers summing up to μ_i) and sorting the resulting list into decreasing order. For instance, $(5,3,2,2,2,1) \underset{\text{ref}}{\leq} (6,5,4)$ (because the tuple $(\lambda^{(1)}, \lambda^{(2)}, \lambda^{(3)}) = ((3,2,1), (5), (2,2))$ satisfies $(3,2,1) \vdash 6$, $(5) \vdash 5$ and $(2,2) \vdash 4$ and $(3,2,1) \cup (5) \cup (2,2) = (5,3,2,2,2,1)$).

It is easy to see that the relation $\underset{\text{ref}}{\leq}$ is transitive, antisymmetric and reflexive[650]. Hence, $\underset{\text{ref}}{\leq}$ is the smaller-or-equal relation of a partial order on the set Par_n. Consider Par_n as a poset, equipped with this partial order. (This partial order is called the *refinement order* on partitions[651].)

Now, for any two partitions λ and μ, let $b_{\mu,\lambda}$ denote the number of tuples $(\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(\ell)}) \in \operatorname{Par}^\ell$ satisfying $\ell = \ell(\mu)$, $(\lambda^{(i)} \vdash \mu_i \text{ for every } i)$ and $\lambda^{(1)} \cup \lambda^{(2)} \cup \cdots \cup \lambda^{(\ell)} = \lambda$. Then, the formula (13.82.5) rewrites as

$$h_\mu = \sum_{\lambda \in \operatorname{Par}} b_{\mu,\lambda} w_\lambda.$$

[649] We allow the "several smaller parts" to be one single part (namely, μ_i).

[650] In proving these properties (specifically, antisymmetry), it helps to observe the following fact: If $\lambda \in \operatorname{Par}_n$ and $\mu \in \operatorname{Par}_n$ satisfy $\lambda \underset{\text{ref}}{\leq} \mu$, then either $\lambda = \mu$ or $\ell(\lambda) > \ell(\mu)$.

[651] Caution: This order is **not** a restriction of the refinement order on compositions.

Thus, for every $\mu \in \operatorname{Par}_n$, we have

$$h_\mu = \sum_{\lambda \in \operatorname{Par}} b_{\mu,\lambda} w_\lambda = \underbrace{\sum_{\substack{\lambda \in \operatorname{Par}; \\ |\lambda| = n}}}_{= \sum_{\lambda \in \operatorname{Par}_n}} b_{\mu,\lambda} w_\lambda + \sum_{\substack{\lambda \in \operatorname{Par}; \\ |\lambda| \neq n}} \underbrace{b_{\mu,\lambda}}_{\substack{=0 \\ \text{(this easily follows} \\ \text{from } |\lambda| \neq n = |\mu|)}} w_\lambda$$
$$= \sum_{\lambda \in \operatorname{Par}_n} b_{\mu,\lambda} w_\lambda + \underbrace{\sum_{\substack{\lambda \in \operatorname{Par}; \\ |\lambda| \neq n}} 0 w_\lambda}_{=0} = \sum_{\lambda \in \operatorname{Par}_n} b_{\mu,\lambda} w_\lambda .$$

Hence, the family $(h_\lambda)_{\lambda \in \operatorname{Par}_n}$ expands[652] in the family $(w_\lambda)_{\lambda \in \operatorname{Par}_n}$ through the matrix $(b_{\mu,\lambda})_{(\mu,\lambda) \in \operatorname{Par}_n \times \operatorname{Par}_n}$. But this matrix $(b_{\mu,\lambda})_{(\mu,\lambda) \in \operatorname{Par}_n \times \operatorname{Par}_n}$ is easily seen to be unitriangular (indeed, $b_{\mu,\lambda} = 0$ for any $(\mu, \lambda) \in \operatorname{Par}_n$ which do not satisfy $\lambda \underset{\text{ref}}{\leq} \mu$, and furthermore, every $\lambda \in \operatorname{Par}_n$ satisfies $b_{\lambda,\lambda} = 1$) and therefore invertibly triangular. Hence, the family $(h_\lambda)_{\lambda \in \operatorname{Par}_n}$ expands invertibly triangularly in the family $(w_\lambda)_{\lambda \in \operatorname{Par}_n}$. Corollary 11.1.19(e) (applied to Λ_n, Par_n, $(h_\lambda)_{\lambda \in \operatorname{Par}_n}$ and $(w_\lambda)_{\lambda \in \operatorname{Par}_n}$) thus shows that the family $(h_\lambda)_{\lambda \in \operatorname{Par}_n}$ is a basis of the **k**-module Λ_n if and only if the family $(w_\lambda)_{\lambda \in \operatorname{Par}_n}$ is a basis of the **k**-module Λ_n. Hence, the family $(w_\lambda)_{\lambda \in \operatorname{Par}_n}$ is a basis of the **k**-module Λ_n (since we know that the family $(h_\lambda)_{\lambda \in \operatorname{Par}_n}$ is a basis of the **k**-module Λ_n).

Now, forget that we fixed n. We thus have shown that, for every $n \in \mathbb{N}$, the family $(w_\lambda)_{\lambda \in \operatorname{Par}_n}$ is a basis of the **k**-module Λ_n. Hence, the disjoint union of the families $(w_\lambda)_{\lambda \in \operatorname{Par}_n}$ over all $n \in \mathbb{N}$ is a basis of the direct sum $\bigoplus_{n \in \mathbb{N}} \Lambda_n$. Since the former disjoint union is the family $(w_\lambda)_{\lambda \in \operatorname{Par}}$, whereas the latter direct sum is $\bigoplus_{n \in \mathbb{N}} \Lambda_n = \Lambda$, this result rewrites as follows: The family $(w_\lambda)_{\lambda \in \operatorname{Par}}$ is a basis of the **k**-module Λ. This solves Exercise 2.9.3(d).

[*Remark:* We could have slightly simplified this argument by using a coarser partial order instead of $\underset{\text{ref}}{\leq}$. Namely, we can define a binary relation $\underset{\text{len}}{\leq}$ on the set Par_n as follows: If $\lambda \in \operatorname{Par}_n$ and $\mu \in \operatorname{Par}_n$, then we let $\lambda \underset{\text{len}}{\leq} \mu$ if and only if either $\ell(\lambda) > \ell(\mu)$ or $\lambda = \mu$. Then, clearly, $\underset{\text{len}}{\leq}$ is the smaller-or-equal relation of a partial order on the set Par_n. If we consider Par_n as a poset equipped with this partial order, then the matrix $(b_{\mu,\lambda})_{(\mu,\lambda) \in \operatorname{Par}_n \times \operatorname{Par}_n}$ is still unitriangular, and thus our argument above still works, but we save ourselves the trouble of proving that $\underset{\text{ref}}{\leq}$ is a partial order.]

Second solution to Exercise 2.9.3(d) (sketched). From part (c), we see that h_n can be written as a polynomial in the $w_1, w_2, w_3, \ldots$ for each $n \in \mathbb{N}$. Therefore, $w_1, w_2, w_3, \ldots$ generate Λ as a **k**-algebra (because $h_1, h_2, h_3, \ldots$ generate Λ as a **k**-algebra). In other words, the family $(w_\lambda)_{\lambda \in \operatorname{Par}}$ spans the **k**-module Λ [653]. Recall that $w_\lambda \in \Lambda_{|\lambda|}$ for each $\lambda \in \operatorname{Par}$. In other words, w_λ is an element of $\Lambda_{|\lambda|}$ for each partition λ. Hence, Exercise 2.5.19 (applied to $v_\lambda = w_\lambda$) yields that the family $(w_\lambda)_{\lambda \in \operatorname{Par}}$ is a graded basis of the graded **k**-module Λ. Thus, in particular, this family $(w_\lambda)_{\lambda \in \operatorname{Par}}$ is a basis of the **k**-module Λ. This solves Exercise 2.9.3(d) again.

[652] Here, we are using the terminology of Section 11.1.

[653] We have used Exercise 2.2.14(b) (applied to w_m and w_λ instead of v_m and v_λ) here. Alternatively, this can also be seen using (13.82.5).

(e) By the definition of the w_n, we have $H(t) = \prod_{n=1}^{\infty} (1 - w_n t^n)^{-1}$. Hence,[654]

$$\frac{d}{dt}(\log H(t)) = \frac{d}{dt}\left(\log \prod_{n=1}^{\infty} (1 - w_n t^n)^{-1}\right) = \frac{d}{dt}\sum_{n=1}^{\infty} \log\left((1 - w_n t^n)^{-1}\right)$$
$$= \sum_{n=1}^{\infty} \underbrace{\frac{d}{dt}\log\left((1 - w_n t^n)^{-1}\right)}_{=\dfrac{n w_n t^{n-1}}{1 - w_n t^n}} = \sum_{n=1}^{\infty} \frac{n w_n t^{n-1}}{1 - w_n t^n}.$$

Compared with

$$\frac{d}{dt}(\log H(t)) = \frac{H'(t)}{H(t)} = \sum_{m \geq 0} p_{m+1} t^m \qquad \text{(by Exercise 2.5.21)},$$

this yields

$$\sum_{m \geq 0} p_{m+1} t^m = \sum_{n=1}^{\infty} \frac{n w_n t^{n-1}}{1 - w_n t^n}.$$

Multiplying this by t, we obtain

$$\sum_{m \geq 0} p_{m+1} t^{m+1} = \sum_{n=1}^{\infty} \underbrace{\frac{n w_n t^n}{1 - w_n t^n}}_{= n \sum_{k=1}^{\infty} w_n^k t^{nk}} = \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} n w_n^k t^{nk}$$
$$= \sum_{m=1}^{\infty} \sum_{d \mid m} d w_d^{m/d} t^m = \sum_{n=1}^{\infty} \sum_{d \mid n} d w_d^{n/d} t^n,$$

so that $\sum_{n=1}^{\infty} \sum_{d \mid n} d w_d^{n/d} t^n = \sum_{m \geq 0} p_{m+1} t^{m+1} = \sum_{n \geq 1} p_n t^n$. Comparing coefficients yields the claim of part (e).

[654] We will be using logarithms here, so prima facie our argument only works when the base ring $\mathbf{k}$ is a $\mathbb{Q}$-algebra. However, it is easy to see that our argument can easily be adapted to work in the general case as well. For example, one can argue that even if the power series $\log t$ is not defined if $\mathbf{k}$ is not a $\mathbb{Q}$-algebra, the notion of the logarithmic derivative $\frac{d}{dt}(\log Q)$ is well-defined for every $\mathbf{k}$ whenever $Q \in \mathbf{k}[[t]]$ is a power series with constant term 1 (for example, one could use the formula $\frac{d}{dt}(\log Q) = \frac{Q'(t)}{Q(t)}$ as the *definition* of the logarithmic derivative) and still has the familiar property of turning products into sums in this generality. See the solution to Exercise 2.5.21 for details about this.

(f) This is done in [55]. For the sake of completeness: First, let $n \geq k \geq 2$. Then, every $i \in \{2, 3, ..., k-1\}$ satisfies $f_{i,i} = \sum_{\substack{\lambda \in \operatorname{Par}_i, \\ \min \lambda \geq i}} w_\lambda = w_{(i)} = w_i$. Thus,

$$\begin{aligned}
\sum_{i=2}^{k-1} \underbrace{f_{i,i}}_{=w_i} \underbrace{f_{n-i,i}}_{=\sum_{\substack{\lambda \in \operatorname{Par}_{n-i}, \\ \min \lambda \geq i}} w_\lambda} &= \sum_{i=2}^{k-1} \underbrace{w_i \sum_{\substack{\lambda \in \operatorname{Par}_{n-i}, \\ \min \lambda \geq i}} w_\lambda}_{= \sum_{\substack{\lambda \in \operatorname{Par}_{n-i}, \\ \min \lambda \geq i}} w_\lambda w_i = \sum_{\substack{\lambda \in \operatorname{Par}_{n}, \\ \min \lambda = i}} w_\lambda} = \sum_{i=2}^{k-1} \sum_{\substack{\lambda \in \operatorname{Par}_{n}, \\ \min \lambda = i}} w_\lambda \\
&= \sum_{\substack{\lambda \in \operatorname{Par}_{n}, \\ 2 \leq \min \lambda < k}} w_\lambda = \underbrace{\sum_{\lambda \in \operatorname{Par}_n} w_\lambda}_{\substack{=h_n \\ \text{(by part (c))}}} - \underbrace{\sum_{\substack{\lambda \in \operatorname{Par}_{n}, \\ \min \lambda = 1}} w_\lambda}_{=w_1 \sum_{\lambda \in \operatorname{Par}_{n-1}} w_\lambda} - \underbrace{\sum_{\substack{\lambda \in \operatorname{Par}_{n}, \\ \min \lambda \geq k}} w_\lambda}_{=f_{n,k}} \\
&= h_n - \underbrace{w_1}_{=h_1} \underbrace{\sum_{\lambda \in \operatorname{Par}_{n-1}} w_\lambda}_{\substack{=h_{n-1} \\ \text{(by part (c))}}} - f_{n,k} \\
&= \underbrace{h_n}_{=s_{(n)}} - h_1 \underbrace{h_{n-1}}_{=s_{(n-1)}} - f_{n,k} \\
&= s_{(n)} - \underbrace{h_1 s_{(n-1)}}_{\substack{=s_{(n)} + s_{(n-1,1)} \\ \text{(by the Pieri rule)}}} - f_{n,k} \\
&= -s_{(n-1,1)} - f_{n,k};
\end{aligned}$$

in other words,

$$-f_{n,k} = s_{(n-1,1)} + \sum_{i=2}^{k-1} f_{i,i} f_{n-i,i}.$$

From this, we can conclude inductively that $-f_{n,k}$ is a sum of Schur functions for every $n \in \mathbb{N}$ and $k \geq 2$ (in fact, the trivial cases with $n < 2$ have to be taken as induction base). Since $f_{n,n} = \sum_{\substack{\lambda \in \operatorname{Par}_n, \\ \min \lambda \geq n}} w_\lambda = w_{(n)} = w_n$, this yields that $-w_n$ is a sum of Schur functions for every $n \geq 2$.

(g) Every partition λ can be uniquely written in the form $\lambda = (1^{m_1} 2^{m_2} 3^{m_3} \cdots)$ for some weak composition $(m_1, m_2, m_3, \ldots)$. Thus,

$$\begin{aligned}
\sum_{\lambda \in \operatorname{Par}} w_\lambda(\mathbf{x}) r_\lambda(\mathbf{y}) &= \sum_{\substack{(m_1,m_2,m_3,\ldots) \\ \text{weak composition}}} \underbrace{w_{(1^{m_1}2^{m_2}3^{m_3}\cdots)}(\mathbf{x})}_{\substack{=\prod_{i\geq 1}(w_i(\mathbf{x}))^{m_i} \\ \text{(by the definition of } w_\lambda\text{)}}} \underbrace{r_{(1^{m_1}2^{m_2}3^{m_3}\cdots)}(\mathbf{y})}_{\substack{=\prod_{i\geq 1} h_{m_i}\left(y_1^i, y_2^i, y_3^i, \ldots\right) \\ \text{(by the definition of } r_\lambda\text{)}}} \\
&= \sum_{\substack{(m_1,m_2,m_3,\ldots) \\ \text{weak composition}}} \left(\prod_{i \geq 1} (w_i(\mathbf{x}))^{m_i} \right) \left(\prod_{i \geq 1} h_{m_i}\left(y_1^i, y_2^i, y_3^i, \ldots\right) \right) \\
&= \sum_{\substack{(m_1,m_2,m_3,\ldots) \\ \text{weak composition}}} \prod_{i \geq 1} (w_i(\mathbf{x}))^{m_i} h_{m_i}\left(y_1^i, y_2^i, y_3^i, \ldots\right) \\
&= \sum_{\substack{(m_1,m_2,m_3,\ldots) \\ \text{weak composition}}} \prod_{i \geq 1} h_{m_i}\left(y_1^i, y_2^i, y_3^i, \ldots\right) (w_i(\mathbf{x}))^{m_i} \\
&= \prod_{i \geq 1} \underbrace{\sum_{m \in \mathbb{N}} h_m\left(y_1^i, y_2^i, y_3^i, \ldots\right) (w_i(\mathbf{x}))^m}_{\substack{=\prod_{j\geq 1}\left(1 - y_j^i w_i(\mathbf{x})\right)^{-1} \\ \text{(by (2.2.18), upon substitution of } (y_1,y_2,y_3,\ldots) \text{ and } w_i(\mathbf{x}) \text{ for } \mathbf{x} \text{ and } t\text{)}}} \\
&= \prod_{i \geq 1} \prod_{j \geq 1} \left(1 - y_j^i w_i(\mathbf{x})\right)^{-1} = \prod_{j \geq 1} \underbrace{\prod_{i \geq 1} \left(1 - y_j^i w_i(\mathbf{x})\right)^{-1}}_{\substack{=\prod_{n\geq 1}\left(1 - y_j^n w_n(\mathbf{x})\right)^{-1} \\ =\prod_{n=1}^{\infty}\left(1 - w_n(\mathbf{x}) y_j^n\right)^{-1} \\ =H(y_j) \\ \text{(since } \prod_{n=1}^{\infty}(1-w_n t^n)^{-1} = H(t)\text{)}}} \\
&= \prod_{j \geq 1} \underbrace{H(y_j)}_{=\prod_{i=1}^{\infty}(1 - x_i y_j)^{-1}} = \prod_{i,j=1}^{\infty} (1 - x_i y_j)^{-1}.
\end{aligned}$$

This proves (g).

(h) For every partition λ, both symmetric functions w_λ and r_λ are homogeneous of degree $|\lambda|$. (In fact, for w_λ this follows from Exercise 2.9.3(c), whereas for r_λ this is easily derived from the definition.)

From Exercise 2.9.3(g), we obtain $\sum_{\lambda \in \operatorname{Par}} w_\lambda(\mathbf{x}) r_\lambda(\mathbf{y}) = \prod_{i,j=1}^{\infty} (1 - x_i y_j)^{-1} = \sum_{\lambda \in \operatorname{Par}} s_\lambda(\mathbf{x}) s_\lambda(\mathbf{y})$ (by (2.5.1)). Thus, we can apply Exercise 2.5.20(a) to $u_\lambda = w_\lambda$ and $v_\lambda = r_\lambda$. As a result, we obtain that $(w_\lambda)_{\lambda \in \operatorname{Par}}$ and $(r_\lambda)_{\lambda \in \operatorname{Par}}$ are $\mathbf{k}$-bases of Λ, and actually are dual bases with respect to the Hall inner product on Λ. Thus we have solved Exercise 2.9.3(h), but also given another proof of Exercise 2.9.3(d) in the process (because we have shown once again that $(w_\lambda)_{\lambda \in \operatorname{Par}}$ is a $\mathbf{k}$-basis of Λ).

13.83. **Solution to Exercise 2.9.4.** *Solution to Exercise 2.9.4.* (a) Let $f \in \Lambda$.

Recall the following fundamental fact from linear algebra: If $\mathbf{k}$ is a commutative ring, if A is a $\mathbf{k}$-module, if $(\cdot, \cdot) : A \times A \to \mathbf{k}$ is a symmetric $\mathbf{k}$-bilinear form on A, and if $(u_\lambda)_{\lambda \in L}$ and $(v_\lambda)_{\lambda \in L}$ are two $\mathbf{k}$-bases of A which are dual to each other with respect to the form $(\cdot, \cdot)$ (where L is some indexing set), then every $a \in A$ satisfies

$$a = \sum_{\lambda \in L} (u_\lambda, a) v_\lambda. \tag{13.83.1}$$

We can apply this fact to $\mathbf{k} = \mathbb{Q}$, $A = \Lambda_{\mathbb{Q}}$, $L = \operatorname{Par}$, $(u_\lambda)_{\lambda \in L} = (p_\lambda)_{\lambda \in \operatorname{Par}}$, $(v_\lambda)_{\lambda \in L} = \left(z_\lambda^{-1} p_\lambda\right)_{\lambda \in \operatorname{Par}}$ and $a = f$ (because the bases $(p_\lambda)_{\lambda \in \operatorname{Par}}$ and $\left(z_\lambda^{-1} p_\lambda\right)_{\lambda \in \operatorname{Par}}$ of $\Lambda_{\mathbb{Q}}$ are dual to each other with respect to the Hall inner product $(\cdot,\cdot)$, as Corollary 2.5.17(b) shows). As a result, we obtain

$$f = \sum_{\lambda\in\operatorname{Par}} (p_\lambda, f)\, z_\lambda^{-1} p_\lambda .$$

Hence,

$$\begin{aligned} Z(f) &= Z\left(\sum_{\lambda\in\operatorname{Par}} (p_\lambda, f)\, z_\lambda^{-1} p_\lambda\right) = \sum_{\lambda\in\operatorname{Par}} (p_\lambda, f)\, z_\lambda^{-1} \underbrace{Z(p_\lambda)}_{=z_\lambda p_\lambda} = \sum_{\lambda\in\operatorname{Par}} (p_\lambda, f)\, z_\lambda^{-1} z_\lambda p_\lambda \\ &= \sum_{\lambda\in\operatorname{Par}} \underbrace{(p_\lambda, f)}_{\substack{\in\mathbb{Z} \\ (\text{since } f\in\Lambda \text{ and } p_\lambda\in\Lambda)}} p_\lambda \in \sum_{\lambda\in\operatorname{Par}} \mathbb{Z} p_\lambda \subset \Lambda . \end{aligned}$$

Since we have proven this for every $f \in \Lambda$, we thus have shown that $Z(\Lambda) \subset \Lambda$. This solves Exercise 2.9.4(a).

(b) *First solution of Exercise 2.9.4(b):* Consider two variable sets $\mathbf{x} = (x_1, x_2, x_3, ...)$ and $\mathbf{y} = (y_1, y_2, y_3, ...)$. Let $\mathbf{xy}$ denote the variable set

$$\begin{aligned} (x_i y_j)_{(i,j)\in\{1,2,3,...\}^2} = (&x_1y_1, x_1y_2, x_1y_3, ..., \\ &x_2y_1, x_2y_2, x_2y_3, ..., \\ &x_3y_1, x_3y_2, x_3y_3, ..., \\ &...) . \end{aligned}$$

Now, we claim that for every $f \in \Lambda_{\mathbb{Q}}$,

$$\text{there is a well-defined element } f(\mathbf{xy}) := f\left((x_iy_j)_{(i,j)\in\{1,2,3,...\}^2}\right) \text{ of } \mathbb{Q}[[\mathbf{x},\mathbf{y}]] . \tag{13.83.2}$$

This claim (13.83.2) is not obvious! For example, there is generally no well-defined element $f\left((x_i+y_j)_{(i,j)\in\{1,2,3,...\}^2}\right)$, because e.g. in the case of $f = e_1$ we would have $e_1\left((x_i+y_j)_{(i,j)\in\{1,2,3,...\}^2}\right) = \sum_{(i,j)\in\{1,2,3,...\}^2} (x_i+y_j)$, which is a sum containing infinitely many x_1's. So there is some subtlety which allows us to make sense of $f\left((x_iy_j)_{(i,j)\in\{1,2,3,...\}^2}\right)$ but not of $f\left((x_i+y_j)_{(i,j)\in\{1,2,3,...\}^2}\right)$.

Here is a sketch of a *proof of (13.83.2):* Consider a new variable set $\mathbf{s} := (s_{i,j})_{(i,j)\in\{1,2,3,...\}^2}$ whose variables are indexed by **pairs** of positive integers. This variable set $\mathbf{s}$ is still countably infinite, and so there is an isomorphism $\Lambda_{\mathbb{Q}} = \Lambda_{\mathbb{Q}}(\mathbf{x}) \to \Lambda_{\mathbb{Q}}(\mathbf{s})$. On the other hand, we can consider the ring $\mathbb{Q}[[\mathbf{s}]]$ of formal power series in the variables from the set $\mathbf{s}$, and then we have $\Lambda_{\mathbb{Q}}(\mathbf{s}) \subset \mathbb{Q}[[\mathbf{s}]]$. It is easy to see that $g\left((x_iy_j)_{(i,j)\in\{1,2,3,...\}^2}\right)$ is a well-defined element of $\mathbb{Q}[[\mathbf{x},\mathbf{y}]]$ for every $g \in \mathbb{Q}[[\mathbf{s}]]$ [655]. Hence, for every $f \in \Lambda_{\mathbb{Q}}$, there is a well-defined element $f(\mathbf{xy}) := f\left((x_iy_j)_{(i,j)\in\{1,2,3,...\}^2}\right)$ of $\mathbb{Q}[[\mathbf{x},\mathbf{y}]]$ (because we can regard f as an element of $\Lambda_{\mathbb{Q}}(\mathbf{s}) \subset \mathbb{Q}[[\mathbf{s}]]$ by means of the isomorphism $\Lambda_{\mathbb{Q}} = \Lambda_{\mathbb{Q}}(\mathbf{x}) \to \Lambda_{\mathbb{Q}}(\mathbf{s})$). This proves (13.83.2).

[655] *Proof.* Let $g \in \mathbb{Q}[[\mathbf{s}]]$. When $(x_iy_j)_{(i,j)\in\{1,2,3,...\}^2}$ is substituted for $(s_{i,j})_{(i,j)\in\{1,2,3,...\}^2}$ in the power series $g \in \mathbb{Q}[[\mathbf{s}]]$, every monomial in the variables $\mathbf{s}$ turns into a monomial in the variables $(\mathbf{x},\mathbf{y})$, and **any given monomial in $(\mathbf{x},\mathbf{y})$ can only be obtained (this way) from finitely many monomials in $\mathbf{s}$** (indeed, a given monomial $\mathbf{x}^\alpha\mathbf{y}^\beta$ in $(\mathbf{x},\mathbf{y})$ can only be obtained from monomials $\prod_{i,j=1,2,3,...} s_{i,j}^{\gamma_{i,j}}$ whose exponents $\gamma_{i,j}$ satisfy the equations

$$\begin{aligned} \sum_{i=1}^{\infty} \gamma_{i,j} &= \beta_j && \text{for all } j \in \{1,2,3,...\} ; \\ \sum_{j=1}^{\infty} \gamma_{i,j} &= \alpha_i && \text{for all } i \in \{1,2,3,...\} ; \end{aligned}$$

but it is easy to see that these equations leave only finitely many possibilities for the monomial $\prod_{i,j=1,2,3,...} s_{i,j}^{\gamma_{i,j}}$). Hence, the substitution yields an infinite sum of monomials in which every monomial occurs **only finitely often**; therefore, this sum converges. This shows that $g\left((x_iy_j)_{(i,j)\in\{1,2,3,...\}^2}\right)$ is a well-defined element of $\mathbb{Q}[[\mathbf{x},\mathbf{y}]]$.

Due to (13.83.2), we can define a map

$$\begin{aligned}\widetilde{\Delta}_\times : \Lambda_{\mathbb{Q}} &\to \mathbb{Q}[[\mathbf{x},\mathbf{y}]],\\ f &\mapsto f(\mathbf{xy}) = f\left((x_i y_j)_{(i,j)\in\{1,2,3,\ldots\}^2}\right).\end{aligned}$$

This map $\widetilde{\Delta}_\times$ is an evaluation map (in an appropriate sense), and thus a $\mathbb{Q}$-algebra homomorphism. It is also clear that $\widetilde{\Delta}_\times(\Lambda) \subset \mathbb{Z}[[\mathbf{x},\mathbf{y}]]$ (where $\mathbb{Z}[[\mathbf{x},\mathbf{y}]]$ is regarded as a subring of $\mathbb{Q}[[\mathbf{x},\mathbf{y}]]$ in the obvious way).

Now, recall the $\mathbb{Q}$-algebra isomorphism $\Lambda_{\mathbb{Q}} \otimes_{\mathbb{Q}} \Lambda_{\mathbb{Q}} \to R_{\mathbb{Q}}(\mathbf{x},\mathbf{y})^{\mathfrak{S}_{(\infty)}\times\mathfrak{S}_{(\infty)}}$ constructed in (2.1.3) (applied to $\mathbf{k} = \mathbb{Q}$). This entails a $\mathbb{Q}$-algebra injection $\Lambda_{\mathbb{Q}} \otimes_{\mathbb{Q}} \Lambda_{\mathbb{Q}} \to \mathbb{Q}[[\mathbf{x},\mathbf{y}]]$ (since $R_{\mathbb{Q}}(\mathbf{x},\mathbf{y})^{\mathfrak{S}_{(\infty)}\times\mathfrak{S}_{(\infty)}} \subset R_{\mathbb{Q}}(\mathbf{x},\mathbf{y}) \subset \mathbb{Q}[[\mathbf{x},\mathbf{y}]]$). Denote this injection by ι.

We shall now show that

$$\iota \circ \Delta_\times = \widetilde{\Delta}_\times. \tag{13.83.3}$$

Proof of (13.83.3): Indeed, (13.83.3) is an equality between $\mathbb{Q}$-algebra homomorphisms (since $\widetilde{\Delta}_\times$, ι and $\Delta_\times$ are $\mathbb{Q}$-algebra homomorphisms), and thus, in order to prove it, we only need to check that it holds on a generating set of the $\mathbb{Q}$-algebra $\Lambda_{\mathbb{Q}}$. We do this on the generating set $(p_n)_{n\geq 1}$, by noticing that every $n \geq 1$ satisfies

$$\begin{aligned}(\iota \circ \Delta_\times)(p_n) &= \iota\Bigg(\underbrace{\Delta_\times(p_n)}_{=p_n\otimes p_n}\Bigg) = \iota(p_n \otimes p_n) = \underbrace{p_n(\mathbf{x})}_{=\sum_{i\geq 1} x_i^n} \underbrace{p_n(\mathbf{y})}_{=\sum_{j\geq 1} y_j^n} \qquad \text{(by the definition of } \iota)\\ &= \left(\sum_{i\geq 1} x_i^n\right)\left(\sum_{j\geq 1} y_j^n\right) = \underbrace{\sum_{i\geq 1}\sum_{j\geq 1}}_{=\sum_{(i,j)\in\{1,2,3,\ldots\}^2}} \underbrace{x_i^n y_j^n}_{=(x_i y_j)^n} = \sum_{(i,j)\in\{1,2,3,\ldots\}^2} (x_i y_j)^n\\ &= p_n\Bigg(\underbrace{(x_i y_j)_{(i,j)\in\{1,2,3,\ldots\}^2}}_{=\mathbf{xy}}\Bigg) = p_n(\mathbf{xy}) = \widetilde{\Delta}_\times(p_n)\\ &\qquad \left(\text{since } \widetilde{\Delta}_\times(p_n) = p_n(\mathbf{xy}) \text{ (by the definition of } \widetilde{\Delta}_\times(p_n))\right).\end{aligned}$$

So (13.83.3) is proven.

From (13.83.3), we obtain $\underbrace{(\iota \circ \Delta_\times)}_{=\widetilde{\Delta}_\times}(\Lambda) = \widetilde{\Delta}_\times(\Lambda) \subset \mathbb{Z}[[\mathbf{x},\mathbf{y}]]$, so that $\iota(\Delta_\times(\Lambda)) = (\iota \circ \Delta_\times)(\Lambda) \subset \mathbb{Z}[[\mathbf{x},\mathbf{y}]]$ and thus $\Delta_\times(\Lambda) \subset \iota^{-1}(\mathbb{Z}[[\mathbf{x},\mathbf{y}]])$.

But it so happens that $\iota^{-1}(\mathbb{Z}[[\mathbf{x},\mathbf{y}]]) = \Lambda \otimes_{\mathbb{Z}} \Lambda$ [656]. Hence, $\Delta_\times(\Lambda) \subset \iota^{-1}(\mathbb{Z}[[\mathbf{x},\mathbf{y}]]) = \Lambda \otimes_{\mathbb{Z}} \Lambda$. This solves Exercise 2.9.4(b).

[656] *Proof.* It is clear that $\iota(\Lambda \otimes_{\mathbb{Z}} \Lambda) \subset \mathbb{Z}[[\mathbf{x},\mathbf{y}]]$, so that $\Lambda \otimes_{\mathbb{Z}} \Lambda \subset \iota^{-1}(\mathbb{Z}[[\mathbf{x},\mathbf{y}]])$. We are now going to prove the reverse inclusion $\iota^{-1}(\mathbb{Z}[[\mathbf{x},\mathbf{y}]]) \subset \Lambda \otimes_{\mathbb{Z}} \Lambda$.

Indeed, let $p \in \iota^{-1}(\mathbb{Z}[[\mathbf{x},\mathbf{y}]])$. Then, p is an element of $\Lambda_{\mathbb{Q}} \otimes_{\mathbb{Q}} \Lambda_{\mathbb{Q}}$ satisfying $\iota(p) \in \mathbb{Z}[[\mathbf{x},\mathbf{y}]]$. All coefficients of the power series $\iota(p)$ are integers (since $\iota(p) \in \mathbb{Z}[[\mathbf{x},\mathbf{y}]]$).

We can write p in the form $p = \sum_{(\lambda,\mu)\in\operatorname{Par}\times\operatorname{Par}} \rho_{\lambda,\mu} m_\lambda \otimes m_\mu$ with $\rho_{\lambda,\mu}$ being elements of $\mathbb{Q}$ (because $(m_\lambda \otimes m_\mu)_{(\lambda,\mu)\in\operatorname{Par}\times\operatorname{Par}}$ is a $\mathbb{Q}$-basis of $\Lambda_{\mathbb{Q}} \otimes_{\mathbb{Q}} \Lambda_{\mathbb{Q}}$ (since $(m_\lambda)_{\lambda\in\operatorname{Par}}$ is a $\mathbb{Q}$-basis of $\Lambda_{\mathbb{Q}}$)). Consider these elements $\rho_{\lambda,\mu}$. Since $p = \sum_{(\lambda,\mu)\in\operatorname{Par}\times\operatorname{Par}} \rho_{\lambda,\mu} m_\lambda \otimes m_\mu$, we have $\iota(p) = \sum_{(\lambda,\mu)\in\operatorname{Par}\times\operatorname{Par}} \rho_{\lambda,\mu} m_\lambda(\mathbf{x}) m_\mu(\mathbf{y})$ (by the definition of ι). Therefore, for every $(\alpha,\beta) \in \operatorname{Par}\times\operatorname{Par}$, the rational number $\rho_{\alpha,\beta}$ is the coefficient of the power series $\iota(p)$ before the monomial $\mathbf{x}^\alpha \mathbf{y}^\beta$ (since there is clearly only one term in the sum $\sum_{(\lambda,\mu)\in\operatorname{Par}\times\operatorname{Par}} \rho_{\lambda,\mu} m_\lambda(\mathbf{x}) m_\mu(\mathbf{y})$ contributing to this coefficient, namely the term for $(\lambda,\mu) = (\alpha,\beta)$, and this term contributes $\rho_{\alpha,\beta}$). In particular, this number $\rho_{\alpha,\beta}$ must be an integer (since all coefficients of the power series $\iota(p)$ are integers). So we have shown that $\rho_{\alpha,\beta}$ is an integer for every $(\alpha,\beta) \in \operatorname{Par}\times\operatorname{Par}$. In other words, $\rho_{\lambda,\mu}$ is an integer for every $(\lambda,\mu) \in \operatorname{Par}\times\operatorname{Par}$. Now, $p = \sum_{(\lambda,\mu)\in\operatorname{Par}\times\operatorname{Par}} \underbrace{\rho_{\lambda,\mu}}_{\text{an integer}} m_\lambda \otimes m_\mu \in \Lambda \otimes_{\mathbb{Z}} \Lambda$.

Forget now that we fixed p. We thus have shown that $p \in \Lambda \otimes_{\mathbb{Z}} \Lambda$ for every $p \in \iota^{-1}(\mathbb{Z}[[\mathbf{x},\mathbf{y}]])$. In other words, $\iota^{-1}(\mathbb{Z}[[\mathbf{x},\mathbf{y}]]) \subset \Lambda \otimes_{\mathbb{Z}} \Lambda$. Combined with $\Lambda \otimes_{\mathbb{Z}} \Lambda \subset \iota^{-1}(\mathbb{Z}[[\mathbf{x},\mathbf{y}]])$, this yields $\iota^{-1}(\mathbb{Z}[[\mathbf{x},\mathbf{y}]]) = \Lambda \otimes_{\mathbb{Z}} \Lambda$, qed.

Second solution of Exercise 2.9.4(b): The following solution of Exercise 2.9.4(b) is a variation on the First one given above. It avoids the use of (13.83.2) in favor of working with finite variable sets. (This has the advantage that we no longer need to bother about technicalities; but more importantly, this approach will be very useful in solving Exercise 2.9.4(f) later on.)

Let $N \in \mathbb{N}$ be arbitrary. We define a $\mathbb{Q}$-linear map

$$\begin{aligned} \mathcal{E}_N : \Lambda_{\mathbb{Q}} \otimes_{\mathbb{Q}} \Lambda_{\mathbb{Q}} &\to \mathbb{Q}\left[x_1, x_2, ..., x_N, y_1, y_2, ..., y_N\right], \\ f \otimes g &\mapsto f\left(x_1, x_2, ..., x_N\right) g\left(y_1, y_2, ..., y_N\right). \end{aligned}$$

This is well-defined because $f\left(x_1, x_2, ..., x_N\right) \in \mathbb{Q}\left[x_1, x_2, ..., x_N\right]$ and $g\left(y_1, y_2, ..., y_N\right) \in \mathbb{Q}\left[y_1, y_2, ..., y_N\right]$ are well-defined polynomials for every $f \in \Lambda_{\mathbb{Q}}$ and $g \in \Lambda_{\mathbb{Q}}$ (by Exercise 2.1.2) and depend linearly on f and g, respectively. It is easy to see that the map $\mathcal{E}_N$ is a $\mathbb{Q}$-algebra homomorphism.

Next, we define a map

$$\begin{aligned} \mathcal{K}_N : \Lambda_{\mathbb{Q}} &\to \mathbb{Q}\left[x_1, x_2, ..., x_N, y_1, y_2, ..., y_N\right], \\ f &\mapsto f\left(\left(x_i y_j\right)_{(i,j)\in\{1,2,...,N\}^2}\right). \end{aligned}$$

Here, $f\left(\left(x_i y_j\right)_{(i,j)\in\{1,2,...,N\}^2}\right)$ is defined as follows: Let $\left(u_1, u_2, ..., u_{N^2}\right)$ be a list of all N^2 elements of the family $\left(x_i y_j\right)_{(i,j)\in\{1,2,...,N\}^2}$ in any arbitrary order, and set $f\left(\left(x_i y_j\right)_{(i,j)\in\{1,2,...,N\}^2}\right) = f\left(u_1, u_2, ..., u_{N^2}\right)$. (The result does not depend on the order chosen, because f is symmetric.)

Again, $\mathcal{K}_N$ is a $\mathbb{Q}$-algebra homomorphism (since $\mathcal{K}_N$ is an evaluation map in an appropriate sense).

We now claim that

$$\mathcal{E}_N \circ \Delta_\times = \mathcal{K}_N. \tag{13.83.4}$$

Proof of (13.83.4): The equality (13.83.4) is an equality between $\mathbb{Q}$-algebra homomorphisms (since $\mathcal{K}_N$, $\mathcal{E}_N$ and $\Delta_\times$ are $\mathbb{Q}$-algebra homomorphisms), and thus, in order to prove it, we only need to check that it holds on a generating set of the $\mathbb{Q}$-algebra $\Lambda_{\mathbb{Q}}$. We do this on the generating set $\left(p_n\right)_{n\geq 1}$, by noticing that every $n \geq 1$ satisfies

$$\begin{aligned} \left(\mathcal{E}_N \circ \Delta_\times\right)\left(p_n\right) &= \mathcal{E}_N\left(\underbrace{\Delta_\times\left(p_n\right)}_{=p_n\otimes p_n}\right) = \mathcal{E}_N\left(p_n \otimes p_n\right) \\ &= \underbrace{p_n\left(x_1, x_2, ..., x_N\right)}_{=\sum_{i=1}^N x_i^n} \underbrace{p_n\left(y_1, y_2, ..., y_N\right)}_{=\sum_{j=1}^N y_j^n} \qquad \text{(by the definition of } \mathcal{E}_N\text{)} \\ &= \left(\sum_{i=1}^N x_i^n\right)\left(\sum_{j=1}^N y_j^n\right) = \underbrace{\sum_{i=1}^N \sum_{j=1}^N}_{=\sum_{(i,j)\in\{1,2,...,N\}^2}} \underbrace{x_i^n y_j^n}_{=(x_i y_j)^n} = \sum_{(i,j)\in\{1,2,...,N\}^2} \left(x_i y_j\right)^n \\ &= p_n\left(\left(x_i y_j\right)_{(i,j)\in\{1,2,...,N\}^2}\right) = \mathcal{K}_N\left(p_n\right) \\ &\qquad \left(\text{since } \mathcal{K}_N\left(p_n\right) = p_n\left(\left(x_i y_j\right)_{(i,j)\in\{1,2,...,N\}^2}\right) \text{ (by the definition of } \mathcal{K}_N\left(p_n\right)\text{)}\right). \end{aligned}$$

This proves (13.83.4).

Now,

$$\mathcal{E}_N\left(\Delta_\times\left(\Lambda\right)\right) = \left(\underbrace{\mathcal{E}_N \circ \Delta_\times}_{\substack{=\mathcal{K}_N \\ \text{(by (13.83.4))}}}\right)\left(\Lambda\right) = \mathcal{K}_N\left(\Lambda\right) \subset \mathbb{Z}\left[x_1, x_2, ..., x_N, y_1, y_2, ..., y_N\right]$$

(the latter inclusion is evident from the definition of $\mathcal{K}_N$). Hence, every $p \in \Delta_\times\left(\Lambda\right)$ satisfies

$$\mathcal{E}_N\left(p\right) \in \mathcal{E}_N\left(\Delta_\times\left(\Lambda\right)\right) = \mathbb{Z}\left[x_1, x_2, ..., x_N, y_1, y_2, ..., y_N\right]. \tag{13.83.5}$$

Now, forget that we fixed N. We thus have defined a $\mathbb{Q}$-algebra homomorphism $\mathcal{E}_N : \Lambda_{\mathbb{Q}} \otimes_{\mathbb{Q}} \Lambda_{\mathbb{Q}} \to \mathbb{Q}[x_1, x_2, ..., x_N, y_1, y_2, ..., y_N]$ for every $N \in \mathbb{N}$, and we have shown that every $p \in \Delta_\times(\Lambda)$ satisfies (13.83.5) for every $N \in \mathbb{N}$.

Now, we are going to show that

$$\left(\begin{array}{c}\text{if some } p \in \Lambda_{\mathbb{Q}} \otimes_{\mathbb{Q}} \Lambda_{\mathbb{Q}} \text{ satisfies} \\ \left(\mathcal{E}_N(p) \in \mathbb{Z}[x_1, x_2, ..., x_N, y_1, y_2, ..., y_N] \text{ for every } N \in \mathbb{N}\right), \\ \text{then } p \in \Lambda \otimes_{\mathbb{Z}} \Lambda\end{array}\right) \tag{13.83.6}$$

Proof of (13.83.6): Let $p \in \Lambda_{\mathbb{Q}} \otimes_{\mathbb{Q}} \Lambda_{\mathbb{Q}}$ be such that

$$\left(\mathcal{E}_N(p) \in \mathbb{Z}[x_1, x_2, ..., x_N, y_1, y_2, ..., y_N] \text{ for every } N \in \mathbb{N}\right). \tag{13.83.7}$$

We can write p in the form $p = \sum_{(\lambda,\mu)\in\mathrm{Par}\times\mathrm{Par}} \rho_{\lambda,\mu} m_\lambda \otimes m_\mu$ with $\rho_{\lambda,\mu}$ being elements of $\mathbb{Q}$ (because $(m_\lambda \otimes m_\mu)_{(\lambda,\mu)\in\mathrm{Par}\times\mathrm{Par}}$ is a $\mathbb{Q}$-basis of $\Lambda_{\mathbb{Q}} \otimes_{\mathbb{Q}} \Lambda_{\mathbb{Q}}$ (since $(m_\lambda)_{\lambda\in\mathrm{Par}}$ is a $\mathbb{Q}$-basis of $\Lambda_{\mathbb{Q}}$)). Consider these elements $\rho_{\lambda,\mu}$.

Now, let $(\alpha, \beta) \in \mathrm{Par} \times \mathrm{Par}$ be arbitrary. Choose some $N \in \mathbb{N}$ satisfying $N \geq \ell(\alpha)$ and $N \geq \ell(\beta)$ (such an N clearly exists). Then, $\alpha = (\alpha_1, \alpha_2, ..., \alpha_N)$ and $\beta = (\beta_1, \beta_2, ..., \beta_N)$. Since $p = \sum_{(\lambda,\mu)\in\mathrm{Par}\times\mathrm{Par}} \rho_{\lambda,\mu} m_\lambda \otimes m_\mu$, we have

$$\mathcal{E}_N(p) = \sum_{(\lambda,\mu)\in\mathrm{Par}\times\mathrm{Par}} \rho_{\lambda,\mu} m_\lambda(x_1, x_2, ..., x_N) m_\mu(y_1, y_2, ..., y_N). \tag{13.83.8}$$

The only addend on the right hand side of this equality which has a nonzero coefficient before $x_1^{\alpha_1} x_2^{\alpha_2} ... x_N^{\alpha_N} y_1^{\beta_1} y_2^{\beta_2} ... y_N^{\beta_N}$ is the addend for $(\lambda, \mu) = (\alpha, \beta)$ (since α and β are partitions, so the only partition λ such that the monomial $x_1^{\alpha_1} x_2^{\alpha_2} ... x_N^{\alpha_N}$ appears in $m_\lambda(x_1, x_2, ..., x_N)$ is α, and the only partition μ such that the monomial $y_1^{\beta_1} y_2^{\beta_2} ... y_N^{\beta_N}$ appears in $m_\mu(y_1, y_2, ..., y_N)$ is β). This coefficient is $\rho_{\alpha,\beta}$. Hence, (13.83.8) shows that the coefficient before $x_1^{\alpha_1} x_2^{\alpha_2} ... x_N^{\alpha_N} y_1^{\beta_1} y_2^{\beta_2} ... y_N^{\beta_N}$ in the power series $\mathcal{E}_N(p)$ is $\rho_{\alpha,\beta}$. But this coefficient must be an integer (in fact, (13.83.7) shows that every coefficient of the power series $\mathcal{E}_N(p)$ is an integer). Thus, $\rho_{\alpha,\beta}$ is an integer.

So we have shown that $\rho_{\alpha,\beta}$ is an integer for every $(\alpha, \beta) \in \mathrm{Par} \times \mathrm{Par}$. In other words, $\rho_{\lambda,\mu}$ is an integer for every $(\lambda, \mu) \in \mathrm{Par} \times \mathrm{Par}$. Now, $p = \sum_{(\lambda,\mu)\in\mathrm{Par}\times\mathrm{Par}} \underbrace{\rho_{\lambda,\mu}}_{\text{an integer}} m_\lambda \otimes m_\mu \in \Lambda \otimes_{\mathbb{Z}} \Lambda$. This proves (13.83.6).

Now, we are almost done. Let $p \in \Delta_\times(\Lambda)$. We know that $\mathcal{E}_N(p) \in \mathbb{Z}[x_1, x_2, ..., x_N, y_1, y_2, ..., y_N]$ for every $N \in \mathbb{N}$ (according to (13.83.5)). Hence, (13.83.6) yields that $p \in \Lambda \otimes_{\mathbb{Z}} \Lambda$. Since we have proven this for every $p \in \Delta_\times(\Lambda)$, we thus obtain $\Delta_\times(\Lambda) \subset \Lambda \otimes_{\mathbb{Z}} \Lambda$. This solves Exercise 2.9.4(b).

Third solution of Exercise 2.9.4(b): Here is another solution of Exercise 2.9.4(b), which entirely gets by without using substitutions. We are going to prove that

$$\Delta_\times(h_n) = \sum_{\lambda\vdash n} s_\lambda \otimes s_\lambda \qquad \text{for every } n \in \mathbb{N}. \tag{13.83.9}$$

[657] Once this is proven, it will follow that $\Delta_\times(h_n) \in \Lambda \otimes_{\mathbb{Z}} \Lambda$ for every $n \in \mathbb{N}$, and therefore

$$\begin{aligned}\Delta_\times\left(\underbrace{h_\lambda}_{=h_{\lambda_1}h_{\lambda_2}h_{\lambda_3}...}\right) &= \Delta_\times(h_{\lambda_1}h_{\lambda_2}h_{\lambda_3}...) \\ &= \underbrace{\Delta_\times(h_{\lambda_1})}_{\in\Lambda\otimes_{\mathbb{Z}}\Lambda} \cdot \underbrace{\Delta_\times(h_{\lambda_2})}_{\in\Lambda\otimes_{\mathbb{Z}}\Lambda} \cdot \underbrace{\Delta_\times(h_{\lambda_3})}_{\in\Lambda\otimes_{\mathbb{Z}}\Lambda} \cdot ... \qquad (\text{since } \Delta_\times \text{ is a } \mathbb{Q}\text{-algebra homomorphism}) \\ &\in (\Lambda \otimes_{\mathbb{Z}} \Lambda) \cdot (\Lambda \otimes_{\mathbb{Z}} \Lambda) \cdot (\Lambda \otimes_{\mathbb{Z}} \Lambda) \cdot ... \\ &\subset \Lambda \otimes_{\mathbb{Z}} \Lambda\end{aligned}$$

for every partition λ; and this will immediately yield that $\Delta_\times(\Lambda) \subset \Lambda \otimes_{\mathbb{Z}} \Lambda$ (because $\Delta_\times$ is $\mathbb{Z}$-linear, and because $(h_\lambda)_{\lambda\in\mathrm{Par}}$ is a basis of the $\mathbb{Z}$-module Λ), which will solve Exercise 2.9.4(b). Hence, in order to solve Exercise 2.9.4(b), it is enough to prove (13.83.9).

[657] The notation "$\lambda \vdash n$" here is a synonym for "$\lambda \in \mathrm{Par}_n$".

We need to prove (13.83.9). In order to do so, it is clearly enough to show that

$$\sum_{n\geq 0} \Delta_\times (h_n) t^n = \sum_{n\geq 0} \left(\sum_{\lambda \vdash n} s_\lambda \otimes s_\lambda \right) t^n \qquad \text{in } (\Lambda_{\mathbb{Q}} \otimes_{\mathbb{Q}} \Lambda_{\mathbb{Q}}) [[t]] \tag{13.83.10}$$

(because comparing coefficients in (13.83.10) immediately yields (13.83.9)).

Recall that for any two $\mathbb{Q}$-algebras $\mathcal{A}$ and $\mathcal{B}$, every $\mathbb{Q}$-algebra homomorphism $\varphi : \mathcal{A} \to \mathcal{B}$ induces a continuous[658] $\mathbb{Q}[[t]]$-algebra homomorphism $\varphi[[t]] : \mathcal{A}[[t]] \to \mathcal{B}[[t]]$ which is given by

$$(\varphi[[t]]) \left(\sum_{k\geq 0} a_k t^k \right) = \sum_{k\geq 0} \varphi(a_k) t^k \qquad \text{for every } (a_k)_{k\geq 0} \in \mathcal{A}^{\mathbb{N}}.$$

In particular, the $\mathbb{Q}$-algebra homomorphism $\Delta_\times : \Lambda_{\mathbb{Q}} \to \Lambda_{\mathbb{Q}} \otimes_{\mathbb{Q}} \Lambda_{\mathbb{Q}}$ induces a $\mathbb{Q}[[t]]$-algebra homomorphism $\Delta_\times[[t]] : \Lambda_{\mathbb{Q}}[[t]] \to (\Lambda_{\mathbb{Q}} \otimes_{\mathbb{Q}} \Lambda_{\mathbb{Q}})[[t]]$. Recall the power series $H(t) \in \Lambda[[t]]$ defined in (2.4.1); it satisfies $H(t) = \sum_{n\geq 0} h_n t^n$. Applying the map $\Delta_\times[[t]]$ to both sides of this equality, we obtain

$$(\Delta_\times[[t]])(H(t)) = (\Delta_\times[[t]]) \left(\sum_{n\geq 0} h_n t^n \right) = \sum_{n\geq 0} \Delta_\times (h_n) t^n \tag{13.83.11}$$

(by the definition of $\Delta_\times[[t]]$).

On the other hand, recall the $\mathbb{Q}$-algebra isomorphism $\Lambda_{\mathbb{Q}} \otimes_{\mathbb{Q}} \Lambda_{\mathbb{Q}} \to R_{\mathbb{Q}}(\mathbf{x},\mathbf{y})^{\mathfrak{S}_{(\infty)} \times \mathfrak{S}_{(\infty)}}$ constructed in (2.1.3) (applied to $\mathbf{k} = \mathbb{Q}$). This entails a $\mathbb{Q}$-algebra injection $\Lambda_{\mathbb{Q}} \otimes_{\mathbb{Q}} \Lambda_{\mathbb{Q}} \to \mathbb{Q}[[\mathbf{x},\mathbf{y}]]$ (since $R_{\mathbb{Q}}(\mathbf{x},\mathbf{y})^{\mathfrak{S}_{(\infty)} \times \mathfrak{S}_{(\infty)}} \subset R_{\mathbb{Q}}(\mathbf{x},\mathbf{y}) \subset \mathbb{Q}[[\mathbf{x},\mathbf{y}]]$). Denote this injection by ι. Then, the $\mathbb{Q}$-algebra homomorphism $\iota : \Lambda_{\mathbb{Q}} \otimes_{\mathbb{Q}} \Lambda_{\mathbb{Q}} \to \mathbb{Q}[[\mathbf{x},\mathbf{y}]]$ induces a $\mathbb{Q}[[t]]$-algebra homomorphism $\iota[[t]] : (\Lambda_{\mathbb{Q}} \otimes_{\mathbb{Q}} \Lambda_{\mathbb{Q}})[[t]] \to \mathbb{Q}[[\mathbf{x},\mathbf{y}]][[t]]$. This homomorphism $\iota[[t]]$ is injective (since ι is injective).

We need to prove (13.83.10). For this, it is enough to prove

$$(\iota[[t]]) \left(\sum_{n\geq 0} \Delta_\times (h_n) t^n \right) = (\iota[[t]]) \left(\sum_{n\geq 0} \left(\sum_{\lambda \vdash n} s_\lambda \otimes s_\lambda \right) t^n \right) \tag{13.83.12}$$

(because the injectivity of $\iota[[t]]$ ensures that (13.83.12) implies (13.83.10)).

In $\mathbb{Q}[[\mathbf{x},\mathbf{y}]][[t]]$, we have

$$\begin{aligned}
(\iota[[t]]) \left(\sum_{n\geq 0} \left(\sum_{\lambda \vdash n} s_\lambda \otimes s_\lambda \right) t^n \right) &= \sum_{n\geq 0} \underbrace{\iota \left(\sum_{\lambda \vdash n} s_\lambda \otimes s_\lambda \right)}_{\substack{= \sum_{\lambda \vdash n} s_\lambda(\mathbf{x}) s_\lambda(\mathbf{y}) \\ \text{(by the definition of } \iota\text{)}}} t^n \\
&= \sum_{n\geq 0} \sum_{\lambda \vdash n} s_\lambda(\mathbf{x}) s_\lambda(\mathbf{y}) t^n = \sum_{\lambda \in \mathrm{Par}} s_\lambda(\mathbf{x}) \underbrace{s_\lambda(\mathbf{y}) t^{|\lambda|}}_{\substack{= s_\lambda(ty_1, ty_2, ty_3, \ldots) \\ \text{(since } s_\lambda \text{ is homogeneous of degree } |\lambda|\text{)}}} \\
&= \sum_{\lambda \in \mathrm{Par}} s_\lambda(\mathbf{x}) s_\lambda(ty_1, ty_2, ty_3, \ldots) = \prod_{i,j=1}^{\infty} (1 - x_i \cdot ty_j)^{-1} \qquad (13.83.13) \\
&\left(\begin{array}{c} \text{since (2.5.1) (applied to } (ty_1, ty_2, ty_3, \ldots) \text{ instead of } \mathbf{y}) \\ \text{yields } \prod_{i,j=1}^{\infty} (1 - x_i \cdot ty_j)^{-1} = \sum_{\lambda \in \mathrm{Par}} s_\lambda(\mathbf{x}) s_\lambda(ty_1, ty_2, ty_3, \ldots) \end{array} \right).
\end{aligned}$$

Meanwhile, exponentiating both sides of the equality (2.5.12) yields

$$H(t) = \exp \left(\sum_{m=1}^{\infty} \frac{1}{m} p_m(\mathbf{x}) t^m \right),$$

[658] The word "continuous" refers to the usual topologies on the power series rings $\mathcal{A}[[t]]$ and $\mathcal{B}[[t]]$.

so that

$$
\begin{aligned}
\left(\Delta_{\times}[[t]]\right)(H(t)) &= \left(\Delta_{\times}[[t]]\right)\left(\exp\left(\sum_{m=1}^{\infty}\frac{1}{m}p_m(\mathbf{x})\,t^m\right)\right) = \exp\left(\underbrace{\left(\Delta_{\times}[[t]]\right)\left(\sum_{m=1}^{\infty}\frac{1}{m}p_m(\mathbf{x})\,t^m\right)}_{=\sum_{m=1}^{\infty}\Delta_{\times}\left(\frac{1}{m}p_m(\mathbf{x})\right)t^m}\right)\\
&\qquad\left(\begin{array}{c}\text{since } \Delta_{\times}[[t]] \text{ is a continuous } \mathbb{Q}\text{-algebra homomorphism,}\\ \text{and thus commutes with } \exp\end{array}\right)\\
&= \exp\left(\sum_{m=1}^{\infty}\underbrace{\Delta_{\times}\left(\frac{1}{m}p_m(\mathbf{x})\right)}_{=\dfrac{1}{m}\Delta_{\times}(p_m(\mathbf{x}))}t^m\right) = \exp\left(\sum_{m=1}^{\infty}\frac{1}{m}\underbrace{\Delta_{\times}(p_m(\mathbf{x}))}_{\substack{=\Delta_{\times}(p_m)=p_m\otimes p_m\\ \text{(by the definition of }\Delta_{\times}\text{)}}}t^m\right)\\
&= \exp\left(\sum_{m=1}^{\infty}\frac{1}{m}p_m\otimes p_m t^m\right).
\end{aligned}
$$

Compared with (13.83.11), this yields

$$
\sum_{n\geq 0}\Delta_{\times}(h_n)\,t^n = \exp\left(\sum_{m=1}^{\infty}\frac{1}{m}p_m\otimes p_m t^m\right).
$$

Applying the map $\iota[[t]]$ to both sides of this equality, we obtain

$$
\begin{aligned}
\left(\iota[[t]]\right)\left(\sum_{n\geq 0}\Delta_{\times}(h_n)\,t^n\right) &= \left(\iota[[t]]\right)\left(\exp\left(\sum_{m=1}^{\infty}\frac{1}{m}p_m\otimes p_m t^m\right)\right) = \exp\left(\underbrace{\left(\iota[[t]]\right)\left(\sum_{m=1}^{\infty}\frac{1}{m}p_m\otimes p_m t^m\right)}_{=\sum_{m=1}^{\infty}\iota\left(\frac{1}{m}p_m\otimes p_m\right)t^m}\right)\\
&\qquad\left(\begin{array}{c}\text{since } \iota[[t]] \text{ is a continuous } \mathbb{Q}\text{-algebra homomorphism,}\\ \text{and thus commutes with } \exp\end{array}\right)\\
&= \exp\left(\sum_{m=1}^{\infty}\underbrace{\iota\left(\frac{1}{m}p_m\otimes p_m\right)}_{\substack{=\dfrac{1}{m}p_m(\mathbf{x})p_m(\mathbf{y})\\ \text{(by the definition of }\iota\text{)}}}t^m\right) = \exp\left(\sum_{m=1}^{\infty}\frac{1}{m}p_m(\mathbf{x})\underbrace{p_m(\mathbf{y})\,t^m}_{\substack{=p_m(ty_1,ty_2,ty_3,\ldots)\\ \text{(since } p_m \text{ is homogeneous of}\\ \text{degree } m\text{)}}}\right)\\
&= \exp\left(\sum_{m=1}^{\infty}\frac{1}{m}p_m(\mathbf{x})\,p_m(ty_1,ty_2,ty_3,\ldots)\right). \qquad (13.83.14)
\end{aligned}
$$

But exponentiating both sides of the equality (2.5.14) yields

$$
\prod_{i,j=1}^{\infty}\left(1-x_iy_j\right)^{-1} = \exp\left(\sum_{m=1}^{\infty}\frac{1}{m}p_m(\mathbf{x})\,p_m(\mathbf{y})\right).
$$

Substituting $(ty_1,ty_2,ty_3,\ldots)$ for $\mathbf{y}$ in this equality, we obtain

$$
\prod_{i,j=1}^{\infty}\left(1-x_i\cdot ty_j\right)^{-1} = \exp\left(\sum_{m=1}^{\infty}\frac{1}{m}p_m(\mathbf{x})\,p_m(ty_1,ty_2,ty_3,\ldots)\right).
$$

Compared with (13.83.14), this yields

$$(\iota\left[\left[t\right]\right])\left(\sum_{n\geq 0}\Delta_{\times}\left(h_n\right)t^n\right)=\prod_{i,j=1}^{\infty}\left(1-x_i\cdot ty_j\right)^{-1}.$$

Compared with (13.83.13), this yields

$$(\iota\left[\left[t\right]\right])\left(\sum_{n\geq 0}\Delta_{\times}\left(h_n\right)t^n\right)=(\iota\left[\left[t\right]\right])\left(\sum_{n\geq 0}\left(\sum_{\lambda\vdash n}s_\lambda\otimes s_\lambda\right)t^n\right).$$

This proves (13.83.12). Thus, Exercise 2.9.4(b) is solved.

Remark: The three solutions we gave for Exercise 2.9.4(b) are not that different. The Second solution is a variation on the First solution which trades the use of a substitution of infinitely many variables (with the technical troubles that come along with it) for the inconvenience of having to consider a "sufficiently high $N\in\mathbb{N}$". The Third solution looks like a different beast, but its main idea – the equality (13.83.9) – is really just an afterthought of the First solution. Indeed, knowing the equality (13.83.3) in the First solution, we can easily prove (13.83.9) as follows: Using the notations of the First solution, we have

$$\begin{aligned}\sum_{n\geq 0}h_n\left(\left(x_iy_j\right)_{(i,j)\in\{1,2,3,\ldots\}^2}\right)t^n&=\prod_{i,j=1}^{\infty}\left(1-tx_iy_j\right)^{-1}&&\text{(by (2.2.18), evaluated on the variable set }\mathbf{xy}\text{)}\\&=\sum_{\lambda\in\operatorname{Par}}t^{|\lambda|}s_\lambda\left(\mathbf{x}\right)s_\lambda\left(\mathbf{y}\right)&&\text{(by (2.5.2))}\\&=\sum_{n\in\mathbb{N}}\left(\sum_{\lambda\vdash n}s_\lambda\left(\mathbf{x}\right)s_\lambda\left(\mathbf{y}\right)\right)t^n\end{aligned}$$

in $\mathbb{Q}\left[\left[\mathbf{x},\mathbf{y}\right]\right]\left[\left[t\right]\right]$. Comparing coefficients before t^n in this equality, we obtain

$$h_n\left(\left(x_iy_j\right)_{(i,j)\in\{1,2,3,\ldots\}^2}\right)=\sum_{\lambda\vdash n}s_\lambda\left(\mathbf{x}\right)s_\lambda\left(\mathbf{y}\right)\qquad\text{for every }n\in\mathbb{N}.$$

Thus, for every $n\in\mathbb{N}$, we have

$$\begin{aligned}\iota\left(\Delta_{\times}\left(h_n\right)\right)&=\underbrace{\left(\iota\circ\Delta_{\times}\right)}_{\substack{=\widetilde{\Delta}_{\times}\\\text{(by (13.83.3))}}}\left(h_n\right)=\widetilde{\Delta}_{\times}\left(h_n\right)=h_n\left(\left(x_iy_j\right)_{(i,j)\in\{1,2,3,\ldots\}^2}\right)=\sum_{\lambda\vdash n}s_\lambda\left(\mathbf{x}\right)s_\lambda\left(\mathbf{y}\right)\\&=\iota\left(\sum_{\lambda\vdash n}s_\lambda\otimes s_\lambda\right),\end{aligned}$$

which (by the injectivity of ι) yields $\Delta_{\times}\left(h_n\right)=\sum_{\lambda\vdash n}s_\lambda\otimes s_\lambda$. Thus, (13.83.9) is proven again.

In a similar vein, one can show that

$$\Delta_{\times}\left(e_n\right)=\sum_{\lambda\vdash n}s_\lambda\otimes s_{\lambda^t}\qquad\text{for every }n\in\mathbb{N}.$$

(One would need to use the dual Cauchy identity, i.e., Exercise 2.7.12(a), instead of (2.5.2) this time.)

(c) We are going to show that

$$\epsilon_r\left(h_n\right)=\left(-1\right)^n\binom{-r}{n}\qquad\text{for every }n\in\mathbb{N}.\tag{13.83.15}$$

This will yield the statement of Exercise 2.9.4(c) in the same way as (13.83.9) yielded the statement of Exercise 2.9.4(b) in the Third solution of Exercise 2.9.4(b) above. Hence, we only need to prove (13.83.15) in order to be done with Exercise 2.9.4(c).

Recall that for any two $\mathbb{Q}$-algebras $\mathcal{A}$ and $\mathcal{B}$, every $\mathbb{Q}$-algebra homomorphism $\varphi : \mathcal{A} \to \mathcal{B}$ induces a continuous[659] $\mathbb{Q}[[t]]$-algebra homomorphism $\varphi[[t]] : \mathcal{A}[[t]] \to \mathcal{B}[[t]]$ which is given by

$$(\varphi[[t]])\left(\sum_{k\geq 0} a_k t^k\right) = \sum_{k\geq 0} \varphi(a_k) t^k \qquad \text{for every } (a_k)_{k\geq 0} \in \mathcal{A}^{\mathbb{N}}.$$

Hence, the $\mathbb{Q}$-algebra homomorphism $\epsilon_r : \Lambda_{\mathbb{Q}} \to \mathbb{Q}$ induces a continuous $\mathbb{Q}[[t]]$-algebra homomorphism $\epsilon_r[[t]] : \Lambda_{\mathbb{Q}}[[t]] \to \mathbb{Q}[[t]]$.

Recall the power series $H(t) \in \Lambda[[t]]$ defined in (2.4.1); it satisfies $H(t) = \sum_{n\geq 0} h_n t^n$. Applying the homomorphism $\epsilon_r[[t]]$ to both sides of this equality, we obtain

$$(\epsilon_r[[t]])(H(t)) = (\epsilon_r[[t]])\left(\sum_{n\geq 0} h_n t^n\right) = \sum_{n\geq 0} \epsilon_r(h_n) t^n. \tag{13.83.16}$$

On the other hand, exponentiating both sides of the equality (2.5.12) yields

$$H(t) = \exp\left(\sum_{m=1}^{\infty} \frac{1}{m} p_m(\mathbf{x}) t^m\right).$$

Applying the map $\epsilon_r[[t]]$ to both sides of this equality, we obtain

$$\begin{aligned}
(\epsilon_r[[t]])(H(t)) &= (\epsilon_r[[t]])\left(\exp\left(\sum_{m=1}^{\infty} \frac{1}{m} p_m(\mathbf{x}) t^m\right)\right) = \exp\left(\underbrace{(\epsilon_r[[t]])\left(\sum_{m=1}^{\infty} \frac{1}{m} p_m(\mathbf{x}) t^m\right)}_{=\sum_{m=1}^{\infty} \epsilon_r\left(\frac{1}{m} p_m(\mathbf{x})\right) t^m}\right) \\
&\qquad \left(\begin{array}{c}\text{since } \epsilon_r[[t]] \text{ is a continuous } \mathbb{Q}\text{-algebra homomorphism,} \\ \text{and thus commutes with } \exp\end{array}\right) \\
&= \exp\left(\sum_{m=1}^{\infty} \underbrace{\epsilon_r\left(\frac{1}{m} p_m(\mathbf{x})\right)}_{=\epsilon_r\left(\frac{1}{m} p_m\right) = \frac{1}{m}\epsilon_r(p_m)} t^m\right) = \exp\left(\sum_{m=1}^{\infty} \frac{1}{m} \underbrace{\epsilon_r(p_m)}_{\substack{=r \\ \text{(by the definition of } \epsilon_r)}} t^m\right) \\
&= \exp\left(\sum_{m=1}^{\infty} \frac{1}{m} r t^m\right) = \exp\left(r \cdot \underbrace{\sum_{m=1}^{\infty} \frac{1}{m} t^m}_{=-\log(1-t)}\right) = \exp(r \cdot (-\log(1-t))) \\
&= \left(\underbrace{\exp(-\log(1-t))}_{=(1-t)^{-1}}\right)^r = \left((1-t)^{-1}\right)^r = (1-t)^{-r} = \sum_{n\geq 0} (-1)^n \binom{-r}{n} t^n
\end{aligned}$$

(by Newton's binomial formula). Comparing this with (13.83.16), we obtain

$$\sum_{n\geq 0} \epsilon_r(h_n) t^n = \sum_{n\geq 0} (-1)^n \binom{-r}{n} t^n.$$

Comparing coefficients in this equality of power series, we see that $\epsilon_r(h_n) = (-1)^n \binom{-r}{n}$ for every $n \in \mathbb{N}$. Thus, (13.83.15) is proven, and so Exercise 2.9.4(c) is solved.

[659] The word "continuous" refers to the usual topologies on the power series rings $\mathcal{A}[[t]]$ and $\mathcal{B}[[t]]$.

(d) Consider the $\mathbb{Q}$-algebra homomorphism ϵ_r defined in Exercise 2.9.4(c), and the $\mathbb{Q}$-algebra homomorphism $\Delta_\times$ defined in Exercise 2.9.4(b). We know that both of these maps ϵ_r and $\Delta_\times$ are $\mathbb{Q}$-algebra homomorphisms. In particular, $\operatorname{id} : \Lambda_{\mathbb{Q}} \to \Lambda_{\mathbb{Q}}$ and $\epsilon_r : \Lambda_{\mathbb{Q}} \to \mathbb{Q}$ are $\mathbb{Q}$-algebra homomorphisms. Thus, $\operatorname{id} \otimes \epsilon_r : \Lambda_{\mathbb{Q}} \otimes_{\mathbb{Q}} \Lambda_{\mathbb{Q}} \to \Lambda_{\mathbb{Q}} \otimes_{\mathbb{Q}} \mathbb{Q}$ is a $\mathbb{Q}$-algebra homomorphism (by Exercise 1.3.6(a)). Also, let $\operatorname{can} : \Lambda_{\mathbb{Q}} \otimes_{\mathbb{Q}} \mathbb{Q} \to \Lambda_{\mathbb{Q}}$ be the canonical $\mathbb{Q}$-vector space isomorphism sending every $f \otimes \alpha \in \Lambda_{\mathbb{Q}} \otimes_{\mathbb{Q}} \mathbb{Q}$ to $\alpha f \in \Lambda_{\mathbb{Q}}$.

Now, we claim that

$$\operatorname{can} \circ (\operatorname{id} \otimes \epsilon_r) \circ \Delta_\times = \mathbf{i}_r. \tag{13.83.17}$$

Proof of (13.83.17): The equality (13.83.17) is an equality between $\mathbb{Q}$-algebra homomorphisms (since can, $\operatorname{id} \otimes \epsilon_r$, $\Delta_\times$ and $\mathbf{i}_r$ are $\mathbb{Q}$-algebra homomorphisms). Consequently, in order to prove it, we only need to check that it holds on a generating set of the $\mathbb{Q}$-algebra $\Lambda_{\mathbb{Q}}$. We do this on the generating set $(p_n)_{n \geq 1}$, by noticing that every $n \geq 1$ satisfies

$$\begin{aligned}\left(\operatorname{can} \circ (\operatorname{id} \otimes \epsilon_r) \circ \Delta_\times\right)(p_n) &= \operatorname{can}\left((\operatorname{id} \otimes \epsilon_r)\left(\underbrace{\Delta_\times(p_n)}_{= p_n \otimes p_n}\right)\right) = \operatorname{can}\left(\underbrace{(\operatorname{id} \otimes \epsilon_r)(p_n \otimes p_n)}_{= p_n \otimes \epsilon_r(p_n)}\right) \\ &= \operatorname{can}(p_n \otimes \epsilon_r(p_n)) = \underbrace{\epsilon_r(p_n)}_{= r} p_n = r p_n = \mathbf{i}_r(p_n).\end{aligned}$$

This proves (13.83.17).

Now, $\underbrace{\left(\operatorname{can} \circ (\operatorname{id} \otimes \epsilon_r) \circ \Delta_\times\right)}_{= \mathbf{i}_r}(\Lambda) = \mathbf{i}_r(\Lambda)$, so that

$$\begin{aligned}\mathbf{i}_r(\Lambda) &= \left(\operatorname{can} \circ (\operatorname{id} \otimes \epsilon_r) \circ \Delta_\times\right)(\Lambda) = \operatorname{can}\left((\operatorname{id} \otimes \epsilon_r) \quad \underbrace{(\Delta_\times(\Lambda))}_{\substack{\subset \Lambda \otimes_{\mathbb{Z}} \Lambda \\ \text{(by Exercise 2.9.4(b))}}}\right) \subset \operatorname{can}\left(\underbrace{(\operatorname{id} \otimes \epsilon_r)(\Lambda \otimes_{\mathbb{Z}} \Lambda)}_{= \Lambda \otimes_{\mathbb{Z}} \epsilon_r(\Lambda)}\right) \\ &= \operatorname{can}\left(\Lambda \otimes_{\mathbb{Z}} \quad \underbrace{\epsilon_r(\Lambda)}_{\substack{\subset \mathbb{Z} \\ \text{(by Exercise 2.9.4(c))}}}\right) \subset \operatorname{can}(\Lambda \otimes \mathbb{Z}) = \mathbb{Z} \cdot \Lambda \qquad \text{(by the definition of can)} \\ &= \Lambda.\end{aligned}$$

This solves Exercise 2.9.4(d).

(e) Consider the $\mathbb{Q}$-algebra homomorphism $\Delta_\times$ defined in Exercise 2.9.4(b). We have $\Delta_\times(p_n) = p_n \otimes p_n$ for every positive integer n. Now, it is easy to see that

$$\Delta_\times(p_\lambda) = p_\lambda \otimes p_\lambda \qquad \text{for every partition } \lambda. \tag{13.83.18}$$

[660]

Now, consider the multiplication map $m_{\Lambda_{\mathbb{Q}}} : \Lambda_{\mathbb{Q}} \otimes_{\mathbb{Q}} \Lambda_{\mathbb{Q}} \to \Lambda_{\mathbb{Q}}$ of the $\mathbb{Q}$-algebra $\Lambda_{\mathbb{Q}}$. We claim that

$$\operatorname{Sq} = m_{\Lambda_{\mathbb{Q}}} \circ \Delta_\times.$$

[660] *Proof of (13.83.18):* Let λ be a partition. Write λ in the form $\lambda = (\lambda_1, \lambda_2, ..., \lambda_\ell)$ with $\ell = \ell(\lambda)$. Then, the definition of p_λ yields $p_\lambda = p_{\lambda_1} p_{\lambda_2} ... p_{\lambda_\ell}$. Applying the map $\Delta_\times$ to both sides of this equality, we obtain

$$\begin{aligned}\Delta_\times(p_\lambda) &= \Delta_\times(p_{\lambda_1} p_{\lambda_2} ... p_{\lambda_\ell}) = \Delta_\times(p_{\lambda_1}) \cdot \Delta_\times(p_{\lambda_2}) \cdot ... \cdot \Delta_\times(p_{\lambda_\ell}) && \text{(since } \Delta_\times \text{ is a } \mathbb{Q}\text{-algebra homomorphism)} \\ &= (p_{\lambda_1} \otimes p_{\lambda_1}) \cdot (p_{\lambda_2} \otimes p_{\lambda_2}) \cdot ... \cdot (p_{\lambda_\ell} \otimes p_{\lambda_\ell}) && \text{(since } \Delta_\times(p_n) = p_n \otimes p_n \text{ for every positive integer } n\text{)} \\ &= \underbrace{(p_{\lambda_1} p_{\lambda_2} ... p_{\lambda_\ell})}_{= p_\lambda} \otimes \underbrace{(p_{\lambda_1} p_{\lambda_2} ... p_{\lambda_\ell})}_{= p_\lambda} = p_\lambda \otimes p_\lambda,\end{aligned}$$

which proves (13.83.18).

Indeed, every partition λ satisfies

$$\operatorname{Sq}(p_\lambda) = p_\lambda^2 = p_\lambda p_\lambda = m_{\Lambda_{\mathbb{Q}}}\left(\underbrace{p_\lambda \otimes p_\lambda}_{\substack{=\Delta_\times(p_\lambda) \\ (\text{by } (13.83.18))}}\right) = m_{\Lambda_{\mathbb{Q}}}\left(\Delta_\times(p_\lambda)\right) = \left(m_{\Lambda_{\mathbb{Q}}} \circ \Delta_\times\right)(p_\lambda).$$

Since $(p_\lambda)_{\lambda\in\operatorname{Par}}$ is a $\mathbb{Q}$-basis of $\Lambda_{\mathbb{Q}}$ (and since Sq and $m_{\Lambda_{\mathbb{Q}}} \circ \Delta_\times$ are $\mathbb{Q}$-linear maps), this shows that $\operatorname{Sq} = m_{\Lambda_{\mathbb{Q}}} \circ \Delta_\times$. Hence,

$$\begin{aligned}\operatorname{Sq}(\Lambda) = \left(m_{\Lambda_{\mathbb{Q}}} \circ \Delta_\times\right)(\Lambda) = m_{\Lambda_{\mathbb{Q}}} &\underbrace{\left(\Delta_\times(\Lambda)\right)}_{\substack{\subset \Lambda\otimes_{\mathbb{Z}}\Lambda \\ (\text{by Exercise 2.9.4(b)})}} \subset m_{\Lambda_{\mathbb{Q}}}\left(\Lambda \otimes_{\mathbb{Z}} \Lambda\right) \\ &= \Lambda \cdot \Lambda \qquad \left(\text{by the definition of } m_{\Lambda_{\mathbb{Q}}}\right) \\ &= \Lambda.\end{aligned}$$

This solves Exercise 2.9.4(e).

(f) For every $N \in \mathbb{N}$, we define a $\mathbb{Q}$-algebra homomorphism $\mathcal{E}_N : \Lambda_{\mathbb{Q}} \otimes_{\mathbb{Q}} \Lambda_{\mathbb{Q}} \to \mathbb{Q}[x_1, x_2, ..., x_N, y_1, y_2, ..., y_N]$ as it was done in our Second solution of Exercise 2.9.4(b).

Recall the definition of the maps $\mathbf{i}_r$ in Exercise 2.9.4(d). We are first going to show that every $a \in \mathbb{Z}$ and $b \in \mathbb{Z}$ satisfy

$$\Delta_a = \Delta_b \star \left(\Delta_{\Lambda_{\mathbb{Q}}} \circ \mathbf{i}_{a-b}\right) \qquad \text{in } \operatorname{Hom}\left(\Lambda_{\mathbb{Q}}, \Lambda_{\mathbb{Q}} \otimes_{\mathbb{Q}} \Lambda_{\mathbb{Q}}\right) \tag{13.83.19}$$

(where $\Delta_{\Lambda_{\mathbb{Q}}} : \Lambda_{\mathbb{Q}} \to \Lambda_{\mathbb{Q}} \otimes_{\mathbb{Q}} \Lambda_{\mathbb{Q}}$ is the usual comultiplication of $\Lambda_{\mathbb{Q}}$).

Proof of (13.83.19): Let $a \in \mathbb{Z}$ and $b \in \mathbb{Z}$. We know that Δ_a and Δ_b are $\mathbb{Q}$-algebra homomorphisms. Also, $\Delta_{\Lambda_{\mathbb{Q}}}$ is a $\mathbb{Q}$-algebra homomorphism (by the axioms of a bialgebra, which we know are satisfied for $\Lambda_{\mathbb{Q}}$), and $\mathbf{i}_{a-b}$ is a $\mathbb{Q}$-algebra homomorphism. Hence, the composition $\Delta_{\Lambda_{\mathbb{Q}}} \circ \mathbf{i}_{a-b}$ is a $\mathbb{Q}$-algebra homomorphism. Thus, the convolution $\Delta_b \star \left(\Delta_{\Lambda_{\mathbb{Q}}} \circ \mathbf{i}_{a-b}\right)$ is a $\mathbb{Q}$-algebra homomorphism (by Exercise 1.5.11(a), applied to $\mathbb{Q}$, $\Lambda_{\mathbb{Q}}$, $\Lambda_{\mathbb{Q}} \otimes_{\mathbb{Q}} \Lambda_{\mathbb{Q}}$, Δ_b and $\Delta_{\Lambda_{\mathbb{Q}}} \circ \mathbf{i}_{a-b}$ instead of $\mathbf{k}$, H, A, f and g). Hence, (13.83.19) is an equality between $\mathbb{Q}$-algebra homomorphisms. Therefore, in order to prove it, we only need to check that it holds on a generating set of the $\mathbb{Q}$-algebra $\Lambda_{\mathbb{Q}}$. We do this on the generating set $(p_n)_{n\geq 1}$, by noticing that every $n \geq 1$ satisfies

$$\begin{aligned}
&\underbrace{\left(\Delta_b \star \left(\Delta_{\Lambda_{\mathbb{Q}}} \circ \mathbf{i}_{a-b}\right)\right)}_{\substack{= m_{\Lambda_{\mathbb{Q}}} \circ \left(\Delta_b \otimes \left(\Delta_{\Lambda_{\mathbb{Q}}} \circ \mathbf{i}_{a-b}\right)\right) \circ \Delta_{\Lambda_{\mathbb{Q}}} \\ \text{(by the definition of convolution)}}} (p_n) \\
&= \left(m_{\Lambda_{\mathbb{Q}}} \circ \left(\Delta_b \otimes \left(\Delta_{\Lambda_{\mathbb{Q}}} \circ \mathbf{i}_{a-b}\right)\right) \circ \Delta_{\Lambda_{\mathbb{Q}}}\right) (p_n) \\
&= m_{\Lambda_{\mathbb{Q}}} \left(\left(\Delta_b \otimes \left(\Delta_{\Lambda_{\mathbb{Q}}} \circ \mathbf{i}_{a-b}\right)\right) \left(\underbrace{\Delta_{\Lambda_{\mathbb{Q}}} (p_n)}_{\substack{= 1 \otimes p_n + p_n \otimes 1 \\ \text{(since } p_n \text{ is primitive)}}} \right) \right) \\
&= m_{\Lambda_{\mathbb{Q}}} \left(\underbrace{\left(\Delta_b \otimes \left(\Delta_{\Lambda_{\mathbb{Q}}} \circ \mathbf{i}_{a-b}\right)\right) (1 \otimes p_n + p_n \otimes 1)}_{= \Delta_b(1) \otimes \left(\Delta_{\Lambda_{\mathbb{Q}}} \circ \mathbf{i}_{a-b}\right)(p_n) + \Delta_b(p_n) \otimes \left(\Delta_{\Lambda_{\mathbb{Q}}} \circ \mathbf{i}_{a-b}\right)(1)} \right) \\
&= m_{\Lambda_{\mathbb{Q}}} \left(\Delta_b (1) \otimes \left(\Delta_{\Lambda_{\mathbb{Q}}} \circ \mathbf{i}_{a-b}\right) (p_n) + \Delta_b (p_n) \otimes \left(\Delta_{\Lambda_{\mathbb{Q}}} \circ \mathbf{i}_{a-b}\right) (1) \right) \\
&= \underbrace{\Delta_b (1)}_{= 1 \otimes 1} \cdot \underbrace{\left(\Delta_{\Lambda_{\mathbb{Q}}} \circ \mathbf{i}_{a-b}\right) (p_n)}_{= \Delta_{\Lambda_{\mathbb{Q}}} (\mathbf{i}_{a-b}(p_n))} + \Delta_b (p_n) \cdot \underbrace{\left(\Delta_{\Lambda_{\mathbb{Q}}} \circ \mathbf{i}_{a-b}\right) (1)}_{\substack{= 1 \otimes 1 \\ \text{(since } \Delta_{\Lambda_{\mathbb{Q}}} \circ \mathbf{i}_{a-b} \text{ is a } \mathbb{Q}\text{-algebra} \\ \text{homomorphism)}}} \\
&= (1 \otimes 1) \cdot \left(\Delta_{\Lambda_{\mathbb{Q}}} \left(\mathbf{i}_{a-b} (p_n)\right)\right) + \Delta_b (p_n) \cdot (1 \otimes 1) = \Delta_{\Lambda_{\mathbb{Q}}} \left(\underbrace{\mathbf{i}_{a-b} (p_n)}_{\substack{= (a-b) p_n \\ \text{(by the definition of } \mathbf{i}_{a-b})}} \right) + \underbrace{\Delta_b (p_n)}_{\substack{= \sum_{i=1}^{n-1} \binom{n}{i} p_i \otimes p_{n-i} + b \otimes p_n + p_n \otimes b \\ \text{(by the definition of } \Delta_b)}} \\
&= \Delta_{\Lambda_{\mathbb{Q}}} \left((a-b) p_n\right) + \sum_{i=1}^{n-1} \binom{n}{i} p_i \otimes p_{n-i} + b \otimes p_n + p_n \otimes b \\
&= (a-b) \underbrace{\Delta_{\Lambda_{\mathbb{Q}}} (p_n)}_{\substack{= 1 \otimes p_n + p_n \otimes 1 \\ \text{(since } p_n \text{ is primitive)}}} + \sum_{i=1}^{n-1} \binom{n}{i} p_i \otimes p_{n-i} + b \otimes p_n + p_n \otimes b \\
&= (a-b) (1 \otimes p_n + p_n \otimes 1) + \sum_{i=1}^{n-1} \binom{n}{i} p_i \otimes p_{n-i} + b \otimes p_n + p_n \otimes b \\
&= \sum_{i=1}^{n-1} \binom{n}{i} p_i \otimes p_{n-i} + \underbrace{b \otimes p_n + p_n \otimes b + (a-b) (1 \otimes p_n + p_n \otimes 1)}_{= a \otimes p_n + p_n \otimes a} \\
&= \sum_{i=1}^{n-1} \binom{n}{i} p_i \otimes p_{n-i} + a \otimes p_n + p_n \otimes a = \Delta_a (p_n) \\
&\qquad \left(\text{since the definition of } \Delta_a \text{ yields } \Delta_a (p_n) = \sum_{i=1}^{n-1} \binom{n}{i} p_i \otimes p_{n-i} + a \otimes p_n + p_n \otimes a \right).
\end{aligned}$$

This proves (13.83.19).

Before we move on, let us record a simple fact about convolution of maps. Namely, if $\mathbf{k}$ is a commutative ring, and C is a $\mathbf{k}$-coalgebra, and A and A' are two $\mathbf{k}$-algebras, and f and g are two $\mathbf{k}$-linear maps $C \to A$, and $\alpha : A \to A'$ is a $\mathbf{k}$-algebra homomorphism, then

$$\alpha \circ (f \star g) = (\alpha \circ f) \star (\alpha \circ g) . \tag{13.83.20}$$

This is merely the particular case of (1.4.2) when $C' = C$ and $\gamma = \operatorname{id}$, and so requires no proof anymore.

Now, let $N \in \mathbb{N}$. Our next goal is to show that

$$(\mathcal{E}_N \circ \Delta_N)(\Lambda) \subset \mathbb{Z}[x_1, x_2, ..., x_N, y_1, y_2, ..., y_N]. \tag{13.83.21}$$

Proof of (13.83.21): Let us define a map

$$\begin{aligned}\mathcal{L}_N : \Lambda_{\mathbb{Q}} &\to \mathbb{Q}[x_1, x_2, ..., x_N, y_1, y_2, ..., y_N], \\ f &\mapsto f\left((x_i + y_j)_{(i,j)\in\{1,2,...,N\}^2}\right).\end{aligned}$$

Here, $f\left((x_i + y_j)_{(i,j)\in\{1,2,...,N\}^2}\right)$ is defined as follows: Let $(u_1, u_2, ..., u_{N^2})$ be a list of all N^2 elements of the family $(x_i + y_j)_{(i,j)\in\{1,2,...,N\}^2}$ in any arbitrary order, and set $f\left((x_i + y_j)_{(i,j)\in\{1,2,...,N\}^2}\right) = f(u_1, u_2, ..., u_{N^2})$. (The result does not depend on the order chosen, because f is symmetric.)

The map $\mathcal{L}_N$ is a $\mathbb{Q}$-algebra homomorphism (since $\mathcal{L}_N$ is an evaluation map in an appropriate sense).

For every positive integer n, we have

$$\begin{aligned}\Delta_N(p_n) &= \sum_{i=1}^{n-1} \binom{n}{i} p_i \otimes p_{n-i} + N \otimes p_n + p_n \otimes N \qquad \text{(by the definition of } \Delta_N(p_n)\text{)} \\ &= \sum_{k=1}^{n-1} \binom{n}{k} p_k \otimes p_{n-k} + N \otimes p_n + p_n \otimes N\end{aligned}$$

and thus

$$\begin{aligned}&(\mathcal{E}_N \circ \Delta_N)(p_n) \\ &= \mathcal{E}_N\left(\underbrace{\Delta_N(p_n)}_{=\sum_{k=1}^{n-1}\binom{n}{k} p_k \otimes p_{n-k} + N \otimes p_n + p_n \otimes N}\right) = \mathcal{E}_N\left(\sum_{k=1}^{n-1} \binom{n}{k} p_k \otimes p_{n-k} + N \otimes p_n + p_n \otimes N\right) \\ &= \sum_{k=1}^{n-1} \binom{n}{k} \underbrace{p_k(x_1, x_2, ..., x_N)}_{=\sum_{i=1}^N x_i^k} \underbrace{p_{n-k}(y_1, y_2, ..., y_N)}_{=\sum_{j=1}^N y_j^{n-k}} + N \underbrace{p_n(y_1, y_2, ..., y_N)}_{=\sum_{j=1}^N y_j^n} + \underbrace{p_n(x_1, x_2, ..., x_N)}_{=\sum_{i=1}^N x_i^n} N \\ &\qquad \text{(by the definition of } \mathcal{E}_N\text{)} \\ &= \sum_{k=1}^{n-1} \binom{n}{k} \left(\sum_{i=1}^N x_i^k\right)\left(\sum_{j=1}^N y_j^{n-k}\right) + N\left(\sum_{j=1}^N y_j^n\right) + \left(\sum_{i=1}^N x_i^n\right) N, \qquad (13.83.22)\end{aligned}$$

while at the same time

$$\begin{aligned}
\mathcal{L}_N(p_n) &= p_n\left((x_i+y_j)_{(i,j)\in\{1,2,\ldots,N\}^2}\right) \qquad \text{(by the definition of } \mathcal{L}_N(p_n)\text{)} \\
&= \sum_{(i,j)\in\{1,2,\ldots,N\}^2} \underbrace{(x_i+y_j)^n}_{=\sum_{k=0}^n \binom{n}{k} x_i^k y_j^{n-k}} = \sum_{(i,j)\in\{1,2,\ldots,N\}^2} \sum_{k=0}^n \binom{n}{k} x_i^k y_j^{n-k} \\
&= \sum_{k=0}^n \binom{n}{k} \underbrace{\sum_{(i,j)\in\{1,2,\ldots,N\}^2} x_i^k y_j^{n-k}}_{=\left(\sum_{i=1}^N x_i^k\right)\left(\sum_{j=1}^N y_j^{n-k}\right)} = \sum_{k=0}^n \binom{n}{k} \left(\sum_{i=1}^N x_i^k\right)\left(\sum_{j=1}^N y_j^{n-k}\right) \\
&= \sum_{k=1}^{n-1} \binom{n}{k} \left(\sum_{i=1}^N x_i^k\right)\left(\sum_{j=1}^N y_j^{n-k}\right) + \underbrace{\binom{n}{0}}_{=1} \underbrace{\left(\sum_{i=1}^N x_i^0\right)}_{=N} \left(\sum_{j=1}^N y_j^n\right) + \underbrace{\binom{n}{n}}_{=1} \left(\sum_{i=1}^N x_i^n\right) \underbrace{\left(\sum_{j=1}^N y_j^0\right)}_{=N} \\
&= \sum_{k=1}^{n-1} \binom{n}{k} \left(\sum_{i=1}^N x_i^k\right)\left(\sum_{j=1}^N y_j^{n-k}\right) + N\left(\sum_{j=1}^N y_j^n\right) + \left(\sum_{i=1}^N x_i^n\right) N. \qquad (13.83.23)
\end{aligned}$$

Comparing (13.83.22) with (13.83.23) reveals that $(\mathcal{E}_N \circ \Delta_N)(p_n) = \mathcal{L}_N(p_n)$ for every positive integer n. In other words, the two maps $\mathcal{E}_N \circ \Delta_N$ and $\mathcal{L}_N$ are equal to each other on the generating set $(p_n)_{n\geq 1}$ of the $\mathbb{Q}$-algebra $\Lambda_{\mathbb{Q}}$. Since these two maps $\mathcal{E}_N \circ \Delta_N$ and $\mathcal{L}_N$ are $\mathbb{Q}$-algebra homomorphisms (since $\mathcal{E}_N$, Δ_N and $\mathcal{L}_N$ are $\mathbb{Q}$-algebra homomorphisms), this yields that these two maps must be identical, i.e., we have $\mathcal{E}_N \circ \Delta_N = \mathcal{L}_N$. Hence,

$$\underbrace{(\mathcal{E}_N \circ \Delta_N)}_{=\mathcal{L}_N}(\Lambda) = \mathcal{L}_N(\Lambda) \subset \mathbb{Z}[x_1,x_2,\ldots,x_N,y_1,y_2,\ldots,y_N]$$

(because the definition of $\mathcal{L}_N$ immediately shows that $\mathcal{L}_N(f) \in \mathbb{Z}[x_1,x_2,\ldots,x_N,y_1,y_2,\ldots,y_N]$ for every $f \in \Lambda$). This proves (13.83.21).

Next, we are going to show that

$$(\mathcal{E}_N \circ \Delta_r)(\Lambda) \subset \mathbb{Z}[x_1,x_2,\ldots,x_N,y_1,y_2,\ldots,y_N]. \qquad (13.83.24)$$

Proof of (13.83.24): Applying (13.83.19) to $a = r$ and $b = N$, we obtain

$$\Delta_r = \Delta_N \star \left(\Delta_{\Lambda_{\mathbb{Q}}} \circ \mathbf{i}_{r-N}\right).$$

Thus,

$$\mathcal{E}_N \circ \underbrace{\Delta_r}_{=\Delta_N \star \left(\Delta_{\Lambda_{\mathbb{Q}}} \circ \mathbf{i}_{r-N}\right)} = \mathcal{E}_N \circ \left(\Delta_N \star \left(\Delta_{\Lambda_{\mathbb{Q}}} \circ \mathbf{i}_{r-N}\right)\right) = (\mathcal{E}_N \circ \Delta_N) \star \left(\mathcal{E}_N \circ \Delta_{\Lambda_{\mathbb{Q}}} \circ \mathbf{i}_{r-N}\right)$$

(by (13.83.20), applied to $\mathbb{Q}$, $\Lambda_{\mathbb{Q}}$, $\Lambda_{\mathbb{Q}} \otimes_{\mathbb{Q}} \Lambda_{\mathbb{Q}}$, $\mathbb{Q}[x_1,x_2,\ldots,x_N,y_1,y_2,\ldots,y_N]$, Δ_N, $\Delta_{\Lambda_{\mathbb{Q}}} \circ \mathbf{i}_{r-N}$ and $\mathcal{E}_N$ instead of $\mathbf{k}$, C, A, A', f, g and α). Thus,

$$\mathcal{E}_N \circ \Delta_r = (\mathcal{E}_N \circ \Delta_N) \star \left(\mathcal{E}_N \circ \Delta_{\Lambda_{\mathbb{Q}}} \circ \mathbf{i}_{r-N}\right) = m_{\mathbb{Q}[x_1,x_2,\ldots,x_N,y_1,y_2,\ldots,y_N]} \circ \left((\mathcal{E}_N \circ \Delta_N) \otimes \left(\mathcal{E}_N \circ \Delta_{\Lambda_{\mathbb{Q}}} \circ \mathbf{i}_{r-N}\right)\right) \circ \Delta_{\Lambda_{\mathbb{Q}}}$$

(by the definition of convolution), so that

$$\begin{aligned}
&\underbrace{\left(\mathcal{E}_N \circ \Delta_r\right)}_{=m_{\mathbb{Q}[x_1,x_2,\ldots,x_N,y_1,y_2,\ldots,y_N]}\circ\left(\left(\mathcal{E}_N\circ\Delta_N\right)\otimes\left(\mathcal{E}_N\circ\Delta_{\Lambda_{\mathbb{Q}}}\circ\mathbf{i}_{r-N}\right)\right)\circ\Delta_{\Lambda_{\mathbb{Q}}}} (\Lambda) \\
&= \left(m_{\mathbb{Q}[x_1,x_2,\ldots,x_N,y_1,y_2,\ldots,y_N]} \circ \left(\left(\mathcal{E}_N \circ \Delta_N\right) \otimes \left(\mathcal{E}_N \circ \Delta_{\Lambda_{\mathbb{Q}}} \circ \mathbf{i}_{r-N}\right)\right) \circ \Delta_{\Lambda_{\mathbb{Q}}}\right)(\Lambda) \\
&= m_{\mathbb{Q}[x_1,x_2,\ldots,x_N,y_1,y_2,\ldots,y_N]}\left(\left(\left(\mathcal{E}_N \circ \Delta_N\right) \otimes \left(\mathcal{E}_N \circ \Delta_{\Lambda_{\mathbb{Q}}} \circ \mathbf{i}_{r-N}\right)\right) \underbrace{\left(\Delta_{\Lambda_{\mathbb{Q}}}(\Lambda)\right)}_{=\Delta_\Lambda(\Lambda)\subset\Lambda\otimes_{\mathbb{Z}}\Lambda}\right) \\
&\subset m_{\mathbb{Q}[x_1,x_2,\ldots,x_N,y_1,y_2,\ldots,y_N]}\left(\underbrace{\left(\left(\mathcal{E}_N \circ \Delta_N\right) \otimes \left(\mathcal{E}_N \circ \Delta_{\Lambda_{\mathbb{Q}}} \circ \mathbf{i}_{r-N}\right)\right)\left(\Lambda\otimes_{\mathbb{Z}}\Lambda\right)}_{=\left(\mathcal{E}_N\circ\Delta_N\right)(\Lambda)\otimes_{\mathbb{Z}}\left(\mathcal{E}_N\circ\Delta_{\Lambda_{\mathbb{Q}}}\circ\mathbf{i}_{r-N}\right)(\Lambda)}\right) \\
&= m_{\mathbb{Q}[x_1,x_2,\ldots,x_N,y_1,y_2,\ldots,y_N]}\left(\left(\mathcal{E}_N \circ \Delta_N\right)(\Lambda) \otimes_{\mathbb{Z}} \left(\mathcal{E}_N \circ \Delta_{\Lambda_{\mathbb{Q}}} \circ \mathbf{i}_{r-N}\right)(\Lambda)\right) \\
&= \left(\mathcal{E}_N \circ \Delta_N\right)(\Lambda) \cdot \underbrace{\left(\mathcal{E}_N \circ \Delta_{\Lambda_{\mathbb{Q}}} \circ \mathbf{i}_{r-N}\right)(\Lambda)}_{=\mathcal{E}_N\left(\Delta_{\Lambda_{\mathbb{Q}}}\left(\mathbf{i}_{r-N}(\Lambda)\right)\right)} \qquad \left(\text{since } m_{\mathbb{Q}[x_1,x_2,\ldots,x_N,y_1,y_2,\ldots,y_N]} \text{ is the multiplication map}\right) \\
&= \left(\mathcal{E}_N \circ \Delta_N\right)(\Lambda) \cdot \mathcal{E}_N\left(\Delta_{\Lambda_{\mathbb{Q}}}\left(\underbrace{\mathbf{i}_{r-N}(\Lambda)}_{\substack{\subset\Lambda \\ \text{(by Exercise 2.9.4(d),} \\ \text{applied to } r-N \text{ instead of } r)}}\right)\right) \\
&\subset \underbrace{\left(\mathcal{E}_N \circ \Delta_N\right)(\Lambda)}_{\substack{\subset\mathbb{Z}[x_1,x_2,\ldots,x_N,y_1,y_2,\ldots,y_N] \\ \text{(by (13.83.21))}}} \cdot \mathcal{E}_N \underbrace{\left(\Delta_{\Lambda_{\mathbb{Q}}}(\Lambda)\right)}_{=\Delta_\Lambda(\Lambda)\subset\Lambda\otimes_{\mathbb{Z}}\Lambda} \\
&\subset \mathbb{Z}\left[x_1,x_2,\ldots,x_N,y_1,y_2,\ldots,y_N\right] \cdot \underbrace{\mathcal{E}_N\left(\Lambda\otimes_{\mathbb{Z}}\Lambda\right)}_{\substack{\subset\mathbb{Z}[x_1,x_2,\ldots,x_N,y_1,y_2,\ldots,y_N] \\ \text{(because the definition of } \mathcal{E}_N \text{ immediately yields} \\ \mathcal{E}_N(f\otimes g)\in\mathbb{Z}[x_1,x_2,\ldots,x_N,y_1,y_2,\ldots,y_N] \text{ for any } f\in\Lambda \text{ and } g\in\Lambda)}} \\
&\subset \mathbb{Z}\left[x_1,x_2,\ldots,x_N,y_1,y_2,\ldots,y_N\right] \cdot \mathbb{Z}\left[x_1,x_2,\ldots,x_N,y_1,y_2,\ldots,y_N\right] = \mathbb{Z}\left[x_1,x_2,\ldots,x_N,y_1,y_2,\ldots,y_N\right].
\end{aligned}$$

This proves (13.83.24).

Now, forget that we fixed N. We thus have proven (13.83.24) to hold for every $N \in \mathbb{N}$.

Now, let $p \in \Delta_r(\Lambda)$. Then,

$$\mathcal{E}_N\left(\underbrace{p}_{\in\Delta_r(\Lambda)}\right) \in \mathcal{E}_N\left(\Delta_r(\Lambda)\right) = \left(\mathcal{E}_N \circ \Delta_r\right)(\Lambda) \in \mathbb{Z}\left[x_1,x_2,\ldots,x_N,y_1,y_2,\ldots,y_N\right] \qquad \text{(by (13.83.24))}$$

for every $N \in \mathbb{N}$. Hence, (13.83.6) yields that $p \in \Lambda \otimes_{\mathbb{Z}} \Lambda$. Since we have shown this for every $p \in \Delta_r(\Lambda)$, we thus conclude that $\Delta_r(\Lambda) \subset \Lambda \otimes_{\mathbb{Z}} \Lambda$. This solves Exercise 2.9.4(f).

Remark. Here is a rough sketch of an alternative way to conclude this solution of Exercise 2.9.4(f) after proving (13.83.21). This is closer to Richard Stanley's suggested solution than the above.

The family $(m_\lambda \otimes m_\mu)_{(\lambda,\mu)\in\operatorname{Par}\times\operatorname{Par}}$ is a $\mathbb{Q}$-basis of $\Lambda_{\mathbb{Q}} \otimes_{\mathbb{Q}} \Lambda_{\mathbb{Q}}$ (since $(m_\lambda)_{\lambda\in\operatorname{Par}}$ is a $\mathbb{Q}$-basis of $\Lambda_{\mathbb{Q}}$). Let us refer to this basis as the *monomial basis* of $\Lambda_{\mathbb{Q}} \otimes_{\mathbb{Q}} \Lambda_{\mathbb{Q}}$. Then, $\Lambda \otimes_{\mathbb{Z}} \Lambda$ is the subset of $\Lambda_{\mathbb{Q}} \otimes_{\mathbb{Q}} \Lambda_{\mathbb{Q}}$ consisting of all elements whose coordinates with respect to the monomial basis all are integers.

We want to prove that $\Delta_r(\Lambda) \subset \Lambda \otimes_{\mathbb{Z}} \Lambda$ for every $r \in \mathbb{Z}$. In other words, we want to prove that $\Delta_r(f) \in \Lambda \otimes_{\mathbb{Z}} \Lambda$ for every $r \in \mathbb{Z}$ and $f \in \Lambda$. Let us fix $f \in \Lambda$, but not fix r. We need to show that $\Delta_r(f) \in \Lambda \otimes_{\mathbb{Z}} \Lambda$ for every $r \in \mathbb{Z}$; in other words, we need to show that for every $(\alpha, \beta) \in \operatorname{Par}$, the $m_\alpha \otimes m_\beta$-coordinate of $\Delta_r(f)$ with respect to the monomial basis is an integer for every $r \in \mathbb{Z}$.

So let us fix $(\alpha, \beta) \in \operatorname{Par}$. It is not hard to see that the $m_\alpha \otimes m_\beta$-coordinate of $\Delta_r(f)$ with respect to the monomial basis is a polynomial in r with rational coefficients.[661] We want to prove that this polynomial is integer-valued (i.e., all its values at integer inputs are integers[662]). To do so, it suffices to show that its values are integers **at all sufficiently high** $r \in \mathbb{Z}$ (because for a polynomial with rational coefficients, the non-integer values appear periodically[663], and therefore if no non-integer values appear from a given integer onwards, then the polynomial is integer-valued). In other words, it suffices to show that the $m_\alpha \otimes m_\beta$-coordinate of $\Delta_r(f)$ with respect to the monomial basis is an integer for all **sufficiently high** $r \in \mathbb{Z}$.

Let us prove this now. Our interpretation of "sufficiently high" will be that $r \geq \ell(\alpha)$ and $r \geq \ell(\beta)$. So what we need to prove is that the $m_\alpha \otimes m_\beta$-coordinate of $\Delta_r(f)$ with respect to the monomial basis is an integer whenever r is an integer satisfying $r \geq \ell(\alpha)$ and $r \geq \ell(\beta)$.

Consider such an r. Set $N = r$ and $p = \Delta_r(f)$. Then, $N = r \geq \ell(\alpha)$ and $N = r \geq \ell(\beta)$ and $\Delta_N(f) = \Delta_r(f) = p$. Then,

$$\mathcal{E}_N\left(\underbrace{p}_{=\Delta_N(f)}\right) = \mathcal{E}_N(\Delta_N(f)) = (\mathcal{E}_N \circ \Delta_N)\left(\underbrace{f}_{\in\Lambda}\right) \in (\mathcal{E}_N \circ \Delta_N)(\Lambda) \subset \mathbb{Z}[x_1, x_2, ..., x_N, y_1, y_2, ..., y_N]$$

(by (13.83.21)). Therefore, arguing precisely as in the proof of (13.83.6) (but using our specific α and β rather than arbitrary α and β as in that proof), we can show that $\rho_{\alpha,\beta}$ is an integer, where p is written in the form $p = \sum_{(\lambda,\mu)\in \operatorname{Par}\times\operatorname{Par}} \rho_{\lambda,\mu} m_\lambda \otimes m_\mu$ with $\rho_{\lambda,\mu}$ being elements of $\mathbb{Q}$. But $\rho_{\alpha,\beta}$ is precisely the $m_\alpha \otimes m_\beta$-coordinate of $\Delta_r(f)$ with respect to the monomial basis. Thus, we have shown that the $m_\alpha \otimes m_\beta$-coordinate of $\Delta_r(f)$ with respect to the monomial basis is an integer (for $r \in \mathbb{Z}$ satisfying $r \geq \ell(\alpha)$ and $r \geq \ell(\beta)$). As we know, this completes our solution of Exercise 2.9.4(f) again.

(g) Let us first show that the $\mathbb{Q}$-module $\Lambda_\mathbb{Q}$, endowed with the comultiplication $\Delta_\times$ and the counit ϵ_1, becomes a $\mathbb{Q}$-coalgebra. In order to do so, we must verifying that the diagrams (1.2.1) and (1.2.2), with C, Δ and ϵ replaced by $\Lambda_\mathbb{Q}$, $\Delta_\times$ and ϵ_1, commute. We will only do this for the diagram (1.2.1), while leaving the diagram (1.2.2) to the reader.

So we must check that the diagram (1.2.1), with C, Δ and ϵ replaced by $\Lambda_\mathbb{Q}$, $\Delta_\times$ and ϵ_1, commutes. In other words, we must prove the identity

$$(\Delta_\times \otimes \operatorname{id}) \circ \Delta_\times = (\operatorname{id} \otimes \Delta_\times) \circ \Delta_\times . \tag{13.83.25}$$

Proof of (13.83.17): The equality (13.83.25) is an equality between $\mathbb{Q}$-algebra homomorphisms (since $\Delta_\times$, $\operatorname{id} \otimes \Delta_\times$ and $\Delta_\times \otimes \operatorname{id}$ are $\mathbb{Q}$-algebra homomorphisms[664]). Consequently, in order to prove it, we only need to check that it holds on a generating set of the $\mathbb{Q}$-algebra $\Lambda_\mathbb{Q}$. But every $n \geq 1$ satisfies

$$((\Delta_\times \otimes \operatorname{id}) \circ \Delta_\times)(p_n) = (\Delta_\times \otimes \operatorname{id})\left(\underbrace{\Delta_\times(p_n)}_{=p_n \otimes p_n}\right) = (\Delta_\times \otimes \operatorname{id})(p_n \otimes p_n) = \underbrace{\Delta_\times(p_n)}_{=p_n \otimes p_n} \otimes p_n = p_n \otimes p_n \otimes p_n$$

and similarly $((\operatorname{id} \otimes \Delta_\times) \circ \Delta_\times)(p_n) = p_n \otimes p_n \otimes p_n$. Hence, every $n \geq 1$ satisfies $((\Delta_\times \otimes \operatorname{id}) \circ \Delta_\times)(p_n) = p_n \otimes p_n \otimes p_n = ((\operatorname{id} \otimes \Delta_\times) \circ \Delta_\times)(p_n)$. Thus, we have checked that the equality (13.83.25) holds on a generating set of the $\mathbb{Q}$-algebra $\Lambda_\mathbb{Q}$ (namely, on the generating set $(p_n)_{n\geq 1}$). As a consequence, the proof of (13.83.25) is complete.

We have thus shown that the $\mathbb{Q}$-module $\Lambda_\mathbb{Q}$, endowed with the comultiplication $\Delta_\times$ and the counit ϵ_1, becomes a $\mathbb{Q}$-coalgebra. This $\mathbb{Q}$-coalgebra becomes a $\mathbb{Q}$-bialgebra when combined with the existing $\mathbb{Q}$-algebra structure on $\Lambda_\mathbb{Q}$ (this is because $\Delta_\times$ and ϵ_1 are $\mathbb{Q}$-algebra homomorphisms), and this $\mathbb{Q}$-bialgebra is cocommutative (this follows from the equality $T \circ \Delta_\times = \Delta_\times$, where $T : \Lambda_\mathbb{Q} \otimes_\mathbb{Q} \Lambda_\mathbb{Q} \to \Lambda_\mathbb{Q} \otimes_\mathbb{Q} \Lambda_\mathbb{Q}$ is the twist map; and this equality can be proven in the same way as we have showed (13.83.25)). This solves Exercise 2.9.4(g).

[661] This can be proven by noticing that it holds whenever $f = p_n$ for some $n \geq 1$ (by inspection of the definition of Δ_r), and if it holds for two given values of f then it holds for any of their $\mathbb{Q}$-linear combinations and also for their product.

[662] This is a weaker statement than saying that it has integer coefficients. (Actually, it does not in general have integer coefficients.)

[663] To see this, just work modulo the common denominator of the coefficients of the polynomial.

[664] Here, we are using Exercise 1.3.6(a) to see that $\operatorname{id} \otimes \Delta_\times$ and $\Delta_\times \otimes \operatorname{id}$ are $\mathbb{Q}$-algebra homomorphisms.

(h) Corollary 2.5.17(b) yields that $(p_\lambda)_{\lambda\in\operatorname{Par}}$ and $\left(z_\lambda^{-1}p_\lambda\right)_{\lambda\in\operatorname{Par}}$ are dual bases with respect to the Hall inner product on Λ. In other words,

$$\left(p_\lambda, z_\mu^{-1}p_\mu\right) = \delta_{\lambda,\mu} \qquad \text{for any partitions } \lambda \text{ and } \mu.$$

Hence,

$$\left(p_\lambda, \underbrace{p_\mu}_{=z_\mu z_\mu^{-1}p_\mu}\right) = \left(p_\lambda, z_\mu z_\mu^{-1}p_\mu\right) = z_\mu \underbrace{\left(p_\lambda, z_\mu^{-1}p_\mu\right)}_{=\delta_{\lambda,\mu}} = z_\mu\delta_{\lambda,\mu} \tag{13.83.26}$$

for any partitions λ and μ.

The Hall inner product $(\cdot,\cdot) : \Lambda_{\mathbb{Q}} \times \Lambda_{\mathbb{Q}} \to \mathbb{Q}$ is a bilinear form on $\Lambda_{\mathbb{Q}}$. Hence, according to Definition 3.1.1(b) (below), this inner product induces a $\mathbb{Q}$-bilinear form $(\cdot,\cdot)_{\Lambda_{\mathbb{Q}}\otimes_{\mathbb{Q}}\Lambda_{\mathbb{Q}}} : (\Lambda_{\mathbb{Q}}\otimes_{\mathbb{Q}}\Lambda_{\mathbb{Q}})\times(\Lambda_{\mathbb{Q}}\otimes_{\mathbb{Q}}\Lambda_{\mathbb{Q}}) \to \mathbb{Q}$. Similarly, the Hall inner product $(\cdot,\cdot) : \Lambda\times\Lambda \to \mathbb{Z}$ induces a $\mathbb{Z}$-bilinear form $(\cdot,\cdot)_{\Lambda\otimes_{\mathbb{Z}}\Lambda} : (\Lambda\otimes_{\mathbb{Z}}\Lambda)\times(\Lambda\otimes_{\mathbb{Z}}\Lambda)\to\mathbb{Z}$, and this latter $\mathbb{Z}$-bilinear form is clearly the restriction of the $\mathbb{Q}$-bilinear form $(\cdot,\cdot)_{\Lambda_{\mathbb{Q}}\otimes_{\mathbb{Q}}\Lambda_{\mathbb{Q}}}$ to $\Lambda\otimes_{\mathbb{Z}}\Lambda$.

We shall now show that

$$(a * b, c) = (a\otimes b, \Delta_\times(c))_{\Lambda_{\mathbb{Q}}\otimes_{\mathbb{Q}}\Lambda_{\mathbb{Q}}} \qquad \text{for all } a\in\Lambda_{\mathbb{Q}},\ b\in\Lambda_{\mathbb{Q}} \text{ and } c\in\Lambda_{\mathbb{Q}}. \tag{13.83.27}$$

Proof of (13.83.27): Let $a\in\Lambda_{\mathbb{Q}}$, $b\in\Lambda_{\mathbb{Q}}$ and $c\in\Lambda_{\mathbb{Q}}$. The equality (13.83.27) is clearly $\mathbb{Q}$-linear in each of a, b and c. Hence, in proving this equality, we can WLOG assume that a, b and c are elements of the basis $(p_\lambda)_{\lambda\in\operatorname{Par}}$ of the $\mathbb{Q}$-module $\Lambda_{\mathbb{Q}}$. Assume this. Thus, $a = p_\lambda$, $b = p_\mu$ and $c = p_\nu$ for some partitions λ, μ and ν; consider these partitions.

We have $\delta_{\lambda,\mu}z_\lambda = \delta_{\lambda,\mu}z_\mu$ (in fact, the two sides of this equality are equal when $\lambda=\mu$, and both vanish otherwise). Hence, $\underbrace{a}_{=p_\lambda} * \underbrace{b}_{=p_\mu} = p_\lambda * p_\mu = \underbrace{\delta_{\lambda,\mu}z_\lambda}_{=\delta_{\lambda,\mu}z_\mu} p_\lambda = \delta_{\lambda,\mu}z_\mu p_\lambda$, so that

$$\left(\underbrace{a*b}_{=\delta_{\lambda,\mu}z_\mu p_\lambda}, \underbrace{c}_{=p_\nu}\right) = \left(\delta_{\lambda,\mu}z_\mu p_\lambda, p_\nu\right) = \delta_{\lambda,\mu}z_\lambda \underbrace{(p_\lambda, p_\nu)}_{\substack{=z_\nu\delta_{\lambda,\nu} \\ \text{(by (13.83.26), applied} \\ \text{to } \nu \text{ instead of } \mu)}} = \delta_{\lambda,\mu}z_\lambda z_\nu\delta_{\lambda,\nu}.$$

On the other hand, $c = p_\nu$, so that

$$\Delta_\times(c) = \Delta_\times(p_\nu) = p_\nu\otimes p_\nu \qquad \text{(by (13.83.18), applied to } \nu \text{ instead of } \lambda\text{)},$$

and thus

$$\begin{aligned}\left(\underbrace{a}_{=p_\lambda}\otimes\underbrace{b}_{=p_\mu}, \underbrace{\Delta_\times(c)}_{=p_\nu\otimes p_\nu}\right)_{\Lambda_{\mathbb{Q}}\otimes_{\mathbb{Q}}\Lambda_{\mathbb{Q}}} &= (p_\lambda\otimes p_\mu, p_\nu\otimes p_\nu)_{\Lambda_{\mathbb{Q}}\otimes_{\mathbb{Q}}\Lambda_{\mathbb{Q}}} = \underbrace{(p_\lambda, p_\nu)}_{\substack{=z_\nu\delta_{\lambda,\nu} \\ \text{(by (13.83.26), applied} \\ \text{to } \nu \text{ instead of } \mu)}} \underbrace{(p_\mu, p_\nu)}_{\substack{=z_\nu\delta_{\mu,\nu} \\ \text{(by (13.83.26), applied} \\ \text{to } \mu \text{ and } \nu \text{ instead of } \lambda \text{ and } \mu)}} \\ &\qquad \left(\text{by the definition of the bilinear form } (\cdot,\cdot)_{\Lambda_{\mathbb{Q}}\otimes_{\mathbb{Q}}\Lambda_{\mathbb{Q}}}\right) \\ &= z_\nu\delta_{\lambda,\nu}z_\nu\delta_{\mu,\nu}.\end{aligned}$$

The equality in question, (13.83.27), thus rewrites as $\delta_{\lambda,\mu}z_\lambda z_\nu\delta_{\lambda,\nu} = z_\nu\delta_{\lambda,\nu}z_\nu\delta_{\mu,\nu}$ (because $(a*b,c) = \delta_{\lambda,\mu}z_\lambda z_\nu\delta_{\lambda,\nu}$ and $(a\otimes b, \Delta_\times(c))_{\Lambda_{\mathbb{Q}}\otimes_{\mathbb{Q}}\Lambda_{\mathbb{Q}}} = z_\nu\delta_{\lambda,\nu}z_\nu\delta_{\mu,\nu}$). But the latter equality is obvious (because both of its sides are z_ν^2 if $\lambda=\mu=\nu$, and vanish otherwise). Hence, (13.83.27) must hold as well.

Now that (13.83.27) is proven, let $f\in\Lambda$ and $g\in\Lambda$ be arbitrary. We can apply (13.83.1) to $\mathbf{k}=\mathbb{Q}$, $A=\Lambda_{\mathbb{Q}}$, $L=\operatorname{Par}$, $(u_\lambda)_{\lambda\in L} = (s_\lambda)_{\lambda\in\operatorname{Par}}$, $(v_\lambda)_{\lambda\in L} = (s_\lambda)_{\lambda\in\operatorname{Par}}$ and $a = f*g$ (because the basis $(s_\lambda)_{\lambda\in\operatorname{Par}}$ of $\Lambda_{\mathbb{Q}}$ is orthonormal with respect to the Hall inner product $(\cdot,\cdot)$, and thus dual to itself with respect to this product). As a result, we obtain

$$f*g=\sum_{\lambda\in\operatorname{Par}}\underbrace{(s_\lambda,f*g)}_{\substack{=(f*g,s_\lambda)\\ \text{(since the Hall inner product}\\ \text{is symmetric)}}}s_\lambda=\sum_{\lambda\in\operatorname{Par}}\underbrace{(f*g,s_\lambda)}_{\substack{=(f\otimes g,\Delta_\times(s_\lambda))_{\Lambda_{\mathbb{Q}}\otimes_{\mathbb{Q}}\Lambda_{\mathbb{Q}}}\\ \text{(by (13.83.27))}}}s_\lambda=\sum_{\lambda\in\operatorname{Par}}(f\otimes g,\Delta_\times(s_\lambda))_{\Lambda_{\mathbb{Q}}\otimes_{\mathbb{Q}}\Lambda_{\mathbb{Q}}}s_\lambda.$$

But every $\lambda\in\operatorname{Par}$ satisfies $\Delta_\times\left(\underbrace{s_\lambda}_{\in\Lambda}\right)\in\Delta_\times(\Lambda)\subset\Lambda\otimes_{\mathbb{Z}}\Lambda$ (by Exercise 2.9.4(b)) and thus

$$(f\otimes g,\Delta_\times(s_\lambda))_{\Lambda_{\mathbb{Q}}\otimes_{\mathbb{Q}}\Lambda_{\mathbb{Q}}}=(f\otimes g,\Delta_\times(s_\lambda))_{\Lambda\otimes_{\mathbb{Z}}\Lambda}\in\mathbb{Z}.$$

Hence,

$$f*g=\sum_{\lambda\in\operatorname{Par}}\underbrace{(f\otimes g,\Delta_\times(s_\lambda))_{\Lambda_{\mathbb{Q}}\otimes_{\mathbb{Q}}\Lambda_{\mathbb{Q}}}}_{\in\mathbb{Z}}s_\lambda\in\sum_{\lambda\in\operatorname{Par}}\mathbb{Z}s_\lambda\subset\Lambda.$$

This solves Exercise 2.9.4(h).

(i) Define a map $U:\Lambda_{\mathbb{Q}}\to\mathbb{Q}$ by

$$U(f)=f(1)\qquad\text{for all }f\in\Lambda_{\mathbb{Q}}.$$

Notice that $U(f)=f(1)$ is the result of substituting $1,0,0,0,\ldots$ for $x_1,x_2,x_3,\ldots$ in f. Hence, U is a $\mathbb{Q}$-algebra homomorphism.

We shall now show that $\epsilon_1=U$.

For every integer $n\geq1$, we have

$$\begin{aligned}U(p_n)&=p_n(1)\qquad\text{(by the definition of }U)\\&=(\text{the result of substituting }1,0,0,0,\ldots\text{ for }x_1,x_2,x_3,\ldots\text{ in }p_n)\\&\qquad\text{(by the definition of }p_n(1))\\&=(\text{the result of substituting }1,0,0,0,\ldots\text{ for }x_1,x_2,x_3,\ldots\text{ in }x_1^n+x_2^n+x_3^n+x_4^n+\cdots)\\&\qquad(\text{since }p_n=x_1^n+x_2^n+x_3^n+x_4^n+\cdots)\\&=1^n+0^n+0^n+0^n+\cdots=1+0+0+0+\cdots=1\\&=\epsilon_1(p_n)\qquad(\text{since }\epsilon_1(p_n)=1\text{ (by the definition of }\epsilon_1)).\end{aligned}$$

In other words, the two $\mathbb{Q}$-algebra homomorphisms U and ϵ_1 are equal to each other on each element of the family $(p_n)_{n\geq1}$. But since this family $(p_n)_{n\geq1}$ is a generating set of the $\mathbb{Q}$-algebra $\Lambda_{\mathbb{Q}}$, this yields that the two $\mathbb{Q}$-algebra homomorphisms U and ϵ_1 must be identical (because if two $\mathbb{Q}$-algebra homomorphisms are equal to each other on each element of a generating set of their domain, then they must be identical). That is, $\epsilon_1=U$. Hence, every $f\in\Lambda_{\mathbb{Q}}$ satisfies $\epsilon_1(f)=U(f)=f(1)$ (by the definition of U). This solves Exercise 2.9.4(i).

13.84. **Solution to Exercise 2.9.6.** *Solution to Exercise 2.9.6.* We are going to be brief; more detailed proofs for everything except of the (very easy) equivalence $\mathcal{D}\iff\mathcal{J}$ can be found at http://www.cip.ifi.lmu.de/~grinberg/algebra/witt5f.pdf (along with generalizations and additional equivalences in the case of $A=\mathbb{Z}$).

As suggested by the hint, we first prove some elementary facts of number theory:

- Every positive integer n satisfies

(13.84.1)
$$\sum_{d\mid n}\phi(d)=n.$$

Proof: Let n be a positive integer. For every positive divisor d of n, there is a bijection

$$\{i\in\{1,2,\ldots,n\}\ \mid\ \gcd(i,n)=d\}\to\left\{j\in\left\{1,2,\ldots,\frac{n}{d}\right\}\ \mid\ j\text{ is coprime to }\frac{n}{d}\right\},$$
$$i\mapsto\frac{i}{d}.$$

Hence, for every positive divisor d of n, we have

$$\begin{aligned} \left|\left\{i \in \{1,2,\ldots,n\} \mid \gcd(i,n) = d\right\}\right| &= \left|\left\{j \in \left\{1,2,\ldots,\frac{n}{d}\right\} \mid j \text{ is coprime to } \frac{n}{d}\right\}\right| \\ &= \phi\left(\frac{n}{d}\right) \qquad \left(\text{because this is how } \phi\left(\frac{n}{d}\right) \text{ was defined}\right). \end{aligned} \tag{13.84.2}$$

Now,

$$n = \left|\{1,2,\ldots,n\}\right| = \sum_{d\mid n} \underbrace{\left|\left\{i \in \{1,2,\ldots,n\} \mid \gcd(i,n) = d\right\}\right|}_{=\phi\left(\frac{n}{d}\right)} = \sum_{d\mid n} \phi\left(\frac{n}{d}\right) = \sum_{d\mid n} \phi(d)$$

(here, we have substituted d for $\frac{n}{d}$ in the sum). This proves (13.84.1).

- Every positive integer n satisfies

$$\sum_{d\mid n} \mu(d) = \delta_{n,1}. \tag{13.84.3}$$

Proof: Let n be a positive integer. Let $n = p_1^{a_1} p_2^{a_2} \cdots p_k^{a_k}$ be the prime factorization of n, with all of $a_1, a_2, \ldots, a_k$ being positive integers (and with $p_1, p_2, \ldots, p_k$ being distinct primes). Then, the **squarefree** positive divisors of n all have the form $\prod_{i\in I} p_i$ for some subset I of $\{1,2,\ldots,k\}$. More precisely, there exists a bijection

$$\begin{aligned} \{I \subset \{1,2,\ldots,k\}\} &\to (\text{the set of all squarefree positive divisors of } n), \\ I &\mapsto \prod_{i\in I} p_i. \end{aligned} \tag{13.84.4}$$

Now, each positive divisor d of n is either squarefree or not. Hence,

$$\begin{aligned} \sum_{d\mid n} \mu(d) &= \sum_{\substack{d\mid n;\\ d \text{ is squarefree}}} \mu(d) + \sum_{\substack{d\mid n;\\ d \text{ is not squarefree}}} \underbrace{\mu(d)}_{\substack{=0\\ \text{(by the definition}\\ \text{of } \mu)}} = \sum_{\substack{d\mid n;\\ d \text{ is squarefree}}} \mu(d) \\ &= \sum_{I\subset\{1,2,\ldots,k\}} \underbrace{\mu\left(\prod_{i\in I} p_i\right)}_{\substack{=(-1)^{|I|}\\ \text{(since } \prod_{i\in I} p_i \text{ is squarefree}\\ \text{and has } |I| \text{ prime factors)}}} \\ &\qquad \left(\begin{array}{c}\text{here, we have substituted } \prod_{i\in I} p_i \text{ for } d\\ \text{due to the bijection (13.84.4)}\end{array}\right) \\ &= \sum_{I\subset\{1,2,\ldots,k\}} (-1)^{|I|} = \begin{cases} 1, & \text{if } \{1,2,\ldots,k\} = \varnothing; \\ 0, & \text{otherwise} \end{cases} = \begin{cases} 1, & \text{if } k = 0; \\ 0, & \text{otherwise} \end{cases} \\ &= \begin{cases} 1, & \text{if } n = 1; \\ 0, & \text{otherwise} \end{cases} \qquad \left(\begin{array}{c}\text{since } k \text{ is the number of distinct prime factors of } n,\\ \text{and thus we have } k = 0 \text{ if and only if } n = 1\end{array}\right) \\ &= \delta_{n,1}, \end{aligned}$$

which proves (13.84.3).

- Every positive integer n satisfies

$$\sum_{d\mid n} \mu(d) \frac{n}{d} = \phi(n). \tag{13.84.5}$$

Proof: This is an elementary fact, but we have a hammer at our disposal, and this looks conspicuously like a nail. Let us define a $\mathbb{Z}$-coalgebra. Namely, let $\mathbf{T}$ be the free $\mathbb{Z}$-module with basis

$(t_n)_{n\geq 1}$. We define a $\mathbb{Z}$-coalgebra structure on $\mathbf{T}$ by setting

$$\Delta(t_n) = \sum_{\substack{d,e\in\{1,2,3,\ldots\};\\ de=n}} t_d \otimes t_e = \sum_{d\mid n} t_d \otimes t_{n/d} \qquad \text{and} \qquad \epsilon(t_n) = \delta_{n,1}$$

for every $n \in \{1,2,3,\ldots\}$. It is straightforward to check that this makes $\mathbf{T}$ into a cocommutative coalgebra. Hence, $(\operatorname{Hom}_{\mathbb{Z}}(\mathbf{T},\mathbb{Z}), \star)$ is a $\mathbb{Z}$-algebra with unity ϵ. [665] This algebra $(\operatorname{Hom}_{\mathbb{Z}}(\mathbf{T},\mathbb{Z}), \star)$ is commutative[666]. We define four elements of $\operatorname{Hom}_{\mathbb{Z}}(\mathbf{T},\mathbb{Z})$:

- a $\mathbb{Z}$-linear map $\widetilde{\phi} : \mathbf{T} \to \mathbb{Z}$ which sends t_n to $\phi(n)$ for every $n \geq 1$;
- a $\mathbb{Z}$-linear map $\widetilde{\mu} : \mathbf{T} \to \mathbb{Z}$ which sends t_n to $\mu(n)$ for every $n \geq 1$;
- a $\mathbb{Z}$-linear map $\widetilde{\operatorname{id}} : \mathbf{T} \to \mathbb{Z}$ which sends t_n to n for every $n \geq 1$;
- a $\mathbb{Z}$-linear map $\widetilde{\mathbf{1}} : \mathbf{T} \to \mathbb{Z}$ which sends t_n to 1 for every $n \geq 1$.

Then, the identity that we want to prove – i.e., the identity (13.84.5) – is equivalent to the claim that $\widetilde{\mu} \star \widetilde{\operatorname{id}} = \widetilde{\phi}$. [667] Similarly, the (already proven) identity (13.84.1) is equivalent to $\widetilde{\phi} \star \widetilde{\mathbf{1}} = \widetilde{\operatorname{id}}$, and the (already proven) identity (13.84.3) is equivalent to $\widetilde{\mu} \star \widetilde{\mathbf{1}} = \epsilon$. Thus,

$$\begin{aligned}\widetilde{\mu} \star \underbrace{\widetilde{\operatorname{id}}}_{=\widetilde{\phi}\star\widetilde{\mathbf{1}}} &= \widetilde{\mu} \star \widetilde{\phi} \star \widetilde{\mathbf{1}} = \underbrace{\widetilde{\mu} \star \widetilde{\mathbf{1}}}_{=\epsilon} \star \widetilde{\phi} \qquad \left(\text{since } (\operatorname{Hom}_{\mathbb{Z}}(\mathbf{T},\mathbb{Z}), \star) \text{ is commutative}\right) \\ &= \epsilon \star \widetilde{\phi} = \widetilde{\phi}.\end{aligned}$$

As we know, this is equivalent to (13.84.5), so that (13.84.5) is proven.

It was not really necessary to phrase this argument in terms of coalgebras; this was only done to illustrate a use of the latter. Our proof can just as well be rewritten as a manipulation of sums, or (as a compromise between concreteness and structure) it can be paraphrased by using the *Dirichlet convolution*, which is the operation taking two maps $f, g : \{1,2,3,\ldots\} \to \mathbb{Z}$ to a third map $h : \{1,2,3,\ldots\} \to \mathbb{Z}$ defined by $h(n) = \sum_{d\mid n} f(d)\, g\left(\dfrac{n}{d}\right)$. Of course, this Dirichlet convolution is the same as our convolution $\star$ on $\operatorname{Hom}_{\mathbb{Z}}(\mathbf{T},\mathbb{Z})$, with the only difference that $\mathbb{Z}$-linear maps $\mathbf{T} \to \mathbb{Z}$ are replaced by arbitrary maps $\{1,2,3,\ldots\} \to \mathbb{Z}$.

- Every positive integer n satisfies

$$\sum_{d\mid n} d\mu(d)\, \phi\left(\frac{n}{d}\right) = \mu(n). \tag{13.84.6}$$

Proof: We use the setup we prepared in the proof of (13.84.5). Additionally, we define $\widetilde{\mu'} : \mathbf{T} \to \mathbb{Z}$ as the $\mathbb{Z}$-linear map which sends t_n to $n\mu(n)$ for every $n \geq 1$. Then, $\widetilde{\mu'} \star \widetilde{\operatorname{id}} = \epsilon$, because every positive integer n satisfies

$$\left(\widetilde{\mu'} \star \widetilde{\operatorname{id}}\right)(t_n) = \sum_{d\mid n} d\mu(d)\, \frac{n}{d} = \sum_{d\mid n} n\mu(d) = n \underbrace{\sum_{d\mid n} \mu(d)}_{\substack{=\delta_{n,1} \\ \text{(by (13.84.3))}}} = n\delta_{n,1} = \delta_{n,1} = \epsilon(t_n).$$

But recall that $\widetilde{\phi} = \widetilde{\mu} \star \widetilde{\operatorname{id}} = \widetilde{\operatorname{id}} \star \widetilde{\mu}$ (since $(\operatorname{Hom}_{\mathbb{Z}}(\mathbf{T},\mathbb{Z}), \star)$ is commutative), and thus

$$\widetilde{\mu'} \star \underbrace{\widetilde{\phi}}_{=\widetilde{\operatorname{id}}\star\widetilde{\mu}} = \underbrace{\widetilde{\mu'} \star \widetilde{\operatorname{id}}}_{=\epsilon} \star \widetilde{\mu} = \epsilon \star \widetilde{\mu} = \widetilde{\mu}.$$

[665] Number theorists will recognize this $\mathbb{Z}$-algebra as an isomorphic version of the so-called *algebra of formal Dirichlet series over* $\mathbb{Z}$. The isomorphism from $(\operatorname{Hom}_{\mathbb{Z}}(\mathbf{T},\mathbb{Z}), \star)$ to the algebra of formal Dirichlet series over $\mathbb{Z}$ takes an element $f \in (\operatorname{Hom}_{\mathbb{Z}}(\mathbf{T},\mathbb{Z}), \star)$ to the formal Dirichlet series $\sum_{n=1}^{\infty} f(n)\, n^{-s}$.

[666] This follows from Exercise 1.5.5 (applied to $\mathbb{Z}$, $\mathbf{T}$ and $\mathbb{Z}$ instead of $\mathbf{k}$, C and A).

[667] This is because the left hand side of (13.84.5) equals $\sum_{d\mid n} \widetilde{\mu}(t_d)\, \widetilde{\operatorname{id}}(t_{n/d}) = m\left(\left(\widetilde{\mu} \otimes \widetilde{\operatorname{id}}\right)\underbrace{\left(\sum_{d\mid n} t_d \otimes t_{n/d}\right)}_{=\Delta(t_n)}\right) = m\left(\left(\widetilde{\mu} \otimes \widetilde{\operatorname{id}}\right)(\Delta(t_n))\right) = \left(\widetilde{\mu} \star \widetilde{\operatorname{id}}\right)(t_n)$, whereas the right hand side equals $\widetilde{\phi}(t_n)$.

This is easily seen to be equivalent to (13.84.6), and so (13.84.6) is proven.

- If k is a positive integer, and if $p \in \mathbb{N}$, $a \in A$ and $b \in A$ are such that $a \equiv b \operatorname{mod} p^k A$, then

$$a^{p^\ell} \equiv b^{p^\ell} \operatorname{mod} p^{k+\ell} A \qquad \text{for every } \ell \in \mathbb{N}. \tag{13.84.7}$$

Proof: Let k be a positive integer. Let $p \in \mathbb{N}$, $a \in A$ and $b \in A$ be such that $a \equiv b \operatorname{mod} p^k A$. We need to prove (13.84.7). It is clearly enough to show that $a^p \equiv b^p \operatorname{mod} p^{k+1} A$, because then (13.84.7) will follow by induction over ℓ. But we have $a \equiv b \operatorname{mod} pA$ (since $a \equiv b \operatorname{mod} p^k A$ and since k is positive) and therefore

$$\begin{aligned} a^{p-1} + a^{p-2} b + \cdots + b^{p-1} &\equiv b^{p-1} + b^{p-2} b + \cdots + b^{p-1} \\ &= \underbrace{b^{p-1} + b^{p-1} + \cdots + b^{p-1}}_{p \text{ terms}} = p b^{p-1} \equiv 0 \operatorname{mod} pA, \end{aligned}$$

so that $a^{p-1} + a^{p-2} b + \cdots + b^{p-1} \in pA$. Thus,

$$a^p - b^p = \underbrace{(a-b)}_{\substack{\in p^k A \\ (\text{since } a \equiv b \operatorname{mod} p^k A)}} \underbrace{\left(a^{p-1} + a^{p-2} b + \cdots + b^{p-1}\right)}_{\in pA} \in \left(p^k A\right)(pA) = p^{k+1} A,$$

so that $a^p \equiv b^p \operatorname{mod} p^{k+1} A$. This proves (13.84.7).

- Here is a slightly more useful corollary of (13.84.7): If a prime number p and two elements a and b of A are such that $a \equiv b \operatorname{mod} pA$, then

$$a^N \equiv b^N \operatorname{mod} p^{v_p(N)+1} A \qquad \text{for every } N \in \{1, 2, 3, \ldots\}. \tag{13.84.8}$$

Proof: Let p be a prime number, and let a and b be two elements of A such that $a \equiv b \operatorname{mod} pA$. Let $N \in \{1, 2, 3, \ldots\}$. Write N in the form $N = p^{v_p(N)} M$ for some positive integer M. Then, (13.84.7) (applied to $\ell = v_p(N)$ and $k = 1$) yields $a^{p^{v_p(N)}} \equiv b^{p^{v_p(N)}} \operatorname{mod} p^{v_p(N)+1} A$, and because of $N = p^{v_p(N)} M$ we have

$$\begin{aligned} a^N &= a^{p^{v_p(N)} M} = \left(a^{p^{v_p(N)}}\right)^M \equiv \left(b^{p^{v_p(N)}}\right)^M \qquad \left(\text{since } a^{p^{v_p(N)}} \equiv b^{p^{v_p(N)}} \operatorname{mod} p^{v_p(N)+1} A\right) \\ &= b^{p^{v_p(N)} M} = b^N \operatorname{mod} p^{v_p(N)+1} A \qquad \left(\text{since } p^{v_p(N)} M = N\right). \end{aligned}$$

This proves (13.84.8).

We shall also use a fact from commutative algebra – namely, one of the versions of the Chinese Remainder Theorem of ring theory:

Theorem 13.84.1. *Let A be a commutative ring. Let $\mathbf{S}$ be a finite set. For every $s \in \mathbf{S}$, let I_s be an ideal of A. Assume that the ideals I_s of A are* comaximal[668]*; this means that every two distinct elements s and t of $\mathbf{S}$ satisfy $I_s + I_t = A$. Then:*

(a) *We have*

$$\bigcap_{s \in \mathbf{S}} I_s = \prod_{s \in \mathbf{S}} I_s.$$

(b) *The canonical ring homomorphism*

$$A / \left(\bigcap_{s \in \mathbf{S}} I_s\right) \to \prod_{s \in \mathbf{S}} (A / I_s), \qquad a + \bigcap_{s \in \mathbf{S}} I_s \mapsto (a + I_s)_{s \in \mathbf{S}}$$

is well-defined and a ring isomorphism.

See http://stacks.math.columbia.edu/tag/00DT (Lemma 10.14.4 (1) in the Stacks Project, as of 4 April 2020) or many other sources for a proof of Theorem 13.84.1. We shall only use part (a) of this theorem.

As a consequence of Theorem 13.84.1(a), we have the following:

[668] Some authors use the word "coprime" instead of "comaximal" here.

- If n is a positive integer and if A is a commutative ring, then

$$\bigcap_{\substack{p \text{ is a prime} \\ \text{factor of } n}} p^{v_p(n)} A = \prod_{\substack{p \text{ is a prime} \\ \text{factor of } n}} \left(p^{v_p(n)} A\right). \tag{13.84.9}$$

Indeed, this follows from Theorem 13.84.1(a), since the ideals $p^{v_p(n)}A$ of A (for varying p) are comaximal (because for any two distinct primes p and q, we can find integers x and y satisfying $p^{v_p(n)}x + q^{v_q(n)}y = 1$, and therefore we have $p^{v_p(n)}A + q^{v_q(n)}A = A$).

Now, we can step to proving the actual equivalences.

Proof of the implication $\mathcal{D} \Longrightarrow \mathcal{J}$: Assume that Assertion $\mathcal{D}$ holds. That is, there exists a family $(\alpha_n)_{n\geq 1} \in A^{\{1,2,3,\ldots\}}$ of elements of A such that every positive integer n satisfies $b_n = \sum_{d|n} d\alpha_d^{n/d}$. Consider this family $(\alpha_n)_{n\geq 1}$. Consider also the family $(w_n)_{n\geq 1} \in \Lambda_{\mathbb{Z}}^{\{1,2,3,\ldots\}}$ defined in Exercise 2.9.3(a). This family $(w_n)_{n\geq 1}$ is an algebraically independent generating set of $\Lambda_{\mathbb{Z}}$ (indeed, this is a restatement of Exercise 2.9.3(d)). Hence, there exists a unique $\mathbb{Z}$-algebra homomorphism $f : \Lambda_{\mathbb{Z}} \to A$ which satisfies

$$f(w_n) = \alpha_n \qquad \text{for every } n \in \{1, 2, 3, \ldots\}.$$

Consider this f. We have $p_n = \sum_{d|n} dw_d^{n/d}$ for every positive integer n (by Exercise 2.9.3(e)). Thus, for every positive integer n, we have

$$f(p_n) = f\left(\sum_{d|n} dw_d^{n/d}\right) = \sum_{d|n} d\left(\underbrace{f(w_d)}_{=\alpha_d}\right)^{n/d} \qquad \text{(since } f \text{ is a } \mathbb{Z}\text{-algebra homomorphism)}$$
$$= \sum_{d|n} d\alpha_d^{n/d} = b_n.$$

Hence, there exists a ring homomorphism $\Lambda_{\mathbb{Z}} \to A$ which, for every positive integer n, sends p_n to b_n (namely, f). That is, Assertion $\mathcal{J}$ holds, and the implication $\mathcal{D} \Longrightarrow \mathcal{J}$ is proven.

Proof of the implication $\mathcal{J} \Longrightarrow \mathcal{D}$: Assume that Assertion $\mathcal{J}$ holds. That is, there exists a ring homomorphism $\Lambda_{\mathbb{Z}} \to A$ which, for every positive integer n, sends p_n to b_n. Let f be such a homomorphism. Consider the family $(w_n)_{n\geq 1} \in \Lambda_{\mathbb{Z}}^{\{1,2,3,\ldots\}}$ defined in Exercise 2.9.3(a). For every positive integer n, we have $p_n = \sum_{d|n} dw_d^{n/d}$ (by Exercise 2.9.3(e)) and thus

$$f(p_n) = f\left(\sum_{d|n} dw_d^{n/d}\right) = \sum_{d|n} d\left(f(w_d)\right)^{n/d} \qquad \text{(since } f \text{ is a ring homomorphism)}.$$

Since $f(p_n) = b_n$ (by the definition of f), this rewrites as $b_n = \sum_{d|n} d\left(f(w_d)\right)^{n/d}$. Thus, there exists a family $(\alpha_n)_{n\geq 1} \in A^{\{1,2,3,\ldots\}}$ of elements of A such that every positive integer n satisfies $b_n = \sum_{d|n} d\alpha_d^{n/d}$ (namely, such a family can be defined by $\alpha_n = f(w_n)$). Assertion $\mathcal{D}$ thus holds. We have now proven the implication $\mathcal{J} \Longrightarrow \mathcal{D}$.

Proof of the implication $\mathcal{D} \Longrightarrow \mathcal{C}$: Assume that Assertion $\mathcal{D}$ holds. That is, there exists a family $(\alpha_n)_{n\geq 1} \in A^{\{1,2,3,\ldots\}}$ of elements of A such that every positive integer n satisfies $b_n = \sum_{d|n} d\alpha_d^{n/d}$. Consider this family $(\alpha_n)_{n\geq 1}$.

We need to prove that Assertion $\mathcal{C}$ holds, i.e., that we have

$$\varphi_p\left(b_{n/p}\right) \equiv b_n \mod p^{v_p(n)}A \tag{13.84.10}$$

for every positive integer n and every prime factor p of n. So let us fix a positive integer n and a prime factor p of n. We need to prove (13.84.10).

The definition of the family $(\alpha_n)_{n\geq 1}$ shows that $b_{n/p} = \sum_{d|n/p} d\alpha_d^{(n/p)/d}$, so that

$$\varphi_p\left(b_{n/p}\right) = \varphi_p\left(\sum_{d|n/p} d\alpha_d^{(n/p)/d}\right) = \sum_{d|n/p} d\left(\varphi_p\left(\alpha_d\right)\right)^{(n/p)/d} \qquad \text{(since } \varphi_p \text{ is a ring endomorphism)}.$$

On the other hand,

$$\sum_{d\mid n} d\alpha_d^{n/d} = \sum_{d\mid n/p} d\alpha_d^{n/d} + \sum_{\substack{d\mid n;\\ d\nmid n/p}} \underbrace{d}_{\substack{\equiv 0 \bmod p^{v_p(n)}A \\ \text{(since } d\mid n \text{ and } d\nmid n/p \\ \text{yield } p^{v_p(n)}\mid d)}} \alpha_d^{n/d} \equiv \sum_{d\mid n/p} d\alpha_d^{n/d} \bmod p^{v_p(n)}A. \tag{13.84.11}$$

Thus, if we succeed to prove that

$$d\left(\varphi_p\left(\alpha_d\right)\right)^{(n/p)/d} \equiv d\alpha_d^{n/d} \bmod p^{v_p(n)}A \qquad \text{for every } d \mid n/p, \tag{13.84.12}$$

then we will obtain

$$\begin{aligned} \varphi_p\left(b_{n/p}\right) &= \sum_{d\mid n/p} \underbrace{d\left(\varphi_p\left(\alpha_d\right)\right)^{(n/p)/d}}_{\equiv d\alpha_d^{n/d} \bmod p^{v_p(n)}A} \equiv \sum_{d\mid n/p} d\alpha_d^{n/d} \equiv \sum_{d\mid n} d\alpha_d^{n/d} \qquad (\text{by } (13.84.11)) \\ &= b_n \bmod p^{v_p(n)}A \qquad \left(\text{since } b_n = \sum_{d\mid n} d\alpha_d^{n/d}\right), \end{aligned}$$

and thus our goal (proving (13.84.10)) will be achieved. Hence, it remains to prove (13.84.12).

So let d be any positive divisor of n/p. Then, $\varphi_p\left(\alpha_d\right) \equiv \alpha_d^p \bmod pA$ (because of the axiom $\varphi_p(a) \equiv a^p \bmod pA$ for every $a \in A$). Thus, (13.84.8) (applied to $a = \varphi_p\left(\alpha_d\right)$, $b = \alpha_p^d$ and $N = (n/p)/d$) yields $\left(\varphi_p\left(\alpha_d\right)\right)^{(n/p)/d} \equiv \left(\alpha_d^p\right)^{(n/p)/d} \bmod p^{v_p((n/p)/d)+1}A$. Since $\left(\alpha_d^p\right)^{(n/p)/d} = \alpha_d^{n/d}$ and $v_p\left((n/p)/d\right)+1 = v_p(n/d)$, this rewrites as $\left(\varphi_p\left(\alpha_d\right)\right)^{(n/p)/d} \equiv \alpha_d^{n/d} \bmod p^{v_p(n/d)}A$. Multiplying this by d results in $d\left(\varphi_p\left(\alpha_d\right)\right)^{(n/p)/d} \equiv d\alpha_d^{n/d} \bmod dp^{v_p(n/d)}A$. Since $dp^{v_p(n/d)}$ is divisible by $p^{v_p(n)}$, this yields (13.84.12). This completes the proof of (13.84.12), and thus also that of the implication $\mathcal{D} \Longrightarrow \mathcal{C}$.

Proof of the implication $\mathcal{C} \Longrightarrow \mathcal{D}$*:* Assume that Assertion $\mathcal{C}$ holds. Thus, for every positive integer n and every prime factor p of n, we have

$$\varphi_p\left(b_{n/p}\right) \equiv b_n \bmod p^{v_p(n)}A. \tag{13.84.13}$$

We now need to prove that Assertion $\mathcal{D}$ holds as well. In other words, we need to show that there exists a family $\left(\alpha_n\right)_{n\geq 1} \in A^{\{1,2,3,\ldots\}}$ of elements of A such that every positive integer n satisfies $b_n = \sum_{d\mid n} d\alpha_d^{n/d}$. In other words (renaming n as m), we need to show that there exists a family $\left(\alpha_m\right)_{m\geq 1} \in A^{\{1,2,3,\ldots\}}$ of elements of A such that every positive integer m satisfies $b_m = \sum_{d\mid m} d\alpha_d^{m/d}$.

We construct this family $\left(\alpha_m\right)_{m\geq 1}$ recursively. So we fix some $n \geq 1$, and assume that α_m is already constructed for every positive integer $m < n$ in such a way that

$$\text{the equality } b_m = \sum_{d\mid m} d\alpha_d^{m/d} \text{ is satisfied for every positive integer } m < n. \tag{13.84.14}$$

We now need to construct an $\alpha_n \in A$ such that $b_m = \sum_{d\mid m} d\alpha_d^{m/d}$ is satisfied for $m = n$. In other words, we need to construct an $\alpha_n \in A$ satisfying $b_n = \sum_{d\mid n} d\alpha_d^{n/d}$.

Let us first choose α_n **arbitrarily** (with the intention to tweak it later). Let p be any prime factor of n. Then, applying (13.84.14) to $m = n/p$, we obtain $b_{n/p} = \sum_{d\mid n/p} d\alpha_d^{(n/p)/d}$. This allows us to prove that $\varphi_p\left(b_{n/p}\right) \equiv \sum_{d\mid n} d\alpha_d^{n/d} \bmod p^{v_p(n)}A$ holds (just as in the proof of the implication $\mathcal{D} \Longrightarrow \mathcal{C}$). Compared with (13.84.13), this yields $b_n \equiv \sum_{d\mid n} d\alpha_d^{n/d} \bmod p^{v_p(n)}A$. That is, $b_n - \sum_{d\mid n} d\alpha_d^{n/d} \in p^{v_p(n)}A$.

Now, let us forget that we fixed p. We thus have shown that (with our arbitrary choice of α_n) we have $b_n - \sum_{d|n} d\alpha_d^{n/d} \in p^{v_p(n)} A$ for every prime factor p of n. As a consequence,

$$b_n - \sum_{d|n} d\alpha_d^{n/d} \in \bigcap_{\substack{p \text{ is a prime} \\ \text{factor of } n}} p^{v_p(n)} A = \prod_{\substack{p \text{ is a prime} \\ \text{factor of } n}} \left(p^{v_p(n)} A\right) \qquad \text{(by (13.84.9))}$$
$$= \underbrace{\left(\prod_{\substack{p \text{ is a prime} \\ \text{factor of } n}} p^{v_p(n)} \right)}_{=n} A = nA.$$

In other words, there exists a $\gamma \in A$ such that $b_n - \sum_{d|n} d\alpha_d^{n/d} = n\gamma$. Consider this γ. Now, if we replace α_n by $\alpha_n + \gamma$, then the sum $\sum_{d|n} d\alpha_d^{n/d}$ increases by $n\gamma = b_n - \sum_{d|n} d\alpha_d^{n/d}$, and therefore becomes precisely b_n. Hence, by replacing α_n by $\alpha_n + \gamma$, we achieve that $b_n = \sum_{d|n} d\alpha_d^{n/d}$ holds. Thus, we have found the α_n we were searching for, and the recursive construction of the family $(\alpha_m)_{m \geq 1}$ has proceeded by one more step. The proof of the implication $\mathcal{C} \Longrightarrow \mathcal{D}$ is thus complete.

Proof of the implication $\mathcal{E} \Longrightarrow \mathcal{C}$: The proof of the implication $\mathcal{E} \Longrightarrow \mathcal{C}$ proceeds exactly as our above proof of the implication $\mathcal{D} \Longrightarrow \mathcal{C}$, with the following changes:

- Every appearance of α_i for some $i \geq 1$ must be replaced by the corresponding β_i.
- Every time an element of A was taken to the k-th power (for some $k \in \{1, 2, 3, \ldots\}$) in our proof of the implication $\mathcal{D} \Longrightarrow \mathcal{C}$, it needs now to be subjected to the ring endomorphism φ_k instead. So, for example, $\alpha_d^{n/d}$ is replaced by $\varphi_{n/d}(\beta_d)$ everywhere (remember that α_d becomes β_d). Note that this concerns only elements of A. We don't replace the power $p^{v_p(n)}$ by anything.
- The equality
$$\varphi_p \left(\sum_{d|n/p} d\alpha_d^{(n/p)/d} \right) = \sum_{d|n/p} d \left(\varphi_p(\alpha_d)\right)^{(n/p)/d}$$
is replaced by
$$\varphi_p \left(\sum_{d|n/p} d\varphi_{(n/p)/d}(\beta_d) \right) = \sum_{d|n/p} d\varphi_{(n/p)/d}(\varphi_p(\beta_d)),$$
whose proof uses the $\mathbb{Z}$-linearity of φ_p and the fact that $\varphi_p \circ \varphi_{(n/p)/d} = \varphi_{n/d} = \varphi_{(n/p)/d} \circ \varphi_p$.
- The proof of (13.84.12) needs to be replaced by a proof of the congruence

$$d\varphi_{(n/p)/d}(\varphi_p(\beta_d)) \equiv d\varphi_{n/d}(\beta_d) \mod p^{v_p(n)} A \qquad \text{for every } d \mid n/p. \tag{13.84.15}$$

Fortunately, the latter congruence is obvious, since $\varphi_{(n/p)/d} \circ \varphi_p = \varphi_{n/d}$.

Proof of the implication $\mathcal{C} \Longrightarrow \mathcal{E}$: The proof of the implication $\mathcal{C} \Longrightarrow \mathcal{E}$ can be obtained from the proof of the implication $\mathcal{C} \Longrightarrow \mathcal{D}$ using the same changes that were made to transform the proof of the implication $\mathcal{D} \Longrightarrow \mathcal{C}$ into a proof of the implication $\mathcal{E} \Longrightarrow \mathcal{C}$. The only new "idea" is to use the fact that $\varphi_1 = \text{id}$ (in showing that if we replace β_n by $\beta_n + \gamma$, then the sum $\sum_{d|n} d\varphi_{n/d}(\beta_d)$ increases by $n\gamma$).

Proof of the implication $\mathcal{E} \Longrightarrow \mathcal{F}$: Assume that Assertion $\mathcal{E}$ holds. That is, there exists a family $(\beta_n)_{n \geq 1} \in A^{\{1,2,3,\ldots\}}$ of elements of A such that every positive integer n satisfies

$$b_n = \sum_{d|n} d\varphi_{n/d}(\beta_d). \tag{13.84.16}$$

Consider this family $(\beta_n)_{n \geq 1}$. We need to prove that Assertion $\mathcal{F}$ holds, i.e., that every positive integer n satisfies

$$\sum_{d|n} \mu(d) \varphi_d(b_{n/d}) \in nA.$$

We fix a positive integer n. Then, every positive divisor d of n satisfies $b_{n/d} = \sum_{e \mid n/d} e\varphi_{(n/d)/e}(\beta_e)$ (by (13.84.16), applied to n/d instead of n, and with the summation index d renamed as e). Hence,

$$\begin{aligned}
\sum_{d \mid n} \mu(d)\, \varphi_d \left(\underbrace{b_{n/d}}_{= \sum_{e \mid n/d} e\varphi_{(n/d)/e}(\beta_e)} \right) &= \sum_{d \mid n} \mu(d)\, \varphi_d \left(\sum_{e \mid n/d} e\varphi_{(n/d)/e}(\beta_e) \right) \\
&= \sum_{d \mid n} \mu(d) \sum_{e \mid n/d} e\varphi_d\left(\varphi_{(n/d)/e}(\beta_e)\right) \qquad \text{(since } \varphi_d \text{ is linear)} \\
&= \underbrace{\sum_{d \mid n} \sum_{e \mid n/d}}_{= \sum_{e \mid n} \sum_{d \mid n/e}} \mu(d)\, e \underbrace{\varphi_d\left(\varphi_{(n/d)/e}(\beta_e)\right)}_{\substack{= \left(\varphi_d \circ \varphi_{(n/d)/e}\right)(\beta_e) \\ = \varphi_{n/e}(\beta_e)}} \\
&= \sum_{e \mid n} \underbrace{\sum_{d \mid n/e} \mu(d)}_{\substack{= \delta_{n/e,1} \\ \text{(by (13.84.3), applied} \\ \text{to } n/e \text{ instead of } n\text{)}}} e\varphi_{n/e}(\beta_e) = \sum_{e \mid n} \underbrace{\delta_{n/e,1}}_{= \delta_{e,n}} e\varphi_{n/e}(\beta_e) \\
&= \sum_{e \mid n} \delta_{e,n} e\varphi_{n/e}(\beta_e) = n \underbrace{\varphi_{n/n}}_{= \varphi_1 = \mathrm{id}} (\beta_n) = n\beta_n \in nA.
\end{aligned}$$

Thus, Assertion $\mathcal{F}$ holds, so that we have proven the implication $\mathcal{E} \Longrightarrow \mathcal{F}$.

Proof of the implication $\mathcal{F} \Longrightarrow \mathcal{E}$*:* Assume that Assertion $\mathcal{F}$ holds. That is, every positive integer n satisfies

$$\sum_{d \mid n} \mu(d)\, \varphi_d\left(b_{n/d}\right) \in nA. \tag{13.84.17}$$

Now we need to prove that Assertion $\mathcal{E}$ holds, i.e., that there exists a family $(\beta_n)_{n \geq 1} \in A^{\{1,2,3,\ldots\}}$ of elements of A such that every positive integer n satisfies

$$b_n = \sum_{d \mid n} d\varphi_{n/d}(\beta_d). \tag{13.84.18}$$

We shall construct such a family $(\beta_n)_{n \geq 1}$ recursively. That is, we fix some $N \in \{1, 2, 3, \ldots\}$, and we assume that we already have constructed a $\beta_n \in A$ for every positive integer $n < N$ in such a way that (13.84.18) is satisfied for every positive integer $n < N$. We now need to find a $\beta_N \in A$ such that (13.84.18) is satisfied for $n = N$ as well.

From (13.84.17) (applied to $n = N$), we have $\sum_{d\mid N} \mu(d)\, \varphi_d\left(b_{N/d}\right) \in nA$. Thus, there exists a $t \in A$ such that $\sum_{d\mid N} \mu(d)\, \varphi_d\left(b_{N/d}\right) = Nt$. Set Consider this t. Set $\beta_N = t$. We have

$$\begin{aligned}
Nt &= \sum_{d\mid N} \mu(d)\, \varphi_d\left(b_{N/d}\right) = \sum_{e\mid N} \mu(e)\, \varphi_e\left(b_{N/e}\right) = \sum_{\substack{e\mid N;\\ e>1}} \mu(e)\, \varphi_e\left(\underbrace{b_{N/e}}_{\substack{=\sum_{d\mid N/e} d\varphi_{(N/e)/d}(\beta_d) \\ \text{(by (13.84.18), applied} \\ \text{to } n=N/e, \text{ because } N/e<n)}}\right) + \underbrace{\mu(1)}_{=1}\underbrace{\varphi_1}_{=\operatorname{id}}\left(\underbrace{b_{N/1}}_{=b_N}\right) \\
&= \sum_{\substack{e\mid N;\\ e>1}} \mu(e)\, \varphi_e\left(\sum_{d\mid N/e} d\varphi_{(N/e)/d}\left(\beta_d\right)\right) + b_N \\
&= \sum_{\substack{e\mid N;\\ e>1}} \mu(e) \sum_{d\mid N/e} d\, \underbrace{\varphi_e\left(\varphi_{(N/e)/d}\left(\beta_d\right)\right)}_{\substack{=\left(\varphi_e\circ\varphi_{(N/e)/d}\right)(\beta_d) \\ =\varphi_{N/d}(\beta_d)}} + b_N \qquad \text{(since } \varphi_e \text{ is linear)} \\
&= \sum_{\substack{e\mid N;\\ e>1}} \mu(e) \sum_{d\mid N/e} d\varphi_{N/d}\left(\beta_d\right) + b_N = \underbrace{\sum_{\substack{e\mid N;\\ e>1}} \sum_{d\mid N/e}}_{= \sum_{\substack{d\mid N;\\ d<N}} \sum_{\substack{e\mid N/d;\\ e>1}}} \mu(e)\, d\varphi_{N/d}\left(\beta_d\right) + b_N \\
&= \sum_{\substack{d\mid N;\\ d<N}} \underbrace{\sum_{\substack{e\mid N/d;\\ e>1}} \mu(e)}_{\substack{=\sum_{e\mid N/d}\mu(e)-\mu(1) \\ \text{(since } 1\mid N/d)}} d\varphi_{N/d}\left(\beta_d\right) + b_N = \sum_{\substack{d\mid N;\\ d<N}} \left(\underbrace{\sum_{e\mid N/d} \mu(e)}_{\substack{=\delta_{N/d,1} \\ \text{(by (13.84.3), applied} \\ \text{to } N/d \text{ instead of } n)}} - \underbrace{\mu(1)}_{=1}\right) d\varphi_{N/d}\left(\beta_d\right) + b_N \\
&= \sum_{\substack{d\mid N;\\ d<N}} \left(\underbrace{\delta_{N/d,1}}_{\substack{=0 \\ \text{(since } N/d>1)}} -1\right) d\varphi_{N/d}\left(\beta_d\right) + b_N = \sum_{\substack{d\mid N;\\ d<N}} (-1)\, d\varphi_{N/d}\left(\beta_d\right) + b_N \\
&= -\sum_{\substack{d\mid N;\\ d<N}} d\varphi_{N/d}\left(\beta_d\right) + b_N .
\end{aligned} \tag{13.84.19}$$

Now,

$$\begin{aligned}
\sum_{d\mid N} d\varphi_{N/d}\left(\beta_d\right) &= \sum_{\substack{d\mid N;\\ d<N}} d\varphi_{N/d}\left(\beta_d\right) + N \underbrace{\varphi_{N/N}}_{=\varphi_1=\operatorname{id}}\left(\underbrace{\beta_N}_{=t}\right) = \sum_{\substack{d\mid N;\\ d<N}} d\varphi_{N/d}\left(\beta_d\right) + \underbrace{Nt}_{\substack{=-\sum_{\substack{d\mid N;\\ d<N}} d\varphi_{N/d}(\beta_d)+b_N \\ \text{(by (13.84.19))}}} \\
&= \sum_{\substack{d\mid N;\\ d<N}} d\varphi_{N/d}\left(\beta_d\right) + \left(-\sum_{\substack{d\mid N;\\ d<N}} d\varphi_{N/d}\left(\beta_d\right) + b_N\right) = b_N .
\end{aligned}$$

Thus, (13.84.18) is satisfied for $n = N$. We have thus completed a step of our recursive construction of the family $\left(\beta_n\right)_{n\geq 1}$; this family therefore exists, and the implication $\mathcal{F} \Longrightarrow \mathcal{E}$ is proven.

Proof of the implication $\mathcal{F} \Longrightarrow \mathcal{G}$: Assume that Assertion $\mathcal{F}$ holds. That is, every positive integer n satisfies

$$\sum_{d\mid n} \mu(d)\,\varphi_d\left(b_{n/d}\right) \in nA. \tag{13.84.20}$$

On the other hand, every positive integer e satisfies

$$\sum_{d\mid e} \mu(d)\,\frac{e}{d} = \phi(e) \tag{13.84.21}$$

(by (13.84.3), applied to $n = e$). Now, every positive integer n satisfies

$$\begin{aligned}
\sum_{d\mid n} \phi(d)\,\varphi_d\left(b_{n/d}\right) &= \sum_{e\mid n} \underbrace{\phi(e)}_{\substack{=\sum_{d\mid e}\mu(d)\frac{e}{d} \\ \text{(by (13.84.21))}}} \varphi_e\left(b_{n/e}\right) \\
&= \underbrace{\sum_{e\mid n}\sum_{d\mid e}}_{=\sum_{d\mid n}\sum_{\substack{e\mid n;\\ d\mid e}}} \mu(d)\,\frac{e}{d}\varphi_e\left(\underbrace{b_{n/e}}_{=b_{(n/d)/(e/d)}}\right) = \sum_{d\mid n}\sum_{\substack{e\mid n;\\ d\mid e}} \mu(d)\,\frac{e}{d}\varphi_e\left(b_{(n/d)/(e/d)}\right) \\
&= \underbrace{\sum_{d\mid n}\sum_{e\mid n/d}}_{=\sum_{e\mid n}\sum_{d\mid n/e}} \mu(d)\,e\,\underbrace{\varphi_{ed}}_{=\varphi_e\circ\varphi_d}\left(\underbrace{b_{(n/d)/e}}_{=b_{(n/e)/d}}\right) \qquad \text{(here, we have substituted } ed \text{ for } e \text{ in the second sum)} \\
&= \sum_{e\mid n}\sum_{d\mid n/e} \mu(d)\,e\left(\varphi_e\circ\varphi_d\right)\left(b_{(n/e)/d}\right) = \sum_{e\mid n}\sum_{d\mid n/e} \mu(d)\,e\varphi_e\left(\varphi_d\left(b_{(n/e)/d}\right)\right) \\
&= \sum_{e\mid n} e \sum_{d\mid n/e} \mu(d)\,\varphi_e\left(\varphi_d\left(b_{(n/e)/d}\right)\right) = \sum_{e\mid n} e\varphi_e\underbrace{\left(\sum_{d\mid n/e}\mu(d)\,\varphi_d\left(b_{(n/e)/d}\right)\right)}_{\substack{\in(n/e)A \\ \text{(by (13.84.20), applied} \\ \text{to } n/e \text{ instead of } n)}} \qquad \text{(since } \varphi_e \text{ is } \mathbb{Z}\text{-linear)} \\
&\in \sum_{e\mid n} e \underbrace{\varphi_e\left((n/e)\,A\right)}_{\substack{\subset(n/e)A \\ \text{(since } \varphi_e \text{ is } \mathbb{Z}\text{-linear)}}} \subset \sum_{e\mid n} \underbrace{e\,(n/e)}_{=n} A = \sum_{e\mid n} nA \subset nA.
\end{aligned}$$

Thus, Assertion $\mathcal{G}$ holds. We have thus proven the implication $\mathcal{F} \Longrightarrow \mathcal{G}$.

Proof of the implication $\mathcal{G} \Longrightarrow \mathcal{F}$: Assume that Assertion $\mathcal{G}$ holds. That is, every positive integer n satisfies

$$\sum_{d\mid n} \phi(d)\,\varphi_d\left(b_{n/d}\right) \in nA. \tag{13.84.22}$$

On the other hand, every positive integer e satisfies $\sum_{d\mid e} d\mu(d)\,\phi\left(\frac{e}{d}\right) = \mu(e)$ (by (13.84.6), applied to $n = e$). In other words, every positive integer e satisfies

$$\mu(e) = \sum_{d\mid e} d\mu(d)\,\phi\left(\frac{e}{d}\right) = \sum_{d\mid e} \frac{e}{d}\mu\left(\frac{e}{d}\right)\phi(d) \tag{13.84.23}$$

(here, we have substituted e/d for d in the sum). Now, every positive integer n satisfies

$$\sum_{d|n} \mu(d)\varphi_d\left(b_{n/d}\right) = \sum_{e|n} \underbrace{\mu(e)}_{\substack{=\sum_{d|e}\frac{e}{d}\mu\left(\frac{e}{d}\right)\phi(d) \\ \text{(by (13.84.23))}}} \varphi_e\left(b_{n/e}\right)$$

$$= \underbrace{\sum_{e|n}\sum_{d|e}}_{=\sum_{d|n}\sum_{\substack{e|n;\\ d|e}}} \frac{e}{d}\mu\left(\frac{e}{d}\right)\phi(d)\varphi_e\left(\underbrace{b_{n/e}}_{=b_{(n/d)/(e/d)}}\right)$$

$$= \sum_{d|n}\sum_{\substack{e|n;\\ d|e}} \frac{e}{d}\mu\left(\frac{e}{d}\right)\phi(d)\varphi_e\left(b_{(n/d)/(e/d)}\right) = \underbrace{\sum_{d|n}\sum_{e|n/d}}_{=\sum_{e|n}\sum_{d|n/e}} e\mu(e)\phi(d)\underbrace{\varphi_{ed}}_{=\varphi_e\circ\varphi_d}\left(\underbrace{b_{(n/d)/e}}_{=b_{(n/e)/d}}\right)$$

(here, we have substituted ed for e in the second sum)

$$= \sum_{e|n}\sum_{d|n/e} e\mu(e)\phi(d)\left(\varphi_e\circ\varphi_d\right)\left(b_{(n/e)/d}\right) = \sum_{e|n}\sum_{d|n/e} e\mu(e)\phi(d)\varphi_e\left(\varphi_d\left(b_{(n/e)/d}\right)\right)$$

$$= \sum_{e|n} e\mu(e)\sum_{d|n/e}\phi(d)\varphi_e\left(\varphi_d\left(b_{(n/e)/d}\right)\right) = \sum_{e|n} e\mu(e)\varphi_e\underbrace{\left(\sum_{d|n/e}\phi(d)\varphi_d\left(b_{(n/e)/d}\right)\right)}_{\substack{\in(n/e)A \\ \text{(by (13.84.22), applied} \\ \text{to } n/e \text{ instead of } n)}} \qquad \text{(since } \varphi_e \text{ is } \mathbb{Z}\text{-linear)}$$

$$\in \sum_{e|n} e\mu(e)\underbrace{\varphi_e\left((n/e)A\right)}_{\substack{\subset(n/e)A \\ \text{(since } \varphi_e \text{ is } \mathbb{Z}\text{-linear)}}} \subset \sum_{e|n}\underbrace{e\mu(e)(n/e)}_{=n\mu(e)}A = \sum_{e|n} n\mu(e)A \subset nA.$$

Thus, Assertion $\mathcal{F}$ holds. We have thus proven the implication $\mathcal{G} \Longrightarrow \mathcal{F}$.

Proof of the equivalence $\mathcal{G} \Longleftrightarrow \mathcal{H}$*:* For every positive integer n, we have

$$\sum_{i=1}^{n}\varphi_{n/\gcd(i,n)}\left(b_{\gcd(i,n)}\right) = \sum_{d|n}\underbrace{\sum_{\substack{i\in\{1,2,\ldots,n\};\\ \gcd(i,n)=d}}\varphi_{n/d}\left(b_d\right)}_{\substack{=\phi\left(\frac{n}{d}\right)\varphi_{n/d}(b_d) \\ \text{(by (13.84.2))}}}$$

$$= \sum_{d|n}\phi\left(\frac{n}{d}\right)\varphi_{n/d}\left(b_d\right) = \sum_{d|n}\phi(d)\varphi_d\left(b_{n/d}\right)$$

(here, we have substituted n/d for d in the sum). This makes it clear that Assertions $\mathcal{G}$ and $\mathcal{H}$ are equivalent.

The implications and equivalences that we have proven, combined, yield the equivalence of all seven assertions $\mathcal{C}$, $\mathcal{D}$, $\mathcal{E}$, $\mathcal{F}$, $\mathcal{G}$, $\mathcal{H}$ and $\mathcal{J}$. This solves Exercise 2.9.6.

13.85. **Solution to Exercise 2.9.8.** *Solution to Exercise 2.9.8.* We know that the seven assertions $\mathcal{C}$, $\mathcal{D}$, $\mathcal{E}$, $\mathcal{F}$, $\mathcal{G}$, $\mathcal{H}$ and $\mathcal{J}$ are equivalent; hence, for each of our families, it suffices to prove one of these assertions. We choose the assertion $\mathcal{C}$, as it is the easiest to prove.

- *Proof of Assertion* $\mathcal{C}$ *for the family* $(b_n)_{n\geq 1} = (q^n)_{n\geq 1}$*, where* q *is a given integer:* Let q be an integer. We need to prove Assertion $\mathcal{C}$ for the family $(b_n)_{n\geq 1} = (q^n)_{n\geq 1}$. This means proving that for every positive integer n and every prime factor p of n, we have

$$\varphi_p\left(q^{n/p}\right) \equiv q^n \operatorname{mod} p^{v_p(n)}\mathbb{Z}. \tag{13.85.1}$$

So let us prove this. Let n be a positive integer, and let p be a prime factor of n. Fermat's little theorem yields $q^p \equiv q \operatorname{mod} p\mathbb{Z}$. Hence, (13.84.8) (applied to $A = \mathbb{Z}$, $a = q^p$, $b = q$ and $N = n/p$) yields $(q^p)^{n/p} \equiv q^{n/p} \operatorname{mod} p^{v_p(n/p)+1}\mathbb{Z}$. Since $(q^p)^{n/p} = q^n$ and $v_p(n/p) + 1 = v_p(n)$, this rewrites as $q^n \equiv q^{n/p} \operatorname{mod} p^{v_p(n)}\mathbb{Z}$. Now, $\varphi_p = \operatorname{id}$, so that $\varphi_p\left(q^{n/p}\right) = q^{n/p} \equiv q^n \operatorname{mod} p^{v_p(n)}\mathbb{Z}$. Thus, (13.85.1) is proven, and we are done with the family $(b_n)_{n\geq 1} = (q^n)_{n\geq 1}$.

- *Proof of Assertion $\mathcal{C}$ for the family $(b_n)_{n\geq 1} = (q)_{n\geq 1}$, where q is a given integer:* Let q be an integer. We need to prove Assertion $\mathcal{C}$ for the family $(b_n)_{n\geq 1} = (q)_{n\geq 1}$. This means proving that for every positive integer n and every prime factor p of n, we have
$$\varphi_p(q) \equiv q \operatorname{mod} p^{v_p(n)}\mathbb{Z}.$$
But this is obvious, since $\varphi_p = \operatorname{id}$. Thus, the family $(b_n)_{n\geq 1} = (q)_{n\geq 1}$ satisfies Assertion $\mathcal{C}$.
- *Proof of Assertion $\mathcal{C}$ for the family $(b_n)_{n\geq 1} = \left(\dbinom{qn}{rn}\right)_{n\geq 1}$, where $r \in \mathbb{Q}$ and $q \in \mathbb{Z}$ are given:* Let $r \in \mathbb{Q}$ and $q \in \mathbb{Z}$. We need to prove Assertion $\mathcal{C}$ for the family $(b_n)_{n\geq 1} = \left(\dbinom{qn}{rn}\right)_{n\geq 1}$. This means proving that for every positive integer n and every prime factor p of n, we have
$$\varphi_p\left(\binom{qn/p}{rn/p}\right) \equiv \binom{qn}{rn} \operatorname{mod} p^{v_p(n)}\mathbb{Z}. \tag{13.85.2}$$
So let us prove this. Let n be a positive integer, and p be a prime factor of n. We need to prove (13.85.2). In other words, we need to prove that
$$\binom{qn/p}{rn/p} \equiv \binom{qn}{rn} \operatorname{mod} p^{v_p(n)}\mathbb{Z} \tag{13.85.3}$$
(since $\varphi_p = \operatorname{id}$). We WLOG assume that $rn \in \mathbb{Z}$ (since otherwise, both sides of this congruence (13.85.3) are 0).

It is well-known that $(1+X)^p \equiv 1 + X^p \operatorname{mod} p\mathbb{Z}[X]$ in the polynomial ring $\mathbb{Z}[X]$. Hence, $((1+X)^p)^{n/p} \equiv (1+X^p)^{n/p} \operatorname{mod} p^{v_p(n/p)+1}\mathbb{Z}[X]$ (by (13.84.8), applied to $\mathbb{Z}[X]$, $(1+X)^p$, $1+X^p$ and n/p instead of A, a, b and N). Since $((1+X)^p)^{n/p} = (1+X)^n$ and $v_p(n/p) + 1 = v_p(n)$, this rewrites as
$$(1+X)^n \equiv (1+X^p)^{n/p} \operatorname{mod} p^{v_p(n)}\mathbb{Z}[X].$$
Hence,
$$(1+X)^n \equiv (1+X^p)^{n/p} \operatorname{mod} p^{v_p(n)}\mathbb{Z}[[X]]$$
(since $\mathbb{Z}[X]$ is a subring of $\mathbb{Z}[[X]]$). We can take both sides of this congruence to the q-th power[669], and thus obtain
$$((1+X)^n)^q \equiv \left((1+X^p)^{n/p}\right)^q \operatorname{mod} p^{v_p(n)}\mathbb{Z}[[X]].$$
In other words,
$$(1+X)^{qn} \equiv (1+X^p)^{qn/p} \operatorname{mod} p^{v_p(n)}\mathbb{Z}[[X]].$$
Comparing the coefficients before X^{rn} on both sides of this congruence, we obtain
$$\binom{qn}{rn} \equiv \binom{qn/p}{rn/p} \operatorname{mod} p^{v_p(n)}\mathbb{Z}.$$
This proves (13.85.3) and thus (13.85.2). Hence, we are done with the family $(b_n)_{n\geq 1} = \left(\dbinom{qn}{rn}\right)_{n\geq 1}$.

[*Remark:* There is an alternative, combinatorial approach to proving (13.85.3) when q is nonnegative. This approach proceeds by counting the (rn)-element subsets of the set $\mathbb{Z}/(qn)$. On the one

[669] This works even if q is negative, since $(1+X)^n$ and $(1+X^p)^{n/p}$ are invertible in $\mathbb{Z}[[X]] / \left(p^{v_p(n)}\mathbb{Z}[[X]]\right)$.

hand, the number of such subsets is clearly $\dbinom{qn}{rn}$. On the other hand, these subsets fall into two classes:

- the subsets which are invariant under the permutation
$$\mathbb{Z}/\left(qn\right) \to \mathbb{Z}/\left(qn\right), \qquad i \mapsto i + qn/p$$
of $\mathbb{Z}/\left(qn\right)$;
- the subsets which are not invariant under this permutation.

It is easy to see that the number of all subsets in the first class is $\dbinom{qn/p}{rn/p}$ (indeed, the intersection of such a subset with $\{0, 1, \ldots, qn/p - 1\} \subset \mathbb{Z}/\left(qn\right)$ must have rn/p elements, and uniquely determines the whole subset by "replication"), whereas the number of all subsets in the second class is divisible by $p^{v_p(n)}$ (because the permutation $\mathbb{Z}/\left(qn\right) \to \mathbb{Z}/\left(qn\right),\ i \mapsto i+1$ acts on these subsets, and thus splits them into orbits, each of which has size divisible by $p^{v_p(n)}$ [670]). Hence, the number of all subsets in both classes together is $\equiv \dbinom{qn/p}{rn/p} \operatorname{mod} p^{v_p(n)}\mathbb{Z}$. Comparing these two answers, we obtain $\dbinom{qn}{rn} \equiv \dbinom{qn/p}{rn/p} \operatorname{mod} p^{v_p(n)}\mathbb{Z}$.

The downside of this nice approach is that it requires a modification in the case when q is negative. Here, the *upper negation formula* $\dbinom{-a}{k} = (-1)^k \dbinom{a+k-1}{k}$ needs to be used, along with the "stars-and-bars" formula $\dbinom{a+k-1}{k}$ for the number of k-element multisets whose elements belong to $\{1, 2, \ldots, a\}$. We leave the details to the reader.]

- *Proof of Assertion $\mathcal{C}$ for the family $\left(b_n\right)_{n\geq 1} = \left(\dbinom{qn-1}{rn-1}\right)_{n\geq 1}$, where $r \in \mathbb{Z}$ and $q \in \mathbb{Z}$ are given:*

Let $r \in \mathbb{Z}$ and $q \in \mathbb{Z}$. We need to prove Assertion $\mathcal{C}$ for the family $\left(b_n\right)_{n\geq 1} = \left(\dbinom{qn-1}{rn-1}\right)_{n\geq 1}$. This means proving that for every positive integer n and every prime factor p of n, we have

$$\varphi_p \left(\dbinom{qn/p-1}{rn/p-1}\right) \equiv \dbinom{qn-1}{rn-1} \operatorname{mod} p^{v_p(n)}\mathbb{Z}. \tag{13.85.4}$$

[670] Here is why: Let O be an orbit under the action of the permutation $\mathbb{Z}/\left(qn\right) \to \mathbb{Z}/\left(qn\right),\ i \mapsto i+1$ on the subsets in the second class. We must prove that O has size divisible by $p^{v_p(n)}$.

Note that $qn \neq 0$, since otherwise there would be no subsets in the second class.

Let ξ be the permutation
$$\mathbb{Z}/\left(qn\right) \to \mathbb{Z}/\left(qn\right), \qquad i \mapsto i+1$$
of $\mathbb{Z}/\left(qn\right)$. Then, O is an orbit under the action of ξ on the subsets in the second class. Note that $\xi^{qn} = \operatorname{id}$, since ξ is a cyclic permutation of a qn-element set.

Fix an element $L \in O$. Thus, L is a subset in the second class. In other words, L is a subset of $\mathbb{Z}/\left(qn\right)$ that is not invariant under the permutation
$$\mathbb{Z}/\left(qn\right) \to \mathbb{Z}/\left(qn\right), \qquad i \mapsto i + qn/p$$
of $\mathbb{Z}/\left(qn\right)$. Since this permutation is $\xi^{qn/p}$, we can restate this as follows: L is a subset of $\mathbb{Z}/\left(qn\right)$ that is not invariant under the action of the permutation $\xi^{qn/p}$ on the subsets of $\mathbb{Z}/\left(qn\right)$. That is, $\xi^{qn/p}\left(L\right) \neq L$. But $\xi^{qn}\left(L\right) = L$ (since $\xi^{qn} = \operatorname{id}$). Now, O is the orbit of L under the action of ξ. Hence, the size of this orbit O is a divisor of qn (since $\xi^{qn}\left(L\right) = L$) but not a divisor of qn/p (since $\xi^{qn/p}\left(L\right) \neq L$). Thus, this size must be divisible by $p^{v_p(qn)}$ (because a divisor of qn that is not divisor of qn/p must necessarily be divisible by $p^{v_p(qn)}$). Hence, it is also divisible by $p^{v_p(n)}$ (since $v_p\left(qn\right) = \underbrace{v_p\left(q\right)}_{\geq 0} + v_p\left(n\right) \geq v_p\left(n\right)$ and therefore $p^{v_p(n)} \mid p^{v_p(qn)}$). Qed.

So let us prove this. Let n be a positive integer, and p be a prime factor of n. We need to prove (13.85.4). In other words, we need to prove that

$$\binom{qn/p-1}{rn/p-1} \equiv \binom{qn-1}{rn-1} \operatorname{mod} p^{v_p(n)}\mathbb{Z} \tag{13.85.5}$$

(since $\varphi_p = \mathrm{id}$). By the recurrence of the binomial coefficients, we have $\binom{qn/p-1}{rn/p-1} = \binom{qn/p}{rn/p} - \binom{qn/p-1}{rn/p}$ and $\binom{qn-1}{rn-1} = \binom{qn}{rn} - \binom{qn-1}{rn}$. Hence, we can obtain the desired congruence (13.85.5) by subtracting the congruence

$$\binom{qn/p-1}{rn/p} \equiv \binom{qn-1}{rn} \operatorname{mod} p^{v_p(n)}\mathbb{Z} \tag{13.85.6}$$

from the congruence (13.85.3). It therefore remains to prove the congruence (13.85.6) (since (13.85.3) has already been proven).

We recall the (easy-to-check) formula $\binom{a-1}{k} = (-1)^k \binom{k-a}{k}$ for every $a \in \mathbb{Z}$ and $k \in \mathbb{N}$. This formula allows us to rewrite both sides of (13.85.6), and thus (13.85.6) becomes

$$(-1)^{rn/p} \binom{rn/p - qn/p}{rn/p} \equiv (-1)^{rn} \binom{rn-qn}{rn} \operatorname{mod} p^{v_p(n)}\mathbb{Z}. \tag{13.85.7}$$

So it remains to prove (13.85.7). We have $(-1)^{rn/p} \equiv (-1)^{rn} \operatorname{mod} p^{v_p(n)}\mathbb{Z}$ [671]. Thus, we can cancel the $(-1)^{rn/p}$ on the left hand side of (13.85.7) against the $(-1)^{rn}$ on the right hand side, and therefore (13.85.7) takes the equivalent form

$$\binom{rn/p - qn/p}{rn/p} \equiv \binom{rn-qn}{rn} \operatorname{mod} p^{v_p(n)}\mathbb{Z}.$$

This further rewrites as

$$\binom{(r-q)\,n/p}{rn/p} \equiv \binom{(r-q)\,n}{rn} \operatorname{mod} p^{v_p(n)}\mathbb{Z}.$$

But this follows from (13.85.3), applied to $r-q$ instead of q. Hence, (13.85.7) is proven, and consequently Assertion $\mathcal{C}$ holds for the family $(b_n)_{n\geq 1} = \left(\binom{qn-1}{rn-1}\right)_{n\geq 1}$.

Exercise 2.9.8 is thus solved.

13.86. **Solution to Exercise 2.9.9.** *Solution to Exercise 2.9.9.* (a) This is obvious since $\mathbf{f}_n$ is an evaluation homomorphism (in an appropriate sense).[672]

(b) Let n and m be positive integers. Then,

$$\begin{aligned}(\mathbf{f}_n \circ \mathbf{f}_m)(a) &= \mathbf{f}_n \left(\underbrace{\mathbf{f}_m(a)}_{=a\left(x_1^m, x_2^m, x_3^m, \ldots\right)} \right) = \mathbf{f}_n \left(a\left(x_1^m, x_2^m, x_3^m, \ldots\right)\right) = \left(a\left(x_1^m, x_2^m, x_3^m, \ldots\right)\right)\left(x_1^n, x_2^n, x_3^n, \ldots\right) \\ &= a\left(\left(x_1^n\right)^m, \left(x_2^n\right)^m, \left(x_3^n\right)^m, \ldots\right) = a\left(x_1^{nm}, x_2^{nm}, x_3^{nm}, \ldots\right) = \mathbf{f}_{nm}(a)\end{aligned}$$

for all $a \in \Lambda$. Thus, $\mathbf{f}_n \circ \mathbf{f}_m = \mathbf{f}_{nm}$, so that Exercise 2.9.9(b) is solved.

[671] *Proof.* The only situation in which this is not obvious is when one of the integers rn/p and rn is even and the other is odd. This means, of course, that rn/p is odd and rn is even (since $rn/p \mid rn$). Hence, in this situation, we have $p = 2$ and $v_p(n) \leq 1$, and therefore $p^{v_p(n)} = 2$, so that $(-1)^{rn/p} \equiv (-1)^{rn} \operatorname{mod} p^{v_p(n)}\mathbb{Z}$ follows immediately from the fact that any two integer powers of -1 are congruent to each other modulo $2\mathbb{Z}$.

[672] Or, in a more down-to-earth fashion, this is obvious because (for example) multiplying two power series and then replacing all variables in the product by their n-th powers gives the same result as first replacing all variables by their n-th powers and then multiplying the resulting two power series.

(c) For every $a \in \Lambda$, we have $\mathbf{f}_1(a) = a\left(x_1^1, x_2^1, x_3^1, \ldots\right) = a$. Thus, $\mathbf{f}_1 = \operatorname{id}$, so that Exercise 2.9.9(c) is solved.

(d) There are several ways to solve Exercise 2.9.9(d):

- One can check $\Delta \circ \mathbf{f}_n = (\mathbf{f}_n \otimes \mathbf{f}_n) \circ \Delta$ rather directly on the basis $(m_\lambda)_{\lambda \in \operatorname{Par}}$ of Λ, using (2.1.4) and the definition of $\mathbf{f}_n$.
- One can check $\Delta \circ \mathbf{f}_n = (\mathbf{f}_n \otimes \mathbf{f}_n) \circ \Delta$ very easily on the elements $p_1, p_2, p_3, \ldots$ of Λ, and then "pretend" that these elements generate Λ (in fact, they do generate Λ when $\mathbf{k} = \mathbb{Q}$, so this solves Exercise 2.9.9(d) in the case of $\mathbf{k} = \mathbb{Q}$; but this easily entails that Exercise 2.9.9(d) holds in the case of $\mathbf{k} = \mathbb{Z}$ [673], and this, in turn, can be used to solve Exercise 2.9.9(d) for arbitrary $\mathbf{k}$ [674]).
- One can derive Exercise 2.9.9(d) from Exercise 2.9.10(e) using the self-duality of Λ and Exercise 2.9.10(f). This is somewhat complicated by the fact that $\mathbf{v}_n$ and $\mathbf{f}_n$ (generally) are not graded maps[675], but with the correct arguments, one can completely avoid any use of gradedness[676].

Probably the following solution of Exercise 2.9.9(d) is the shortest:

[673] Getting from $\mathbf{k} = \mathbb{Q}$ to $\mathbf{k} = \mathbb{Z}$ requires a standard functoriality argument. See Step 2 of the solution of Exercise 2.9.10(g) further below for an example of such an argument.

[674] Getting from $\mathbf{k} = \mathbb{Z}$ to arbitrary $\mathbf{k}$ requires a standard functoriality argument. See Step 3 of the solution of Exercise 2.9.10(g) further below for an example of such an argument.

[675] They nevertheless have a property close to being graded. Namely, let us say that a $\mathbf{k}$-linear map φ between two graded $\mathbf{k}$-modules $V = \bigoplus_{n \geq 0} V_n$ and $W = \bigoplus_{n \geq 0} W_n$ *scales the degree by* q (where q is some fixed rational number) if it has the property that $\varphi(V_n) \subset W_{qn}$ for all $n \in \mathbb{N}$ (where W_{qn} is understood to be 0 when qn is not an integer). Then, $\mathbf{v}_n$ scales the degree by $1/n$, whereas $\mathbf{f}_n$ scales the degree by n. When a $\mathbf{k}$-linear map φ between two graded $\mathbf{k}$-modules V and W scales the degree by a nonzero rational number q, the adjoint map $\varphi^* : W^* \to V^*$ restricts to a $\mathbf{k}$-linear map $W^o \to V^o$, which scales the degree by $1/q$; thus a large part of the theory of adjoint maps which is usually formulated for graded maps can be carried over to the case of maps scaling the degree by q.

[676] Here is a sketch of how these correct arguments look like.

The Hall inner product $(\cdot, \cdot)_\Lambda$ on Λ gives rise to a symmetric bilinear form $(\cdot, \cdot)_{\Lambda \otimes \Lambda}$ on $\Lambda \otimes \Lambda$ (according to Definition 3.1.2(b)). It is easy to see that every three elements a, b and c of Λ satisfy

$$(a, m(b \otimes c))_\Lambda = (\Delta(a), b \otimes c)_{\Lambda \otimes \Lambda}. \tag{13.86.1}$$

[*Proof of* (13.86.1)*:* Let a, b and c be three elements of Λ. We need to prove the equality (13.86.1). Since this equality is $\mathbf{k}$-linear in each of a, b and c, we can WLOG assume that each of a, b and c belongs to the basis $(s_\lambda)_{\lambda \in \operatorname{Par}}$ of the $\mathbf{k}$-module Λ. Assume this. Then, there exist three elements α, β and γ of Par such that $a = s_\alpha$, $b = s_\beta$ and $c = s_\gamma$. Consider these α, β and γ. Now, let us use the notations of Corollary 2.5.7. Applying Corollary 2.5.7 to $\lambda = \alpha$, $\mu = \beta$ and $\nu = \gamma$ yields $c^\alpha_{\beta,\gamma} = \hat{c}^\alpha_{\beta,\gamma}$. But the definition of the map m yields

$$m(b \otimes c) = \underbrace{b}_{=s_\beta}\underbrace{c}_{=s_\gamma} = s_\beta s_\gamma = \sum_{\lambda \in \operatorname{Par}} c^\lambda_{\beta,\gamma} s_\lambda$$

(by (2.5.6), applied to $\mu = \beta$ and $\nu = \gamma$). Hence,

$$\left(\underbrace{a}_{=s_\alpha}, \underbrace{m(b \otimes c)}_{=\sum_{\lambda \in \operatorname{Par}} c^\lambda_{\beta,\gamma} s_\lambda}\right)_\Lambda = \left(s_\alpha, \sum_{\lambda \in \operatorname{Par}} c^\lambda_{\beta,\gamma} s_\lambda\right)_\Lambda = \sum_{\lambda \in \operatorname{Par}} c^\lambda_{\beta,\gamma} \underbrace{(s_\alpha, s_\lambda)_\Lambda}_{=\delta_{\alpha,\lambda}} = \sum_{\lambda \in \operatorname{Par}} c^\lambda_{\beta,\gamma} \delta_{\alpha,\lambda} = c^\alpha_{\beta,\gamma} = \hat{c}^\alpha_{\beta,\gamma}.$$

On the other hand, applying the map Δ to both sides of the equality $a = s_\alpha$, we obtain

$$\Delta a = \Delta s_\alpha = \sum_{\mu \in \operatorname{Par};\ \nu \in \operatorname{Par}} \hat{c}^\alpha_{\mu,\nu} s_\mu \otimes s_\nu$$

(by (2.5.7), applied to $\lambda = \alpha$). Hence,

$$\begin{aligned}
\left(\underbrace{\Delta a}_{=\sum_{\mu \in \operatorname{Par};\ \nu \in \operatorname{Par}} \hat{c}^\alpha_{\mu,\nu} s_\mu \otimes s_\nu}, \underbrace{b}_{=s_\beta} \otimes \underbrace{c}_{=s_\gamma}\right)_{\Lambda \otimes \Lambda} &= \left(\sum_{\mu \in \operatorname{Par};\ \nu \in \operatorname{Par}} \hat{c}^\alpha_{\mu,\nu} s_\mu \otimes s_\nu, s_\beta \otimes s_\gamma\right)_{\Lambda \otimes \Lambda} \\
&= \sum_{\mu \in \operatorname{Par};\ \nu \in \operatorname{Par}} \hat{c}^\alpha_{\mu,\nu} \underbrace{(s_\mu \otimes s_\nu, s_\beta \otimes s_\gamma)_{\Lambda \otimes \Lambda}}_{=(s_\mu, s_\beta)_\Lambda (s_\nu, s_\gamma)_\Lambda} = \sum_{\mu \in \operatorname{Par};\ \nu \in \operatorname{Par}} \hat{c}^\alpha_{\mu,\nu} \underbrace{(s_\mu, s_\beta)_\Lambda}_{=\delta_{\mu,\beta}} \underbrace{(s_\nu, s_\gamma)_\Lambda}_{=\delta_{\nu,\gamma}} \\
&= \sum_{\mu \in \operatorname{Par};\ \nu \in \operatorname{Par}} \hat{c}^\alpha_{\mu,\nu} \delta_{\mu,\beta} \delta_{\nu,\gamma} = \sum_{\mu \in \operatorname{Par}} \delta_{\mu,\beta} \underbrace{\sum_{\nu \in \operatorname{Par}} \hat{c}^\alpha_{\mu,\nu} \delta_{\nu,\gamma}}_{=\hat{c}^\alpha_{\mu,\gamma}} = \sum_{\mu \in \operatorname{Par}} \delta_{\mu,\beta} \hat{c}^\alpha_{\mu,\gamma} = \hat{c}^\alpha_{\beta,\gamma}.
\end{aligned}$$

Compared with $(a, m(b \otimes c))_\Lambda = \hat{c}^\alpha_{\beta,\gamma}$, this yields $(a, m(b \otimes c))_\Lambda = (\Delta(a), b \otimes c)_{\Lambda \otimes \Lambda}$. Thus, (13.86.1) is proven.]

Fix $n \in \{1, 2, 3, \ldots\}$. We need to prove that $\mathbf{f}_n : \Lambda \to \Lambda$ is a Hopf algebra homomorphism. We already know that $\mathbf{f}_n$ is a $\mathbf{k}$-algebra homomorphism. Therefore, if we can show that $\mathbf{f}_n$ is a $\mathbf{k}$-coalgebra homomorphism, then it will immediately follow that $\mathbf{f}_n$ is a $\mathbf{k}$-bialgebra homomorphism and thus a Hopf algebra homomorphism (due to Corollary 1.4.27). Therefore, it remains to show that $\mathbf{f}_n$ is a $\mathbf{k}$-coalgebra homomorphism. To do so, we need to check that $\epsilon \circ \mathbf{f}_n = \epsilon$ and $\Delta \circ \mathbf{f}_n = (\mathbf{f}_n \otimes \mathbf{f}_n) \circ \Delta$. We shall only prove $\Delta \circ \mathbf{f}_n = (\mathbf{f}_n \otimes \mathbf{f}_n) \circ \Delta$, while the easy proof of $\epsilon \circ \mathbf{f}_n = \epsilon$ is left to the reader.

Let us first notice that Λ is a bialgebra, and therefore Δ is a $\mathbf{k}$-algebra homomorphism (by the axioms of a bialgebra). Also, $\mathbf{f}_n$ is a $\mathbf{k}$-algebra homomorphism. Thus, $\Delta \circ \mathbf{f}_n$ and $(\mathbf{f}_n \otimes \mathbf{f}_n) \circ \Delta$ are $\mathbf{k}$-algebra homomorphisms.

For every $m \in \mathbb{N}$, let us define an element $\widetilde{h}_m \in \mathbf{k}[[\mathbf{x}]]$ by $\widetilde{h}_m = \sum_{i_1 \leq i_2 \leq \cdots \leq i_m} x_{i_1}^n x_{i_2}^n \cdots x_{i_m}^n$. Then, it is easy to see that

$$\mathbf{f}_n(h_m) = \widetilde{h}_m \qquad \text{for every } m \in \mathbb{N}. \tag{13.86.6}$$

Now, we can show that every $a \in \Lambda$ and $B \in \Lambda \otimes \Lambda$ satisfy

$$(a, m(B))_\Lambda = (\Delta(a), B)_{\Lambda \otimes \Lambda}. \tag{13.86.2}$$

[*Proof of* (13.86.2)*:* Let $a \in \Lambda$ and $B \in \Lambda \otimes \Lambda$. We need to prove the equality (13.86.2). Since this equality is $\mathbf{k}$-linear in B, we can WLOG assume that B is a pure tensor (because the pure tensors span the $\mathbf{k}$-module $\Lambda \otimes \Lambda$). Assume this. Then, $B = b \otimes c$ for two elements b and c of Λ. Consider these b and c. Then,

$$\begin{aligned}\left(a, m\left(\underbrace{B}_{=b\otimes c}\right)\right)_\Lambda &= (a, m(b \otimes c))_\Lambda = \left(\Delta(a), \underbrace{b \otimes c}_{=B}\right)_{\Lambda \otimes \Lambda} \qquad (\text{by } (13.86.1)) \\ &= (\Delta(a), B)_{\Lambda \otimes \Lambda}.\end{aligned}$$

This proves (13.86.2).]

Exercise 2.9.10(f) yields that the maps $\mathbf{f}_n : \Lambda \to \Lambda$ and $\mathbf{v}_n : \Lambda \to \Lambda$ are adjoint with respect to the Hall inner product on Λ. Thus,

$$(\mathbf{f}_n a, b)_\Lambda = (a, \mathbf{v}_n b)_\Lambda \qquad \text{for every } a \in \Lambda \text{ and } b \in \Lambda. \tag{13.86.3}$$

From this, it is easy to see that

$$((\mathbf{f}_n \otimes \mathbf{f}_n) A, B)_{\Lambda \otimes \Lambda} = (A, (\mathbf{v}_n \otimes \mathbf{v}_n) B)_{\Lambda \otimes \Lambda} \qquad \text{for every } A \in \Lambda \otimes \Lambda \text{ and } B \in \Lambda \otimes \Lambda. \tag{13.86.4}$$

[*Proof of (13.86.4):* Let $A \in \Lambda \otimes \Lambda$ and $B \in \Lambda \otimes \Lambda$. We have to prove the equality (13.86.4). Since this equality is $\mathbf{k}$-linear in each of A and B, we can WLOG assume that A and B are pure tensors (since pure tensors span $\Lambda \otimes \Lambda$ as a $\mathbf{k}$-module). Assume this. Then, we can write A and B in the forms $A = a_1 \otimes a_2$ and $B = b_1 \otimes b_2$ for some $a_1, a_2 \in \Lambda$ and $b_1, b_2 \in \Lambda$. Consider these a_1, a_2 and b_1, b_2. Now,

$$\begin{aligned}\left((\mathbf{f}_n \otimes \mathbf{f}_n) \underbrace{A}_{=a_1 \otimes a_2}, \underbrace{B}_{=b_1 \otimes b_2}\right)_{\Lambda \otimes \Lambda} &= \left(\underbrace{(\mathbf{f}_n \otimes \mathbf{f}_n)(a_1 \otimes a_2)}_{=(\mathbf{f}_n a_1) \otimes (\mathbf{f}_n a_2)}, b_1 \otimes b_2\right)_{\Lambda \otimes \Lambda} = ((\mathbf{f}_n a_1) \otimes (\mathbf{f}_n a_2), b_1 \otimes b_2)_{\Lambda \otimes \Lambda} \\ &= \underbrace{(\mathbf{f}_n a_1, b_1)_\Lambda}_{\substack{=(a_1, \mathbf{v}_n b_1)_\Lambda \\ (\text{by } (13.86.3), \text{ applied to} \\ a=a_1 \text{ and } b=b_1)}} \underbrace{(\mathbf{f}_n a_2, b_2)_\Lambda}_{\substack{=(a_2, \mathbf{v}_n b_2)_\Lambda \\ (\text{by } (13.86.3), \text{ applied to} \\ a=a_2 \text{ and } b=b_2)}} \\ &= (a_1, \mathbf{v}_n b_1)_\Lambda (a_2, \mathbf{v}_n b_2)_\Lambda.\end{aligned}$$

Compared with

$$\begin{aligned}\left(\underbrace{A}_{=a_1 \otimes a_2}, (\mathbf{v}_n \otimes \mathbf{v}_n) \underbrace{B}_{=b_1 \otimes b_2}\right)_{\Lambda \otimes \Lambda} &= \left(a_1 \otimes a_2, \underbrace{(\mathbf{v}_n \otimes \mathbf{v}_n)(b_1 \otimes b_2)}_{=\mathbf{v}_n(b_1) \otimes \mathbf{v}_n(b_2)}\right)_{\Lambda \otimes \Lambda} \\ &= (a_1 \otimes a_2, \mathbf{v}_n(b_1) \otimes \mathbf{v}_n(b_2))_{\Lambda \otimes \Lambda} = (a_1, \mathbf{v}_n b_1)_\Lambda (a_2, \mathbf{v}_n b_2)_\Lambda,\end{aligned}$$

this yields $((\mathbf{f}_n \otimes \mathbf{f}_n) A, B)_{\Lambda \otimes \Lambda} = (A, (\mathbf{v}_n \otimes \mathbf{v}_n) B)_{\Lambda \otimes \Lambda}$. This proves (13.86.4).]

Let us now show that $(\mathbf{f}_n \otimes \mathbf{f}_n) \circ \Delta = \Delta \circ \mathbf{f}_n$ (this is one of the axioms that need to be checked in order to show that $\mathbf{f}_n$ is a Hopf algebra homomorphism).

Indeed, $\mathbf{v}_n$ is a $\mathbf{k}$-algebra homomorphism. Hence, $m \circ (\mathbf{v}_n \otimes \mathbf{v}_n) = \mathbf{v}_n \circ m$.

[677] As a consequence, $\widetilde{h}_m = \mathbf{f}_n(h_m) \in \Lambda$ for every $m \in \mathbb{N}$. (Of course, this was obvious anyway.)

Fix a positive integer r. Proposition 2.3.6(iii) (applied to r instead of n) states that $\Delta h_r = \sum_{i+j=r} h_i \otimes h_j$. The proof of this proposition (which proceeded by observing that $h_r(\mathbf{x}, \mathbf{y}) = \sum_{i+j=r} h_i(\mathbf{x}) h_j(\mathbf{y})$ in $\mathbf{k}[[\mathbf{x}, \mathbf{y}]]$) can be easily modified to obtain a proof of the equality $\Delta \widetilde{h}_r = \sum_{i+j=r} \widetilde{h}_i \otimes \widetilde{h}_j$. Thus, we have

$$\Delta \widetilde{h}_r = \sum_{i+j=r} \widetilde{h}_i \otimes \widetilde{h}_j . \tag{13.86.7}$$

Now,

$$(\Delta \circ \mathbf{f}_n)(h_r) = \Delta \left(\underbrace{\mathbf{f}_n(h_r)}_{\substack{= \widetilde{h}_r \\ \text{(by (13.86.6), applied to } m=r)}} \right) = \Delta \widetilde{h}_r = \sum_{i+j=r} \widetilde{h}_i \otimes \widetilde{h}_j \qquad \text{(by (13.86.7))} .$$

Now, let $a \in \Lambda$. Every $B \in \Lambda \otimes \Lambda$ satisfies

$$\begin{aligned}
\Big(\underbrace{((\mathbf{f}_n \otimes \mathbf{f}_n) \circ \Delta)(a)}_{= (\mathbf{f}_n \otimes \mathbf{f}_n)(\Delta a)}, B \Big)_{\Lambda \otimes \Lambda} &= ((\mathbf{f}_n \otimes \mathbf{f}_n)(\Delta a), B)_{\Lambda \otimes \Lambda} = (\Delta a, (\mathbf{v}_n \otimes \mathbf{v}_n) B)_{\Lambda \otimes \Lambda} \\
&\qquad \text{(by (13.86.4), applied to } A = \Delta a) \\
&= (\Delta a, (\mathbf{v}_n \otimes \mathbf{v}_n) B)_{\Lambda \otimes \Lambda} = \Big(a, \underbrace{m((\mathbf{v}_n \otimes \mathbf{v}_n) B)}_{= (m \circ (\mathbf{v}_n \otimes \mathbf{v}_n))(B)} \Big)_{\Lambda} \\
&\qquad \left(\begin{array}{c} \text{because } (a, m((\mathbf{v}_n \otimes \mathbf{v}_n) B))_{\Lambda} = (\Delta a, (\mathbf{v}_n \otimes \mathbf{v}_n) B)_{\Lambda \otimes \Lambda} \\ \text{(by (13.86.2), applied to } (\mathbf{v}_n \otimes \mathbf{v}_n) B \text{ instead of } B) \end{array} \right) \\
&= \Big(a, \underbrace{(m \circ (\mathbf{v}_n \otimes \mathbf{v}_n))}_{= \mathbf{v}_n \circ m}(B) \Big)_{\Lambda} = \Big(a, \underbrace{(\mathbf{v}_n \circ m)(B)}_{= \mathbf{v}_n(mB)} \Big)_{\Lambda} = (a, \mathbf{v}_n(mB))_{\Lambda} = (\mathbf{f}_n a, mB)_{\Lambda} \\
&\qquad \left(\begin{array}{c} \text{because } (\mathbf{f}_n a, mB)_{\Lambda} = (a, \mathbf{v}_n(mB))_{\Lambda} \\ \text{(by (13.86.3), applied to } mB \text{ instead of } b) \end{array} \right) \\
&= \Big(\underbrace{\Delta(\mathbf{f}_n a)}_{= (\Delta \circ \mathbf{f}_n)(a)}, B \Big)_{\Lambda \otimes \Lambda} \qquad \text{(by (13.86.2), applied to } \mathbf{f}_n a \text{ instead of } a) \\
&= ((\Delta \circ \mathbf{f}_n)(a), B)_{\Lambda \otimes \Lambda} .
\end{aligned} \tag{13.86.5}$$

Now, the bilinear form $(\cdot, \cdot)_{\Lambda \otimes \Lambda}$ is nondegenerate (in fact, $(s_\mu \otimes s_\nu)_{(\mu,\nu) \in \operatorname{Par} \times \operatorname{Par}}$ is an orthonormal basis with respect to this bilinear form). Hence, if U and V are two elements of $\Lambda \otimes \Lambda$ such that every $B \in \Lambda \otimes \Lambda$ satisfies $(U, B)_{\Lambda \otimes \Lambda} = (V, B)_{\Lambda \otimes \Lambda}$, then $U = V$. Applying this to $U = ((\mathbf{f}_n \otimes \mathbf{f}_n) \circ \Delta) a$ and $V = (\Delta \circ \mathbf{f}_n) a$, we obtain $((\mathbf{f}_n \otimes \mathbf{f}_n) \circ \Delta) a = (\Delta \circ \mathbf{f}_n) a$ (because of (13.86.5)). Since we have proven this for every $a \in \Lambda$, this yields that $(\mathbf{f}_n \otimes \mathbf{f}_n) \circ \Delta = \Delta \circ \mathbf{f}_n$. Thus, $(\mathbf{f}_n \otimes \mathbf{f}_n) \circ \Delta = \Delta \circ \mathbf{f}_n$ is proven.

[677] *Proof of (13.86.6):* Let $m \in \mathbb{N}$. Then, (2.2.3) (applied to m instead of n) yields $h_m = \sum_{i_1 \leq i_2 \leq \cdots \leq i_m} x_{i_1} x_{i_2} \cdots x_{i_m}$. Thus,

$$h_m(x_1^n, x_2^n, x_3^n, \ldots) = \left(\sum_{i_1 \leq i_2 \leq \cdots \leq i_m} x_{i_1} x_{i_2} \cdots x_{i_m} \right) (x_1^n, x_2^n, x_3^n, \ldots) = \sum_{i_1 \leq i_2 \leq \cdots \leq i_m} x_{i_1}^n x_{i_2}^n \cdots x_{i_m}^n .$$

But the definition of $\mathbf{f}_n$ yields $\mathbf{f}_n(h_m) = h_m(x_1^n, x_2^n, x_3^n, \ldots) = \sum_{i_1 \leq i_2 \leq \cdots \leq i_m} x_{i_1}^n x_{i_2}^n \cdots x_{i_m}^n$. This proves (13.86.6).

Compared with

$$\begin{aligned}\left(\left(\mathbf{f}_n\otimes\mathbf{f}_n\right)\circ\Delta\right)\left(h_r\right)&=\left(\mathbf{f}_n\otimes\mathbf{f}_n\right)\left(\underbrace{\Delta h_r}_{=\sum_{i+j=r}h_i\otimes h_j}\right)=\left(\mathbf{f}_n\otimes\mathbf{f}_n\right)\left(\sum_{i+j=r}h_i\otimes h_j\right)\\&=\sum_{i+j=r}\underbrace{\mathbf{f}_n\left(h_i\right)}_{\substack{=\widetilde{h}_i\\ \text{(by (13.86.6), applied to } m=i)}}\otimes\underbrace{\mathbf{f}_n\left(h_j\right)}_{\substack{=\widetilde{h}_j\\ \text{(by (13.86.6), applied to } m=j)}}=\sum_{i+j=r}\widetilde{h}_i\otimes\widetilde{h}_j,\end{aligned}$$

this yields $\left(\Delta\circ\mathbf{f}_n\right)\left(h_r\right)=\left(\left(\mathbf{f}_n\otimes\mathbf{f}_n\right)\circ\Delta\right)\left(h_r\right)$.

Now, let us forget that we fixed r. We thus have proven that

$$(13.86.8)\qquad \left(\Delta\circ\mathbf{f}_n\right)\left(h_r\right)=\left(\left(\mathbf{f}_n\otimes\mathbf{f}_n\right)\circ\Delta\right)\left(h_r\right)\qquad\text{for every positive integer } r.$$

Now, recall that the family $\left(h_r\right)_{r\geq 1}$ generates the $\mathbf{k}$-algebra Λ (according to Proposition 2.4.1). In other words, $\left(h_r\right)_{r\geq 1}$ is a generating set of the $\mathbf{k}$-algebra Λ. The two $\mathbf{k}$-algebra homomorphisms $\Delta\circ\mathbf{f}_n$ and $\left(\mathbf{f}_n\otimes\mathbf{f}_n\right)\circ\Delta$ are equal to each other on this generating set (according to (13.86.8)), and therefore must be identical (because if two $\mathbf{k}$-algebra homomorphisms from the same domain are equal to each other on a generating set of their domain, then these two homomorphisms must be identical). In other words, $\Delta\circ\mathbf{f}_n=\left(\mathbf{f}_n\otimes\mathbf{f}_n\right)\circ\Delta$. This completes the solution of Exercise 2.9.9(d).

(e) Let $m\in\mathbb{N}$. From (2.2.18), we have $\prod_{i=1}^{\infty}\left(1-x_it\right)^{-1}=\sum_{n\geq 0}h_n\left(\mathbf{x}\right)t^n$ in the ring $\Lambda\left[\left[t\right]\right]$. Substituting x_i^2 for x_i and t^2 for t in this equality, we obtain

$$(13.86.9)\qquad \prod_{i=1}^{\infty}\left(1-x_i^2t^2\right)^{-1}=\sum_{n\geq 0}h_n\left(x_1^2,x_2^2,x_3^2,\ldots\right)\left(t^2\right)^n.$$

But

$$(13.86.10)\qquad \prod_{i=1}^{\infty}\left(1-x_it\right)^{-1}=\sum_{n\geq 0}\underbrace{h_n\left(\mathbf{x}\right)}_{=h_n}t^n=\sum_{n\geq 0}h_nt^n.$$

Substituting $-t$ for t in this equality, we obtain $\prod_{i=1}^{\infty}\left(1-x_i\left(-t\right)\right)^{-1}=\sum_{n\geq 0}h_n\left(-t\right)^n=\sum_{n\geq 0}\left(-1\right)^nh_nt^n$. Thus,

$$(13.86.11)\qquad \sum_{n\geq 0}\left(-1\right)^nh_nt^n=\prod_{i=1}^{\infty}\left(\underbrace{1-x_i\left(-t\right)}_{=1+x_it}\right)^{-1}=\prod_{i=1}^{\infty}\left(1+x_it\right)^{-1}.$$

Now,

$$\begin{aligned}\sum_{n\geq 0}\underbrace{\mathbf{f}_2\left(h_n\right)}_{\substack{=h_n\left(x_1^2,x_2^2,x_3^2,\ldots\right)\\ \text{(by the definition of } \mathbf{f}_2)}}\underbrace{t^{2n}}_{=\left(t^2\right)^n}&=\sum_{n\geq 0}h_n\left(x_1^2,x_2^2,x_3^2,\ldots\right)\left(t^2\right)^n=\prod_{i=1}^{\infty}\left(\underbrace{1-x_i^2t^2}_{=\left(1-x_it\right)\left(1+x_it\right)}\right)^{-1}\qquad\text{(by (13.86.9))}\\&=\prod_{i=1}^{\infty}\left(\left(1-x_it\right)\left(1+x_it\right)\right)^{-1}=\underbrace{\left(\prod_{i=1}^{\infty}\left(1-x_it\right)^{-1}\right)}_{\substack{=\sum_{n\geq 0}h_nt^n\\ \text{(by (13.86.10))}}}\underbrace{\left(\prod_{i=1}^{\infty}\left(1+x_it\right)^{-1}\right)}_{\substack{=\sum_{n\geq 0}\left(-1\right)^nh_nt^n\\ \text{(by (13.86.11))}}}\\&=\left(\sum_{n\geq 0}h_nt^n\right)\left(\sum_{n\geq 0}\left(-1\right)^nh_nt^n\right)=\sum_{n\geq 0}\left(\sum_{i=0}^{n}h_i\cdot\left(-1\right)^{n-i}h_{n-i}\right)t^n\\&\qquad\text{(by the definition of the product of two formal power series)}.\end{aligned}$$

Comparing coefficients before t^{2m} on both sides of this equality, we obtain

$$\mathbf{f}_2\left(h_m\right) = \sum_{i=0}^{2m} h_i \cdot \underbrace{(-1)^{2m-i}}_{\substack{=(-1)^i \\ (\text{since } 2m-i \equiv i \bmod 2)}} h_{2m-i} = \sum_{i=0}^{2m} h_i \cdot (-1)^i h_{2m-i} = \sum_{i=0}^{2m} (-1)^i h_i h_{2m-i}.$$

Exercise 2.9.9(e) is thus solved.

(f) Let p be a prime number, and let $a \in \Lambda$. We need to prove that $\mathbf{f}_p\left(a\right) \equiv a^p \bmod p\Lambda$.

Indeed, let us first check that $\mathbf{f}_p\left(a\right) \equiv a^p \bmod p\mathbb{Z}\left[\left[\mathbf{x}\right]\right]$. Since $\mathbf{f}_p\left(a\right) = a\left(x_1^p, x_2^p, x_3^p, \ldots\right)$, this is equivalent to showing that $a\left(x_1^p, x_2^p, x_3^p, \ldots\right) \equiv a^p \bmod p\mathbb{Z}\left[\left[\mathbf{x}\right]\right]$. This, in turn, is equivalent to proving that $\overline{a}\left(x_1^p, x_2^p, x_3^p, \ldots\right) = \overline{a}^p$, where $\overline{a}$ denotes the projection of the power series $a \in \mathbb{Z}\left[\left[\mathbf{x}\right]\right]$ onto the ring $\left(\mathbb{Z}/p\mathbb{Z}\right)\left[\left[\mathbf{x}\right]\right]$ (by reducing every coefficient modulo p). So let us prove $\overline{a}\left(x_1^p, x_2^p, x_3^p, \ldots\right) = \overline{a}^p$ now.

Write the power series $\overline{a}$ in the form $\overline{a} = \sum_\beta \kappa_\beta \mathbf{x}^\beta$, where the sum ranges over all weak compositions β, and where κ_β is an element of $\mathbb{Z}/p\mathbb{Z}$ for every weak composition β. Taking both sides of this equality to the p-th power, we obtain

$$\overline{a}^p = \left(\sum_\beta \kappa_\beta \mathbf{x}^\beta\right)^p . \qquad (13.86.12)$$

In the commutative ring $\left(\mathbb{Z}/p\mathbb{Z}\right)\left[\left[\mathbf{x}\right]\right]$, we have $p \cdot 1_{\left(\mathbb{Z}/p\mathbb{Z}\right)\left[\left[\mathbf{x}\right]\right]} = 0$. Thus, taking the p-th power is a ring endomorphism of $\left(\mathbb{Z}/p\mathbb{Z}\right)\left[\left[\mathbf{x}\right]\right]$. This ring endomorphism is moreover continuous (with respect to the usual topology on $\left(\mathbb{Z}/p\mathbb{Z}\right)\left[\left[\mathbf{x}\right]\right]$) and $\left(\mathbb{Z}/p\mathbb{Z}\right)$-linear (by virtue of being a ring endomorphism). Thus, this endomorphism respects infinite $\left(\mathbb{Z}/p\mathbb{Z}\right)$-linear combinations; hence,

$$\left(\sum_\beta \kappa_\beta \mathbf{x}^\beta\right)^p = \sum_\beta \kappa_\beta \left(\mathbf{x}^\beta\right)^p . \qquad (13.86.13)$$

But $\overline{a} = \sum_\beta \kappa_\beta \mathbf{x}^\beta$, so that $\overline{a}\left(x_1^p, x_2^p, x_3^p, \ldots\right) = \sum_\beta \kappa_\beta \left(\mathbf{x}^\beta\right)^p$ (because replacing all variables $x_1, x_2, x_3, \ldots$ by their p-th powers transforms every monomial $\mathbf{x}^\beta$ into $\left(\mathbf{x}^\beta\right)^p$). Thus,

$$\begin{aligned} \overline{a}\left(x_1^p, x_2^p, x_3^p, \ldots\right) &= \sum_\beta \kappa_\beta \left(\mathbf{x}^\beta\right)^p = \left(\sum_\beta \kappa_\beta \mathbf{x}^\beta\right)^p \qquad (\text{by } (13.86.13)) \\ &= \overline{a}^p \qquad (\text{by } (13.86.12)) . \end{aligned}$$

We thus have proven that $\overline{a}\left(x_1^p, x_2^p, x_3^p, \ldots\right) = \overline{a}^p$. As explained above, this yields $\mathbf{f}_p\left(a\right) \equiv a^p \bmod p\mathbb{Z}\left[\left[\mathbf{x}\right]\right]$. In other words, the power series $\dfrac{\mathbf{f}_p\left(a\right) - a^p}{p}$ (this is, a priori, an element of $\mathbb{Q}\left[\left[\mathbf{x}\right]\right]$) belongs to $\mathbb{Z}\left[\left[\mathbf{x}\right]\right]$. Since this power series is also of bounded degree and symmetric (because so are $\mathbf{f}_p\left(a\right)$ and a^p), it follows that it lies in Λ. So we have $\dfrac{\mathbf{f}_p\left(a\right) - a^p}{p} \in \Lambda$, thus $\mathbf{f}_p\left(a\right) - a^p \in p\Lambda$ and hence $\mathbf{f}_p\left(a\right) \equiv a^p \bmod p\Lambda$. This solves Exercise 2.9.9(f).

(g) Set $\mathbf{k} = \mathbb{Z}$. Thus, the sign $\otimes$ will mean $\otimes_{\mathbb{Z}}$ in the remainder of this solution. Also, $\Lambda = \Lambda_{\mathbb{Z}}$. We define the notation $v_p\left(n\right)$ as in Exercise 2.9.6.

Let us introduce a notion from commutative algebra:

In the following, a *special Ψ-ring* will mean a pair $\left(A, \left(\varphi_n\right)_{n\in\{1,2,3,\ldots\}}\right)$, where A is a commutative ring and $\left(\varphi_n\right)_{n\in\{1,2,3,\ldots\}}$ is a family of ring endomorphisms $\varphi_n : A \to A$ of A satisfying the following properties:

- We have $\varphi_n \circ \varphi_m = \varphi_{nm}$ for any two positive integers n and m.
- We have $\varphi_1 = \mathrm{id}$.
- We have $\varphi_p\left(a\right) \equiv a^p \bmod pA$ for every $a \in A$ and every prime number p.

The tensor product of two special Ψ-rings $\left(A, \left(\varphi_n\right)_{n\in\{1,2,3,\ldots\}}\right)$ and $\left(B, \left(\psi_n\right)_{n\in\{1,2,3,\ldots\}}\right)$ is defined to be the pair $\left(A \otimes B, \left(\varphi_n \otimes \psi_n\right)_{n\in\{1,2,3,\ldots\}}\right)$. This pair $\left(A \otimes B, \left(\varphi_n \otimes \psi_n\right)_{n\in\{1,2,3,\ldots\}}\right)$ is a special Ψ-ring again[678].

[678] *Proof.* We need to prove the following five statements:

Here are three examples of special Ψ-rings which we will need:

- The pair $\left(\mathbb{Z}, (\mathrm{id})_{n\in\{1,2,3,\ldots\}}\right)$ is a special Ψ-ring. (The proof of this relies on Fermat's little theorem.)
- The pair $\left(\Lambda, (\mathbf{f}_n)_{n\in\{1,2,3,\ldots\}}\right)$ is a special Ψ-ring. (This follows from parts (a), (b), (c) and (f) of Exercise 2.9.9.)
- The pair $\left(\Lambda\otimes\Lambda, (\mathbf{f}_n\otimes\mathbf{f}_n)_{n\in\{1,2,3,\ldots\}}\right)$ is a special Ψ-ring. (Indeed, this pair is the tensor product of the special Ψ-ring $\left(\Lambda, (\mathbf{f}_n)_{n\in\{1,2,3,\ldots\}}\right)$ with itself.)

Statement 1: The ring $A\otimes B$ is a commutative ring.

Statement 2: The family $(\varphi_n\otimes\psi_n)_{n\in\{1,2,3,\ldots\}}$ is a family of ring endomorphisms $\varphi_n\otimes\psi_n : A\otimes B\to A\otimes B$ of $A\otimes B$.

Statement 3: We have $(\varphi_n\otimes\psi_n)\circ(\varphi_m\otimes\psi_n) = \varphi_{nm}\otimes\psi_{nm}$ for any two positive integers n and m.

Statement 4: We have $\varphi_1\otimes\psi_1 = \mathrm{id}$.

Statement 5: We have $(\varphi_p\otimes\psi_p)(a)\equiv a^p \bmod p(A\otimes B)$ for every $a\in A\otimes B$ and every prime number p.

Proof of Statement 1: This is obvious.

Proof of Statement 2: The family $(\varphi_n)_{n\in\{1,2,3,\ldots\}}$ is a family of ring endomorphisms of A (since $\left(A, (\varphi_n)_{n\in\{1,2,3,\ldots\}}\right)$ is a special Ψ-ring). Thus, φ_n is a ring endomorphism of A for every positive integer n. Similarly, ψ_n is a ring endomorphism of B for every positive integer n. Thus, $\varphi_n\otimes\psi_n$ is a ring endomorphism of $A\otimes B$ for every positive integer n. In other words, Statement 2 holds.

Proof of Statement 3: Let n and m be positive integers. Then, $\varphi_n\circ\varphi_m = \varphi_{nm}$ (since $\left(A, (\varphi_n)_{n\in\{1,2,3,\ldots\}}\right)$ is a special Ψ-ring) and $\psi_n\circ\psi_m = \psi_{nm}$ (similarly). Now, $(\varphi_n\otimes\psi_n)\circ(\varphi_m\otimes\psi_n) = \underbrace{(\varphi_n\circ\varphi_m)}_{=\varphi_{nm}}\otimes\underbrace{(\psi_n\circ\psi_m)}_{=\psi_{nm}} = \varphi_{nm}\otimes\psi_{nm}$, and thus Statement 3 is proven.

Proof of Statement 4: Since $\left(A, (\varphi_n)_{n\in\{1,2,3,\ldots\}}\right)$ is a special Ψ-ring, we have $\varphi_1 = \mathrm{id}$. Similarly, $\psi_1 = \mathrm{id}$. Thus, $\varphi_1\otimes\psi_1 = \mathrm{id}\otimes\mathrm{id} = \mathrm{id}$. This proves Statement 4.

Proof of Statement 5: Fix a prime number p. For every commutative ring R, we introduce three pieces of notation:

- We let $\overline{R}$ denote the commutative ring R/pR.
- We let π_R denote the canonical projection $R\to R/pR$.
- We let pow_R denote the map $R\to R$ which sends every $r\in R$ to r^p. This is not a linear map in general, but when $p\cdot 1_R = 0$, the map pow_R is a ring endomorphism of R.

Now, $\left(A, (\varphi_n)_{n\in\{1,2,3,\ldots\}}\right)$ is a special Ψ-ring. Hence, $\varphi_p(a)\equiv a^p \bmod pA$ for every $a\in A$. In other words, $\pi_A(\varphi_p(a)) = \pi_A\left(\underbrace{a^p}_{=\mathrm{pow}_A a}\right) = \pi_A(\mathrm{pow}_A a)$ for every $a\in A$. In other words, $\pi_A\circ\varphi_p = \pi_A\circ\mathrm{pow}_A$. But $\pi_A\circ\mathrm{pow}_A = \mathrm{pow}_{\overline{A}}\circ\pi_A$ (this is just saying that taking the p-th power commutes with the projection π_A). Hence,

$$\pi_A\circ\varphi_p = \pi_A\circ\mathrm{pow}_A = \mathrm{pow}_{\overline{A}}\circ\pi_A.$$

Similarly,

$$\pi_B\circ\psi_p = \pi_B\circ\mathrm{pow}_B = \mathrm{pow}_{\overline{B}}\circ\pi_B.$$

For every commutative ring R, the ring $\overline{R}$ satisfies $p\cdot 1_{\overline{R}} = 0$, and thus $\mathrm{pow}_{\overline{R}}$ is a ring endomorphism of $\overline{R}$. Hence, $\mathrm{pow}_{\overline{A}}$, $\mathrm{pow}_{\overline{B}}$ and $\mathrm{pow}_{\overline{A\otimes B}}$ are ring endomorphisms, thus $\mathbb{Z}$-algebra endomorphisms; consequently, the tensor product $\mathrm{pow}_{\overline{A}}\otimes\mathrm{pow}_{\overline{B}}$ makes sense. Moreover, we have $\mathrm{pow}_{\overline{A}}\otimes\mathrm{pow}_{\overline{B}} = \mathrm{pow}_{\overline{A\otimes B}}$ (because both $\mathrm{pow}_{\overline{A}}\otimes\mathrm{pow}_{\overline{B}}$ and $\mathrm{pow}_{\overline{A\otimes B}}$ are ring endomorphisms of $\overline{A}\otimes\overline{B}$ sending pure tensors $\alpha\otimes\beta$ to $\alpha^p\otimes\beta^p = (\alpha\otimes\beta)^p$, and this characterizes them uniquely).

There is a canonical isomorphism $\overline{A}\otimes\overline{B}\to\overline{A\otimes B}$ (because tensoring is right-exact). We identify $\overline{A\otimes B}$ with $\overline{A}\otimes\overline{B}$ along this isomorphism. Then, $\pi_{A\otimes B} = \pi_A\otimes\pi_B$. Thus,

$$\begin{aligned}
\underbrace{\pi_{A\otimes B}}_{=\pi_A\otimes\pi_B}\circ(\varphi_p\otimes\psi_p) &= (\pi_A\otimes\pi_B)\circ(\varphi_p\otimes\psi_p) = \underbrace{(\pi_A\circ\varphi_p)}_{=\mathrm{pow}_{\overline{A}}\circ\pi_A}\otimes\underbrace{(\pi_B\circ\psi_p)}_{=\mathrm{pow}_{\overline{B}}\circ\pi_B}\\
&= \left(\mathrm{pow}_{\overline{A}}\circ\pi_A\right)\otimes\left(\mathrm{pow}_{\overline{B}}\circ\pi_B\right) = \underbrace{\left(\mathrm{pow}_{\overline{A}}\otimes\mathrm{pow}_{\overline{B}}\right)}_{=\mathrm{pow}_{\overline{A\otimes B}}}\circ\underbrace{(\pi_A\otimes\pi_B)}_{=\pi_{A\otimes B}}\\
&= \mathrm{pow}_{\overline{A\otimes B}}\circ\pi_{A\otimes B} = \pi_{A\otimes B}\circ\mathrm{pow}_{A\otimes B}\\
&\qquad\text{(since taking the p-th power commutes with the projection $\pi_{A\otimes B}$)}.
\end{aligned}$$

Hence, every $a\in A\otimes B$ satisfies $\pi_{A\otimes B}((\varphi_p\otimes\psi_p)(a)) = \pi_{A\otimes B}\left(\mathrm{pow}_{A\otimes B}(a)\right)$. In other words, every $a\in A\otimes B$ satisfies $(\varphi_p\otimes\psi_p)(a)\equiv\mathrm{pow}_{A\otimes B}(a) = a^p \bmod p(A\otimes B)$. Thus, Statement 5 is proven.

It thus follows that $\left(A\otimes B, (\varphi_n\otimes\psi_n)_{n\in\{1,2,3,\ldots\}}\right)$ is a special Ψ-ring, qed.

Now, we will establish a pattern which we will follow in our new solutions to parts (b), (c), (d), (e) and (f) of Exercise 2.9.4. Namely, let $\left(A, (\varphi_n)_{n\in\{1,2,3,\ldots\}}\right)$ be a special Ψ-ring such that the $\mathbb{Z}$-module A is free. Then, A canonically injects into $A \otimes \mathbb{Q}$. Identify A with a subring of $A \otimes \mathbb{Q}$ using this injection. Also, consider $\Lambda = \Lambda_{\mathbb{Z}}$ as a subring of $\Lambda_{\mathbb{Q}}$ as in Exercise 2.9.4. Let $f : \Lambda_{\mathbb{Q}} \to A \otimes \mathbb{Q}$ be a $\mathbb{Q}$-algebra homomorphism such that every $n \in \{1, 2, 3, \ldots\}$ satisfies $f(p_n) \in A$. We can then ask for a criterion for $f(\Lambda) \subset A$. Using Exercise 2.9.6, we can obtain such an answer:

$$\left(\begin{array}{c}\text{We have } f(\Lambda) \subset A \text{ if every positive integer } n \text{ and every}\\ \text{prime factor } p \text{ of } n \text{ satisfy } \varphi_p\left(f\left(p_{n/p}\right)\right) \equiv f(p_n) \bmod p^{v_p(n)} A\end{array}\right). \tag{13.86.14}$$

[679] (The “if” here can be extended to “if and only if”, but we do not need the “only if” part.)

Let us now give alternative solutions to parts (b), (c), (d), (e) and (f) of Exercise 2.9.4:

Alternative solution to part (b) of Exercise 2.9.4: The $\mathbb{Z}$-module $\Lambda \otimes \Lambda$ is free, and thus canonically injects into $(\Lambda \otimes \Lambda) \otimes \mathbb{Q}$. We use this identification to regard $\Lambda \otimes \Lambda$ as a subring of $(\Lambda \otimes \Lambda) \otimes \mathbb{Q}$. We also identify $\Lambda_{\mathbb{Q}} \otimes_{\mathbb{Q}} \Lambda_{\mathbb{Q}}$ with $(\Lambda \otimes \Lambda) \otimes \mathbb{Q}$. Notice that every $n \in \{1, 2, 3, \ldots\}$ satisfies $\Delta_\times(p_n) = p_n \otimes p_n \in \Lambda \otimes \Lambda$.

We need to prove that $\Delta_\times(\Lambda) \subset \Lambda \otimes_{\mathbb{Z}} \Lambda$. In other words, we need to prove that $\Delta_\times(\Lambda) \subset \Lambda \otimes \Lambda$. This will follow from (13.86.14) (applied to $\left(A, (\varphi_n)_{n\in\{1,2,3,\ldots\}}\right) = \left(\Lambda \otimes \Lambda, (\mathbf{f}_n \otimes \mathbf{f}_n)_{n\in\{1,2,3,\ldots\}}\right)$ and $f = \Delta_\times$) once we have showed that every positive integer n and every prime factor p of n satisfy

$$(\mathbf{f}_p \otimes \mathbf{f}_p)\left(\Delta_\times\left(p_{n/p}\right)\right) \equiv \Delta_\times(p_n) \bmod p^{v_p(n)} (\Lambda \otimes \Lambda). \tag{13.86.15}$$

Thus, it remains to prove (13.86.15).

Let n be a positive integer, and let p be a prime factor of n. The definition of $\mathbf{f}_p\left(p_{n/p}\right)$ yields

$$\begin{aligned}\mathbf{f}_p\left(p_{n/p}\right) &= p_{n/p}\left(x_1^p, x_2^p, x_3^p, \ldots\right) = \left(x_1^{n/p}\right)^p + \left(x_2^{n/p}\right)^p + \left(x_3^{n/p}\right)^p + \cdots \\ &= x_1^n + x_2^n + x_3^n + \cdots = p_n.\end{aligned} \tag{13.86.16}$$

Now,

$$\begin{aligned}(\mathbf{f}_p \otimes \mathbf{f}_p)\left(\underbrace{\Delta_\times\left(p_{n/p}\right)}_{\substack{=p_{n/p}\otimes p_{n/p}\\ \text{(by the definition of } \Delta_\times)}}\right) &= (\mathbf{f}_p \otimes \mathbf{f}_p)\left(p_{n/p} \otimes p_{n/p}\right) = \underbrace{\mathbf{f}_p\left(p_{n/p}\right)}_{=p_n} \otimes \underbrace{\mathbf{f}_p\left(p_{n/p}\right)}_{=p_n} \\ &= p_n \otimes p_n = \Delta_\times(p_n) \qquad \text{(by the definition of } \Delta_\times).\end{aligned}$$

Hence, (13.86.15) holds. Thus, Exercise 2.9.4(b) is solved again.

Alternative solution to part (c) of Exercise 2.9.4: We identify $\mathbb{Z} \otimes \mathbb{Q}$ with $\mathbb{Q}$. Every $n \in \{1, 2, 3, \ldots\}$ satisfies $\epsilon_r(p_n) = r \in \mathbb{Z}$.

[679] *Proof of (13.86.14):* Exercise 2.9.6 can be applied to the family $(b_n)_{n\geq 1} = (f(p_n))_{n\geq 1}$ (indeed, the conditions of Exercise 2.9.6 are satisfied because $\left(A, (\varphi_n)_{n\in\{1,2,3,\ldots\}}\right)$ is a special Ψ-ring). As a result, we see that the Assertions $\mathcal{C}$, $\mathcal{D}$, $\mathcal{E}$, $\mathcal{F}$, $\mathcal{G}$, $\mathcal{H}$ and $\mathcal{J}$ for $(b_n)_{n\geq 1} = (f(p_n))_{n\geq 1}$ are equivalent.

Assume that every positive integer n and every prime factor p of n satisfy $\varphi_p\left(f\left(p_{n/p}\right)\right) \equiv f(p_n) \bmod p^{v_p(n)} A$. Then, the family $(f(p_n))_{n\geq 1} \in A^{\{1,2,3,\ldots\}}$ satisfies the Assertion $\mathcal{C}$ of Exercise 2.9.6. But since Assertion $\mathcal{C}$ is equivalent to Assertion $\mathcal{J}$, this yields that the family $(f(p_n))_{n\geq 1} \in A^{\{1,2,3,\ldots\}}$ satisfies the Assertion $\mathcal{J}$ as well. In other words, there exists a ring homomorphism $\Lambda_{\mathbb{Z}} \to A$ which, for every positive integer n, sends p_n to $f(p_n)$. Let g be such a ring homomorphism. Then, $g(p_n) = f(p_n)$ for every positive integer n.

The $\mathbb{Z}$-algebra homomorphism $g : \Lambda_{\mathbb{Z}} \to A$ can be extended to a $\mathbb{Q}$-algebra homomorphism $\Lambda_{\mathbb{Z}} \otimes \mathbb{Q} \to A \otimes \mathbb{Q}$ (by base change). Since $\Lambda_{\mathbb{Z}} \otimes \mathbb{Q} \cong \Lambda_{\mathbb{Q}}$ canonically, we can regard this latter $\mathbb{Q}$-algebra homomorphism as a $\mathbb{Q}$-algebra homomorphism $\Lambda_{\mathbb{Q}} \to A \otimes \mathbb{Q}$. Let us denote this $\mathbb{Q}$-algebra homomorphism $\Lambda_{\mathbb{Q}} \to A \otimes \mathbb{Q}$ by $\widetilde{g}$. Thus, $\widetilde{g}\mid_\Lambda = g$.

Now, both f and $\widetilde{g}$ are $\mathbb{Q}$-algebra homomorphisms $\Lambda_{\mathbb{Q}} \to A \otimes \mathbb{Q}$. These homomorphisms f and $\widetilde{g}$ are equal to each other on the elements $p_1, p_2, p_3, \ldots$ of $\Lambda_{\mathbb{Q}}$ (because for every positive integer n, we have $\widetilde{g}(p_n) = \underbrace{(\widetilde{g}\mid_\Lambda)}_{=g}(p_n) = g(p_n) = f(p_n)$). Since the elements $p_1, p_2, p_3, \ldots$ generate the $\mathbb{Q}$-algebra $\Lambda_{\mathbb{Q}}$, this forces said homomorphisms f and $\widetilde{g}$ to be identical. That is, we have $f = \widetilde{g}$. Hence, $f(\Lambda) = \widetilde{g}(\Lambda) = \underbrace{(\widetilde{g}\mid_\Lambda)}_{=g}(\Lambda) = g(\Lambda) \subset A$ (since the target of g is A). This proves (13.86.14).

We need to prove that $\epsilon_r(\Lambda) \subset \mathbb{Z}$. This will follow from (13.86.14) (applied to $\left(A, (\varphi_n)_{n\in\{1,2,3,\ldots\}}\right) = \left(\mathbb{Z}, (\mathrm{id})_{n\in\{1,2,3,\ldots\}}\right)$ and $f = \epsilon_r$) once we have showed that every positive integer n and every prime factor p of n satisfy

$$\mathrm{id}\left(\epsilon_r\left(p_{n/p}\right)\right) \equiv \epsilon_r(p_n) \bmod p^{v_p(n)}\mathbb{Z}.$$

But this congruence follows immediately from

$$\begin{aligned}\mathrm{id}\left(\epsilon_r\left(p_{n/p}\right)\right) = \epsilon_r\left(p_{n/p}\right) = r \qquad & \text{(by the definition of } \epsilon_r\text{)}\\ = \epsilon_r(p_n) \qquad & \text{(by the definition of } \epsilon_r\text{)}.\end{aligned}$$

Thus, Exercise 2.9.4(c) is solved again.

Alternative solution to part (d) of Exercise 2.9.4: The $\mathbb{Z}$-module Λ is free, and thus canonically injects into $\Lambda \otimes \mathbb{Q}$. We use this identification to regard Λ as a subring of $\Lambda \otimes \mathbb{Q}$. We also identify $\Lambda_{\mathbb{Q}}$ with $\Lambda \otimes \mathbb{Q}$. Notice that every $n \in \{1,2,3,\ldots\}$ satisfies $\mathbf{i}_r(p_n) = rp_n \in \Lambda$.

We need to prove that $\mathbf{i}_r(\Lambda) \subset \Lambda$. This will follow from (13.86.14) (applied to $\left(A, (\varphi_n)_{n\in\{1,2,3,\ldots\}}\right) = \left(\Lambda, (\mathbf{f}_n)_{n\in\{1,2,3,\ldots\}}\right)$ and $f = \mathbf{i}_r$) once we have showed that every positive integer n and every prime factor p of n satisfy

$$\mathbf{f}_p\left(\mathbf{i}_r\left(p_{n/p}\right)\right) \equiv \mathbf{i}_r(p_n) \bmod p^{v_p(n)}\Lambda.$$

But this congruence follows from

$$\mathbf{f}_p\left(\underbrace{\mathbf{i}_r\left(p_{n/p}\right)}_{\substack{=rp_{n/p}\\ \text{(by the definition of } \mathbf{i}_r\text{)}}}\right) = \mathbf{f}_p\left(rp_{n/p}\right) = r\underbrace{\mathbf{f}_p\left(p_{n/p}\right)}_{\substack{=p_n\\ \text{(by (13.86.16))}}} = rp_n = \mathbf{i}_r(p_n) \qquad \text{(by the definition of } \mathbf{i}_r\text{)}.$$

This solves Exercise 2.9.4(d) again.

Alternative solution to part (e) of Exercise 2.9.4: The $\mathbb{Z}$-module Λ is free, and thus canonically injects into $\Lambda \otimes \mathbb{Q}$. We use this identification to regard Λ as a subring of $\Lambda \otimes \mathbb{Q}$. We also identify $\Lambda_{\mathbb{Q}}$ with $\Lambda \otimes \mathbb{Q}$.

It is easy to see that the map Sq is a $\mathbb{Q}$-algebra homomorphism[680]. Every $n \in \{1,2,3,\ldots\}$ satisfies $\mathrm{Sq}(p_n) = p_n^2 \in \Lambda$.

We need to prove that $\mathrm{Sq}(\Lambda) \subset \Lambda$. This will follow from (13.86.14) (applied to $\left(A, (\varphi_n)_{n\in\{1,2,3,\ldots\}}\right) = \left(\Lambda, (\mathbf{f}_n)_{n\in\{1,2,3,\ldots\}}\right)$ and $f = \mathrm{Sq}$) once we have showed that every positive integer n and every prime factor p of n satisfy

$$\mathbf{f}_p\left(\mathrm{Sq}\left(p_{n/p}\right)\right) \equiv \mathrm{Sq}(p_n) \bmod p^{v_p(n)}\Lambda.$$

[680] *Proof.* We need to check that $\mathrm{Sq}(ab) = (\mathrm{Sq}\, a)(\mathrm{Sq}\, b)$ for any $a \in \Lambda_{\mathbb{Q}}$ and $b \in \Lambda_{\mathbb{Q}}$. Since Sq is $\mathbb{Q}$-linear, this only needs to be checked on a basis of the $\mathbb{Q}$-module $\Lambda_{\mathbb{Q}}$. For this we use the basis $(p_\lambda)_{\lambda\in\mathrm{Par}}$ of $\Lambda_{\mathbb{Q}}$. Checking the identity $\mathrm{Sq}(ab) = (\mathrm{Sq}\, a)(\mathrm{Sq}\, b)$ on this basis amounts to proving that $\mathrm{Sq}(p_\lambda p_\mu) = (\mathrm{Sq}(p_\lambda))(\mathrm{Sq}(p_\mu))$ for any two partitions λ and μ. So let λ and μ be two partitions. It is clear that there exists a partition ν such that $p_\lambda p_\mu = p_\nu$ (indeed, this ν is the partition obtained by sorting the list $\left(\lambda_1, \lambda_2, \ldots, \lambda_{\ell(\lambda)}, \mu_1, \mu_2, \ldots, \mu_{\ell(\mu)}\right)$ in decreasing order). Consider this ν. We have

$$\begin{aligned}\mathrm{Sq}\left(\underbrace{p_\lambda p_\mu}_{=p_\nu}\right) &= \mathrm{Sq}(p_\nu) = p_\nu^2 \qquad \text{(by the definition of Sq)}\\ &= (p_\lambda p_\mu)^2 \qquad \text{(since } p_\nu = p_\lambda p_\mu\text{)}\\ &= \underbrace{p_\lambda^2}_{\substack{=\mathrm{Sq}(p_\lambda)\\ \text{(by the definition of Sq)}}}\underbrace{p_\mu^2}_{\substack{=\mathrm{Sq}(p_\mu)\\ \text{(by the definition of Sq)}}} = (\mathrm{Sq}(p_\lambda))(\mathrm{Sq}(p_\mu)),\end{aligned}$$

which is what we wanted to prove. Thus we have checked that $\mathrm{Sq}(ab) = (\mathrm{Sq}\, a)(\mathrm{Sq}\, b)$ for any $a \in \Lambda_{\mathbb{Q}}$ and $b \in \Lambda_{\mathbb{Q}}$. Hence, Sq is a $\mathbb{Q}$-algebra homomorphism (since $\mathrm{Sq}(1) = 1$), qed.

But this congruence follows from

$$\mathbf{f}_p\left(\underbrace{\operatorname{Sq}\left(p_{n/p}\right)}_{\substack{=p_{n/p}^2\\ \text{(by the definition of Sq)}}}\right)=\mathbf{f}_p\left(p_{n/p}^2\right)=\left(\underbrace{\mathbf{f}_p\left(p_{n/p}\right)}_{\substack{=p_n\\ \text{(by (13.86.16))}}}\right)^2 \qquad \text{(since } \mathbf{f}_p \text{ is a ring homomorphism)}$$
$$=p_n^2=\operatorname{Sq}\left(p_n\right).$$

This solves Exercise 2.9.4(e) again.

Alternative solution to part (f) of Exercise 2.9.4: The $\mathbb{Z}$-module $\Lambda\otimes\Lambda$ is free, and thus canonically injects into $(\Lambda\otimes\Lambda)\otimes\mathbb{Q}$. We use this identification to regard $\Lambda\otimes\Lambda$ as a subring of $(\Lambda\otimes\Lambda)\otimes\mathbb{Q}$. We also identify $\Lambda_{\mathbb{Q}}\otimes_{\mathbb{Q}}\Lambda_{\mathbb{Q}}$ with $(\Lambda\otimes\Lambda)\otimes\mathbb{Q}$. Notice that every $n\in\{1,2,3,\ldots\}$ satisfies $\Delta_r(p_n)=\sum_{i=1}^{n-1}\binom{n}{i}p_i\otimes p_{n-i}+r\otimes p_n+p_n\otimes r\in\Lambda\otimes\Lambda$.

We need to prove that $\Delta_r(\Lambda)\subset\Lambda\otimes_{\mathbb{Z}}\Lambda$. In other words, we need to prove that $\Delta_r(\Lambda)\subset\Lambda\otimes\Lambda$. This will follow from (13.86.14) (applied to $\left(A,(\varphi_n)_{n\in\{1,2,3,\ldots\}}\right)=\left(\Lambda\otimes\Lambda,(\mathbf{f}_n\otimes\mathbf{f}_n)_{n\in\{1,2,3,\ldots\}}\right)$ and $f=\Delta_r$) once we have showed that every positive integer n and every prime factor p of n satisfy

$$(\mathbf{f}_p\otimes\mathbf{f}_p)\left(\Delta_r\left(p_{n/p}\right)\right)\equiv\Delta_r(p_n)\operatorname{mod}p^{v_p(n)}(\Lambda\otimes\Lambda). \tag{13.86.17}$$

Thus, it remains to prove (13.86.17).

Let n be a positive integer, and let p be a prime factor of n. For the sake of brevity, we denote r by p_0. Then,

$$\Delta_r(p_n)=\sum_{i=1}^{n-1}\binom{n}{i}p_i\otimes p_{n-i}+\underbrace{r}_{=p_0}\otimes p_n+p_n\otimes\underbrace{r}_{=p_0}$$
$$=\sum_{i=1}^{n-1}\binom{n}{i}p_i\otimes p_{n-i}+p_0\otimes p_n+p_n\otimes p_0=\sum_{i=0}^{n}\binom{n}{i}p_i\otimes p_{n-i}. \tag{13.86.18}$$

Applying this to n/p instead of n, we obtain

$$\Delta_r\left(p_{n/p}\right)=\sum_{i=0}^{n/p}\binom{n/p}{i}p_i\otimes p_{n/p-i}. \tag{13.86.19}$$

Now, we will need two elementary congruences for binomial coefficients:

- For any $i\in\mathbb{N}$ satisfying $p\mid i$, we have

$$\binom{n/p}{i/p}\equiv\binom{n}{i}\operatorname{mod}p^{v_p(n)}. \tag{13.86.20}$$

  (This follows from (13.85.3), applied to $q=1$ and $r=i/n$.)

- For any $i\in\mathbb{N}$ satisfying $p\nmid i$, we have

$$0\equiv\binom{n}{i}\operatorname{mod}p^{v_p(n)}. \tag{13.86.21}$$

  (This follows from (13.85.3), applied to $q=1$ and $r=i/n$, keeping in mind that $\binom{a}{b}=0$ if $b\notin\mathbb{N}$.)

Now, applying the map $\mathbf{f}_p \otimes \mathbf{f}_p$ to both sides of the equality (13.86.19), we obtain

$$\begin{aligned}
\left(\mathbf{f}_p \otimes \mathbf{f}_p\right)\left(\Delta_r\left(p_{n/p}\right)\right) &= \left(\mathbf{f}_p \otimes \mathbf{f}_p\right)\left(\sum_{i=0}^{n/p}\binom{n/p}{i} p_i \otimes p_{n/p-i}\right) \\
&= \sum_{i=0}^{n/p}\binom{n/p}{i} \underbrace{\mathbf{f}_p\left(p_i\right)}_{\substack{=\mathbf{f}_p\left(p_{pi/p}\right)=p_{pi} \\ \text{(by (13.86.16), applied to} \\ pi \text{ instead of } n)}} \otimes \underbrace{\mathbf{f}_p\left(p_{n/p-i}\right)}_{\substack{=\mathbf{f}_p\left(p_{(n-pi)/p}\right)=p_{n-pi} \\ \text{(by (13.86.16), applied to} \\ n-pi \text{ instead of } n)}} \\
&= \sum_{i=0}^{n/p}\binom{n/p}{i} p_{pi} \otimes p_{n-pi} = \sum_{i\in\{0,1,\ldots,n/p\}}\binom{n/p}{i} p_{pi} \otimes p_{n-pi} = \sum_{\substack{i\in\{0,1,\ldots,n\}; \\ p\mid i}}\binom{n/p}{i/p} p_i \otimes p_{n-i}
\end{aligned}$$

(here, we have substituted i/p for i in the sum). Comparing this with

$$\begin{aligned}
\Delta_r\left(p_n\right) &= \sum_{i=0}^{n}\binom{n}{i} p_i \otimes p_{n-i} \qquad \text{(by (13.86.18))} \\
&= \sum_{\substack{i\in\{0,1,\ldots,n\}; \\ p\mid i}} \underbrace{\binom{n}{i}}_{\substack{\equiv\binom{n/p}{i/p} \bmod p^{v_p(n)}(\Lambda\otimes\Lambda) \\ \text{(by (13.86.20))}}} p_i \otimes p_{n-i} + \sum_{\substack{i\in\{0,1,\ldots,n\}; \\ p\nmid i}} \underbrace{\binom{n}{i}}_{\substack{\equiv 0 \bmod p^{v_p(n)}(\Lambda\otimes\Lambda) \\ \text{(by (13.86.21))}}} p_i \otimes p_{n-i} \\
&\equiv \sum_{\substack{i\in\{0,1,\ldots,n\}; \\ p\mid i}}\binom{n/p}{i/p} p_i \otimes p_{n-i} + \underbrace{\sum_{\substack{i\in\{0,1,\ldots,n\}; \\ p\nmid i}} 0 p_i \otimes p_{n-i}}_{=0} \\
&= \sum_{\substack{i\in\{0,1,\ldots,n\}; \\ p\mid i}}\binom{n/p}{i/p} p_i \otimes p_{n-i} \bmod p^{v_p(n)}\left(\Lambda\otimes\Lambda\right),
\end{aligned}$$

we obtain

$$\left(\mathbf{f}_p \otimes \mathbf{f}_p\right)\left(\Delta_r\left(p_{n/p}\right)\right) \equiv \Delta_r\left(p_n\right) \bmod p^{v_p(n)}\left(\Lambda\otimes\Lambda\right).$$

Hence, (13.86.17) holds. Thus, Exercise 2.9.4(f) is solved again.

13.87. **Solution to Exercise 2.9.10.** *Solution to Exercise 2.9.10.* Let us first notice that every positive integer n satisfies

$$\text{(13.87.1)} \qquad \mathbf{v}_n\left(h_m\right) = \begin{cases} h_{m/n}, & \text{if } n \mid m; \\ 0, & \text{if } n \nmid m \end{cases} \qquad \text{for every } m \in \mathbb{N}$$

[681].

(b) Let n be a positive integer. Recall that a ring of formal power series $R\left[\left[t\right]\right]$ over a commutative ring R has a canonical topology which makes it into a topological R-algebra. Thus, in particular, $\Lambda\left[\left[t\right]\right]$ becomes a topological Λ-algebra. We shall be considering this topology when we speak of continuity.

[681] *Proof of (13.87.1):* Let n be a positive integer. Let $m \in \mathbb{N}$. We need to prove that (13.87.1) holds.

If $m \neq 0$, then m is a positive integer. Hence, if $m \neq 0$, then (13.87.1) follows immediately from the definition of $\mathbf{v}_n\left(h_m\right)$. Thus, for the rest of this proof of (13.87.1), we can WLOG assume that we don't have $m \neq 0$. Assume this.

We don't have $m \neq 0$. Thus, we have $m = 0$, so that $h_m = h_0 = 1$ and therefore $\mathbf{v}_n\left(\underbrace{h_m}_{=1}\right) = \mathbf{v}_n\left(1\right) = 1$ (since $\mathbf{v}_n$ is a $\mathbf{k}$-algebra homomorphism).

The $\mathbf{k}$-algebra homomorphism $\mathbf{v}_n : \Lambda \to \Lambda$ induces a continuous $\mathbf{k}$-algebra homomorphism $\mathbf{v}_n [[t]] : \Lambda [[t]] \to \Lambda [[t]]$ given by

$$(\mathbf{v}_n [[t]]) \left(\sum_{i \in \mathbb{N}} a_i t^i \right) = \sum_{i \in \mathbb{N}} \mathbf{v}_n (a_i) t^i \qquad \text{for all } (a_i)_{i \in \mathbb{N}} \in \Lambda^{\mathbb{N}}.$$

Consider the power series $H(t) \in \Lambda [[t]]$ and $E(t) \in \Lambda [[t]]$ defined in the proof of Proposition 2.4.1. We have

$$H(t) = 1 + h_1(\mathbf{x}) t + h_2(\mathbf{x}) t^2 + \cdots = \sum_{i \geq 0} h_i(\mathbf{x}) t^i = \sum_{i \geq 0} h_i t^i$$

and similarly $E(t) = \sum_{i \geq 0} e_i t^i$.

Substituting t^n for t in the equality $H(t) = \sum_{i \geq 0} h_i t^i$, we obtain $H(t^n) = \sum_{i \geq 0} h_i (t^n)^i$. Substituting $-t$ for t in the equality $E(t) = \sum_{i \geq 0} e_i t^i$, we obtain $E(-t) = \sum_{i \geq 0} e_i \underbrace{(-t)^i}_{=(-1)^i t^i} = \sum_{i \geq 0} (-1)^i e_i t^i$. Substituting $-t^n$ for t in the equality $E(t) = \sum_{i \geq 0} e_i t^i$, we obtain $E(-t^n) = \sum_{i \geq 0} e_i (-t^n)^i$.

Applying the map $\mathbf{v}_n [[t]]$ to both sides of the equality $H(t) = \sum_{i \geq 0} h_i t^i$, we obtain

$$\begin{aligned}
&(\mathbf{v}_n [[t]]) (H(t)) \\
&= (\mathbf{v}_n [[t]]) \left(\sum_{i \geq 0} h_i t^i \right) = \sum_{i \geq 0} \underbrace{\mathbf{v}_n (h_i)}_{\substack{= \begin{cases} h_{i/n}, & \text{if } n \mid i; \\ 0, & \text{if } n \nmid i \end{cases} \\ \text{(by (13.87.1), applied to } m = i)}} t^i \qquad \text{(by the definition of } \mathbf{v}_n [[t]]) \\
&= \sum_{i \geq 0} \begin{cases} h_{i/n}, & \text{if } n \mid i; \\ 0, & \text{if } n \nmid i \end{cases} \cdot t^i = \sum_{\substack{i \geq 0; \\ n \mid i}} \underbrace{\begin{cases} h_{i/n}, & \text{if } n \mid i; \\ 0, & \text{if } n \nmid i \end{cases}}_{\substack{= h_{i/n} \\ \text{(since } n \mid i)}} \cdot t^i + \sum_{\substack{i \geq 0; \\ n \nmid i}} \underbrace{\begin{cases} h_{i/n}, & \text{if } n \mid i; \\ 0, & \text{if } n \nmid i \end{cases}}_{\substack{= 0 \\ \text{(since } n \nmid i)}} \cdot t^i \\
&= \sum_{\substack{i \geq 0; \\ n \mid i}} h_{i/n} t^i + \underbrace{\sum_{\substack{i \geq 0; \\ n \nmid i}} 0 t^i}_{= 0} = \sum_{\substack{i \geq 0; \\ n \mid i}} h_{i/n} t^i = \sum_{i \geq 0} h_i \underbrace{t^{ni}}_{= (t^n)^i} \qquad \text{(here, we substituted } ni \text{ for } i \text{ in the sum)} \\
&= \sum_{i \geq 0} h_i (t^n)^i = H(t^n).
\end{aligned} \tag{13.87.2}$$

Also, $\begin{cases} h_{m/n}, & \text{if } n \mid m; \\ 0, & \text{if } n \nmid m \end{cases} = h_{m/n}$ (since $n \mid 0 = m$). Now,

$$\begin{aligned}
\mathbf{v}_n (h_m) = 1 = h_0 = h_{m/n} \qquad & \left(\text{since } h_{m/n} = h_0 \text{ (since } \underbrace{m}_{=0} / n = 0/n = 0) \right) \\
= \begin{cases} h_{m/n}, & \text{if } n \mid m; \\ 0, & \text{if } n \nmid m \end{cases} .
\end{aligned}$$

This proves (13.87.1).

But (2.4.3) yields $1 = E(-t) H(t)$, so that $E(-t) = \frac{1}{H(t)}$. Applying the map $\mathbf{v}_n[[t]]$ to both sides of this equality, we obtain

$$\begin{aligned}
(\mathbf{v}_n[[t]])(E(-t)) &= (\mathbf{v}_n[[t]])\left(\frac{1}{H(t)}\right) = \frac{1}{(\mathbf{v}_n[[t]])(H(t))} \qquad (\text{since } \mathbf{v}_n[[t]] \text{ is a } \mathbf{k}\text{-algebra homomorphism}) \\
&= \frac{1}{H(t^n)} \qquad (\text{by (13.87.2)}) \\
&= E(-t^n) \qquad \left(\begin{array}{c}\text{because } E(-t^n) = \dfrac{1}{H(t^n)} \text{ (this follows by substituting } t^n \\ \text{for } t \text{ in the equality } E(-t) = \dfrac{1}{H(t)})\end{array}\right) \\
&= \sum_{i \geq 0} e_i \underbrace{(-t^n)^i}_{=(-1)^i t^{ni}} = \sum_{i \geq 0} (-1)^i e_i t^{ni}. \qquad (13.87.3)
\end{aligned}$$

Now,

$$\begin{aligned}
&\sum_{i \geq 0} \begin{cases} (-1)^{i/n} e_{i/n}, & \text{if } n \mid i; \\ 0, & \text{if } n \nmid i \end{cases} \cdot t^i \\
&= \sum_{\substack{i \geq 0; \\ n \mid i}} \underbrace{\begin{cases} (-1)^{i/n} e_{i/n}, & \text{if } n \mid i; \\ 0, & \text{if } n \nmid i \end{cases}}_{\substack{=(-1)^{i/n} e_{i/n} \\ (\text{since } n \mid i)}} \cdot t^i + \sum_{\substack{i \geq 0; \\ n \nmid i}} \underbrace{\begin{cases} (-1)^{i/n} e_{i/n}, & \text{if } n \mid i; \\ 0, & \text{if } n \nmid i \end{cases}}_{\substack{=0 \\ (\text{since } n \nmid i)}} \cdot t^i \\
&= \sum_{\substack{i \geq 0; \\ n \mid i}} (-1)^{i/n} e_{i/n} t^i + \underbrace{\sum_{\substack{i \geq 0; \\ n \nmid i}} 0 t^i}_{=0} = \sum_{\substack{i \geq 0; \\ n \mid i}} (-1)^{i/n} e_{i/n} t^i \\
&= \sum_{i \geq 0} (-1)^i e_i t^{ni} \qquad (\text{here, we substituted } ni \text{ for } i \text{ in the sum}) \\
&= (\mathbf{v}_n[[t]])(E(-t)) \qquad (\text{by (13.87.3)}). \qquad (13.87.4)
\end{aligned}$$

But applying the map $\mathbf{v}_n[[t]]$ to both sides of the equality $E(-t) = \sum_{i \geq 0} (-1)^i e_i t^i$, we obtain

$$\begin{aligned}
(\mathbf{v}_n[[t]])(E(-t)) &= (\mathbf{v}_n[[t]])\left(\sum_{i \geq 0} (-1)^i e_i t^i\right) = \sum_{i \geq 0} \underbrace{\mathbf{v}_n\left((-1)^i e_i\right)}_{\substack{=(-1)^i \mathbf{v}_n(e_i) \\ (\text{since } \mathbf{v}_n \text{ is } \mathbf{k}\text{-linear})}} t^i \qquad (\text{by the definition of } \mathbf{v}_n[[t]]) \\
&= \sum_{i \geq 0} (-1)^i \mathbf{v}_n(e_i) t^i.
\end{aligned}$$

Hence,

$$\sum_{i \geq 0} (-1)^i \mathbf{v}_n(e_i) t^i = (\mathbf{v}_n[[t]])(E(-t)) = \sum_{i \geq 0} \begin{cases} (-1)^{i/n} e_{i/n}, & \text{if } n \mid i; \\ 0, & \text{if } n \nmid i \end{cases} \cdot t^i \qquad (13.87.5)$$

(by (13.87.4)).

Now, let m be a positive integer. Comparing coefficients before t^m in the equality (13.87.5), we obtain

$$(-1)^m \mathbf{v}_n(e_m) = \begin{cases} (-1)^{m/n} e_{m/n}, & \text{if } n \mid m; \\ 0, & \text{if } n \nmid m \end{cases}.$$

Dividing this by $(-1)^m$, we obtain

$$\mathbf{v}_n(e_m) = \frac{1}{(-1)^m} \begin{cases} (-1)^{m/n} e_{m/n}, & \text{if } n \mid m; \\ 0, & \text{if } n \nmid m \end{cases} = \begin{cases} \dfrac{1}{(-1)^m} (-1)^{m/n} e_{m/n}, & \text{if } n \mid m; \\ 0, & \text{if } n \nmid m \end{cases}$$
$$= \begin{cases} (-1)^{m-m/n} e_{m/n}, & \text{if } n \mid m; \\ 0, & \text{if } n \nmid m \end{cases}$$

(because if $n \mid m$, then $\dfrac{1}{(-1)^m}(-1)^{m/n} = (-1)^{m/n-m} = (-1)^{m-m/n}$). This solves Exercise 2.9.10(b).

(a) Let n be a positive integer. We define the continuous $\mathbf{k}$-algebra homomorphism $\mathbf{v}_n[[t]] : \Lambda[[t]] \to [[t]]$ as in the solution of Exercise 2.9.10(a).

Consider the power series $H(t) \in \Lambda[[t]]$ defined in the proof of Proposition 2.4.1. We have $H(t) = \sum_{i\geq 0} h_i t^i$ (this can be proven just as in the solution of Exercise 2.9.10(a)), so that $H'(t) = \sum_{i\geq 1} i h_i t^{i-1}$ (by the definition of the derivative of a power series). We can also see that the equality (13.87.2) holds (this can be proven just as in the solution of Exercise 2.9.10(a)). The power series $H(t)$ is invertible (since its constant term is $h_0 = 1$).

Exercise 2.5.21 yields $\sum_{m\geq 0} p_{m+1} t^m = \dfrac{H'(t)}{H(t)}$, so that

$$\frac{H'(t)}{H(t)} = \sum_{m\geq 0} p_{m+1} t^m = \sum_{i\geq 1} p_i t^{i-1} \qquad \text{(here, we substituted } i \text{ for } m+1 \text{ in the sum)}.$$

Multiplying this equality with t, we obtain

$$t \cdot \frac{H'(t)}{H(t)} = t \cdot \sum_{i\geq 1} p_i t^{i-1} = \sum_{i\geq 1} p_i \underbrace{t t^{i-1}}_{=t^i} = \sum_{i\geq 1} p_i t^i.$$

Multiplying this equality with $H(t)$, we obtain

$$t \cdot H'(t) = H(t) \cdot \sum_{i\geq 1} p_i t^i.$$

Hence,

$$H(t) \cdot \sum_{i\geq 1} p_i t^i = t \cdot \underbrace{H'(t)}_{=\sum_{i\geq 1} i h_i t^{i-1}} = t \cdot \sum_{i\geq 1} i h_i t^{i-1} = \sum_{i\geq 1} i h_i \underbrace{t t^{i-1}}_{=t^i} = \sum_{i\geq 1} i h_i t^i. \tag{13.87.6}$$

Substituting t^n for t in this equality, we obtain

$$H(t^n) \cdot \sum_{i\geq 1} p_i (t^n)^i = \sum_{i\geq 0} i h_i (t^n)^i. \tag{13.87.7}$$

Applying the map $\mathbf{v}_n[[t]]$ to both sides of the equality (13.87.6), we obtain

$$\begin{aligned}
&(\mathbf{v}_n[[t]])\left(H(t)\cdot\sum_{i\geq1}p_it^i\right)\\
&=(\mathbf{v}_n[[t]])\left(\sum_{i\geq1}ih_it^i\right)=\sum_{i\geq1}i\underbrace{(\mathbf{v}_n[[t]])(h_i)}_{\substack{=\mathbf{v}_n(h_i)\\ \text{(by the definition of }\mathbf{v}_n[[t]])}}\underbrace{(\mathbf{v}_n[[t]])\left(t^i\right)}_{\substack{=t^i\\ \text{(by the definition of }\mathbf{v}_n[[t]])}}\\
&\qquad\qquad\text{(since $\mathbf{v}_n[[t]]$ is a continuous $\mathbf{k}$-algebra homomorphism)}\\
&=\sum_{i\geq1}i\underbrace{\mathbf{v}_n(h_i)}_{\substack{=\begin{cases}h_{i/n}, & \text{if } n\mid i;\\ 0, & \text{if } n\nmid i\end{cases}\\ \text{(by the definition of }\mathbf{v}_n)}}t^i\\
&=\sum_{i\geq1}i\begin{cases}h_{i/n}, & \text{if } n\mid i;\\ 0, & \text{if } n\nmid i\end{cases}\cdot t^i=\sum_{\substack{i\geq1;\\ n\mid i}}i\underbrace{\begin{cases}h_{i/n}, & \text{if } n\mid i;\\ 0, & \text{if } n\nmid i\end{cases}}_{\substack{=h_{i/n}\\ \text{(since } n\mid i)}}\cdot t^i+\sum_{\substack{i\geq1;\\ n\nmid i}}i\underbrace{\begin{cases}h_{i/n}, & \text{if } n\mid i;\\ 0, & \text{if } n\nmid i\end{cases}}_{\substack{=0\\ \text{(since } n\nmid i)}}\cdot t^i\\
&=\sum_{\substack{i\geq1;\\ n\mid i}}ih_{i/n}t^i+\underbrace{\sum_{\substack{i\geq1;\\ n\nmid i}}0t^i}_{=0}=\sum_{\substack{i\geq1;\\ n\mid i}}ih_{i/n}t^i=\sum_{\substack{i\geq0;\\ n\mid i}}ih_{i/n}t^i\\
&\qquad\qquad\left(\text{since }\sum_{\substack{i\geq0;\\ n\mid i}}ih_{i/n}t^i=\sum_{\substack{i\geq1;\\ n\mid i}}ih_{i/n}t^i+\underbrace{0h_{0/n}t^0}_{=0}=\sum_{\substack{i\geq1;\\ n\mid i}}ih_{i/n}t^i\right)\\
&=\sum_{i\geq0}nih_i\underbrace{t^{ni}}_{=(t^n)^i}\qquad\qquad\text{(here, we substituted ni for i in the sum)}\\
&=\sum_{i\geq0}nih_i\left(t^n\right)^i=n\underbrace{\sum_{i\geq0}ih_i\left(t^n\right)^i}_{\substack{=H(t^n)\cdot\sum_{i\geq1}p_i(t^n)^i\\ \text{(by (13.87.7))}}}=n\cdot H\left(t^n\right)\cdot\sum_{i\geq1}p_i\left(t^n\right)^i=H\left(t^n\right)\cdot n\cdot\sum_{i\geq1}p_i\left(t^n\right)^i.
\end{aligned}$$

Compared with

$$\begin{aligned}
&(\mathbf{v}_n[[t]])\left(H(t)\cdot\sum_{i\geq1}p_it^i\right)\\
&=\underbrace{(\mathbf{v}_n[[t]])(H(t))}_{\substack{=H(t^n)\\ \text{(by (13.87.2))}}}\cdot(\mathbf{v}_n[[t]])\left(\sum_{i\geq1}p_it^i\right)\qquad\qquad\text{(since $\mathbf{v}_n[[t]]$ is a $\mathbf{k}$-algebra homomorphism)}\\
&=H\left(t^n\right)\cdot(\mathbf{v}_n[[t]])\left(\sum_{i\geq1}p_it^i\right),
\end{aligned}$$

this becomes

$$H\left(t^n\right)\cdot n\cdot\sum_{i\geq1}p_i\left(t^n\right)^i=H\left(t^n\right)\cdot(\mathbf{v}_n[[t]])\left(\sum_{i\geq1}p_it^i\right).$$

We can divide both sides of this equality by $H\left(t^{n}\right)$ (since $H\left(t^{n}\right)$ is invertible[682]). As a result, we obtain

$$n\cdot\sum_{i\geq 1}p_{i}\left(t^{n}\right)^{i}=\left(\mathbf{v}_{n}\left[\left[t\right]\right]\right)\left(\sum_{i\geq 1}p_{i}t^{i}\right)=\sum_{i\geq 1}\underbrace{\left(\mathbf{v}_{n}\left[\left[t\right]\right]\right)\left(p_{i}\right)}_{\substack{=\mathbf{v}_{n}\left(p_{i}\right)\\ \text{(by the definition of }\mathbf{v}_{n}[[t]])}}\underbrace{\left(\mathbf{v}_{n}\left[\left[t\right]\right]\right)\left(t^{i}\right)}_{\substack{=t^{i}\\ \text{(by the definition of }\mathbf{v}_{n}[[t]])}}$$

$$\text{(since }\mathbf{v}_{n}\left[\left[t\right]\right]\text{ is a continuous }\mathbf{k}\text{-algebra homomorphism)}$$

$$=\sum_{i\geq 1}\mathbf{v}_{n}\left(p_{i}\right)t^{i}. \tag{13.87.8}$$

On the other hand,

$$\begin{aligned}
&\sum_{i\geq 1}\begin{cases}np_{i/n}, & \text{if } n\mid i;\\ 0, & \text{if } n\nmid i\end{cases}\cdot t^{i}\\
&=\sum_{\substack{i\geq 1;\\ n\mid i}}\underbrace{\begin{cases}np_{i/n}, & \text{if } n\mid i;\\ 0, & \text{if } n\nmid i\end{cases}}_{\substack{=np_{i/n}\\ \text{(since } n\mid i)}}\cdot t^{i}+\sum_{\substack{i\geq 1;\\ n\nmid i}}\underbrace{\begin{cases}np_{i/n}, & \text{if } n\mid i;\\ 0, & \text{if } n\nmid i\end{cases}}_{\substack{=0\\ \text{(since } n\nmid i)}}\cdot t^{i}\\
&=\sum_{\substack{i\geq 1;\\ n\mid i}}np_{i/n}t^{i}+\underbrace{\sum_{\substack{i\geq 1;\\ n\nmid i}}0t^{i}}_{=0}=\sum_{\substack{i\geq 1;\\ n\mid i}}np_{i/n}t^{i}\\
&=\sum_{i\geq 1}np_{i}\underbrace{t^{ni}}_{=\left(t^{n}\right)^{i}}\\
&\qquad\text{(here, we have substituted } ni \text{ for } i \text{ in the sum)}\\
&=\sum_{i\geq 1}np_{i}\left(t^{n}\right)^{i}=n\cdot\sum_{i\geq 1}p_{i}\left(t^{n}\right)^{i}\\
&=\sum_{i\geq 1}\mathbf{v}_{n}\left(p_{i}\right)t^{i}\qquad\text{(by (13.87.8))}.
\end{aligned}\tag{13.87.9}$$

Now, let m be a positive integer. Comparing coefficients before t^{m} in the equality (13.87.9), we obtain

$$\begin{cases}np_{m/n}, & \text{if } n\mid m;\\ 0, & \text{if } n\nmid m\end{cases}=\mathbf{v}_{n}\left(p_{m}\right).$$

This solves Exercise 2.9.10(a).

(c) Fix two positive integers n and m. We need to prove that $\mathbf{v}_{n}\circ\mathbf{v}_{m}=\mathbf{v}_{nm}$. But both $\mathbf{v}_{n}\circ\mathbf{v}_{m}$ and $\mathbf{v}_{nm}$ are $\mathbf{k}$-algebra homomorphisms (since $\mathbf{v}_{n}$, $\mathbf{v}_{m}$ and $\mathbf{v}_{nm}$ are $\mathbf{k}$-algebra homomorphisms).

Let us now show that

$$\left(\mathbf{v}_{n}\circ\mathbf{v}_{m}\right)\left(h_{r}\right)=\mathbf{v}_{nm}\left(h_{r}\right)\qquad\text{for every positive integer } r. \tag{13.87.10}$$

[682] *Proof.* We know that the power series $H\left(t\right)$ is invertible. That is, there exists a power series $A\left(t\right)\in\Lambda\left[\left[t\right]\right]$ such that $A\left(t\right)\cdot H\left(t\right)=1$. Consider this A. Substituting t^{n} for t in the equality $A\left(t\right)\cdot H\left(t\right)=1$, we obtain $A\left(t^{n}\right)\cdot H\left(t^{n}\right)=1$. Hence, the power series $H\left(t^{n}\right)$ is invertible (and its inverse is $A\left(t^{n}\right)$), qed.

Proof of (13.87.10): Let r be a positive integer. If $m \nmid r$, then (13.87.10) holds[683]. Thus, for the rest of this proof, we can WLOG assume that we don't have $m \nmid r$. Assume this.

We have $m \mid r$ (since we don't have $m \nmid r$). Thus, r/m is a positive integer. By the definition of $\mathbf{v}_m$, we have $\mathbf{v}_m(h_r) = \begin{cases} h_{r/m}, & \text{if } m \mid r; \\ 0, & \text{if } m \nmid r \end{cases} = h_{r/m}$ (since $m \mid r$). Now,

$$\begin{aligned}
(\mathbf{v}_n \circ \mathbf{v}_m)(h_r) &= \mathbf{v}_n \left(\underbrace{\mathbf{v}_m(h_r)}_{=h_{r/m}} \right) = \mathbf{v}_n\left(h_{r/m}\right) \\
&= \begin{cases} h_{(r/m)/n}, & \text{if } n \mid r/m; \\ 0, & \text{if } n \nmid r/m \end{cases} \qquad \text{(by the definition of } \mathbf{v}_n\text{)} \\
&= \begin{cases} h_{(r/m)/n}, & \text{if } nm \mid r; \\ 0, & \text{if } nm \nmid r \end{cases} \\
&\qquad \left(\begin{array}{c} \text{since the condition } n \mid r/m \text{ is equivalent to the condition } nm \mid r, \\ \text{and since the condition } n \nmid r/m \text{ is equivalent to the condition } nm \nmid r \end{array} \right) \\
&= \begin{cases} h_{r/(nm)}, & \text{if } nm \mid r; \\ 0, & \text{if } nm \nmid r \end{cases} \qquad \left(\text{since } (r/m)/n = r/(nm)\right).
\end{aligned}$$

Compared with

$$\mathbf{v}_{nm}(h_r) = \begin{cases} h_{r/(nm)}, & \text{if } nm \mid r; \\ 0, & \text{if } nm \nmid r \end{cases} \qquad \left(\text{by the definition of } \mathbf{v}_{nm}\right),$$

this yields $(\mathbf{v}_n \circ \mathbf{v}_m)(h_r) = \mathbf{v}_{nm}(h_r)$. This proves (13.87.10).

Now, recall that the family $(h_r)_{r\geq 1}$ generates the $\mathbf{k}$-algebra Λ (according to Proposition 2.4.1). In other words, $(h_r)_{r\geq 1}$ is a generating set of the $\mathbf{k}$-algebra Λ. The two $\mathbf{k}$-algebra homomorphisms $\mathbf{v}_n \circ \mathbf{v}_m$ and $\mathbf{v}_{nm}$ are equal to each other on this generating set (according to (13.87.10)), and therefore must be identical (because if two $\mathbf{k}$-algebra homomorphisms from the same domain are equal to each other on a generating set of their domain, then these two homomorphisms must be identical). In other words, $\mathbf{v}_n \circ \mathbf{v}_m = \mathbf{v}_{nm}$. This solves Exercise 2.9.10(c).

(d) For every positive integer r, we have

$$\begin{aligned}
\mathbf{v}_1(h_r) &= \begin{cases} h_{r/1}, & \text{if } 1 \mid r; \\ 0, & \text{if } 1 \nmid r \end{cases} \qquad \left(\text{by the definition of } \mathbf{v}_1\right) \\
&= h_{r/1} \qquad (\text{since } 1 \mid r) \\
&= h_r = \operatorname{id}(h_r). \qquad \qquad (13.87.11)
\end{aligned}$$

Now, recall that the family $(h_r)_{r\geq 1}$ generates the $\mathbf{k}$-algebra Λ (according to Proposition 2.4.1). In other words, $(h_r)_{r\geq 1}$ is a generating set of the $\mathbf{k}$-algebra Λ. The two $\mathbf{k}$-algebra homomorphisms $\mathbf{v}_1$ and id are equal to each other on this generating set (according to (13.87.11)), and therefore must be identical (because if two $\mathbf{k}$-algebra homomorphisms from the same domain are equal to each other on a generating set of their domain, then these two homomorphisms must be identical). In other words, $\mathbf{v}_1 = \operatorname{id}$. This solves Exercise 2.9.10(d).

(e) Fix a positive integer n. We need to show that $\mathbf{v}_n : \Lambda \to \Lambda$ is a Hopf algebra homomorphism.

[683] *Proof.* Assume that $m \nmid r$. Then, $nm \nmid r$ (because otherwise, we would have $nm \mid r$, and therefore $m \mid nm \mid r$, which would contradict $m \nmid r$).

By the definition of $\mathbf{v}_m$, we have $\mathbf{v}_m(h_r) = \begin{cases} h_{r/m}, & \text{if } m \mid r; \\ 0, & \text{if } m \nmid r \end{cases} = 0$ (since $m \nmid r$). By the definition of $\mathbf{v}_{nm}$, we have $\mathbf{v}_{nm}(h_r) = \begin{cases} h_{r/(nm)}, & \text{if } nm \mid r; \\ 0, & \text{if } nm \nmid r \end{cases} = 0$ (since $nm \nmid r$). Now, $(\mathbf{v}_n \circ \mathbf{v}_m)(h_r) = \mathbf{v}_n \left(\underbrace{\mathbf{v}_m(h_r)}_{=0} \right) = \mathbf{v}_n(0) = 0$ (since $\mathbf{v}_n$ is $\mathbf{k}$-linear). Compared with $\mathbf{v}_{nm}(h_r) = 0$, this yields $(\mathbf{v}_n \circ \mathbf{v}_m)(h_r) = \mathbf{v}_{nm}(h_r)$. Hence, (13.87.10) holds is proven under the assumption that $m \nmid r$. Qed.

We know that $\mathbf{v}_n : \Lambda \to \Lambda$ is a $\mathbf{k}$-algebra homomorphism. Thus, $\mathbf{v}_n \otimes \mathbf{v}_n : \Lambda \otimes \Lambda \to \Lambda \otimes \Lambda$ is a $\mathbf{k}$-algebra homomorphism. Also, $\Delta : \Lambda \to \Lambda \otimes \Lambda$ is a $\mathbf{k}$-algebra homomorphism (due to the axioms of a bialgebra, since Λ is a bialgebra). Hence, $\Delta \circ \mathbf{v}_n$ and $(\mathbf{v}_n \otimes \mathbf{v}_n) \circ \Delta$ are $\mathbf{k}$-algebra homomorphisms.

For every $q \in \mathbb{N}$, we have

$$\Delta\left(h_q\right) = \sum_{i \in \{0,1,\ldots,q\}} h_i \otimes h_{q-i}. \tag{13.87.12}$$

[684]. Furthermore, for every $q \in \mathbb{N}$, we have

$$\mathbf{v}_n\left(h_{nq}\right) = h_q \tag{13.87.13}$$

[685].

Now, we shall prove that

$$\left(\Delta \circ \mathbf{v}_n\right)\left(h_r\right) = \left(\left(\mathbf{v}_n \otimes \mathbf{v}_n\right) \circ \Delta\right)\left(h_r\right) \qquad \text{for every positive integer } r. \tag{13.87.14}$$

Proof of (13.87.14): Let r be a positive integer. We need to prove (13.87.14).

Let us first notice that

$$\left(\mathbf{v}_n \otimes \mathbf{v}_n\right)\left(h_i \otimes h_{r-i}\right) = 0 \qquad \text{for every } i \in \{0,1,\ldots,r\} \text{ which does not satisfy } (n \mid i \text{ and } n \mid r-i) \tag{13.87.15}$$

[686]. In particular,

$$\left(\mathbf{v}_n \otimes \mathbf{v}_n\right)\left(h_i \otimes h_{r-i}\right) = 0 \qquad \text{for every } i \in \{0,1,\ldots,r\} \text{ satisfying } n \nmid i \tag{13.87.16}$$

[687].

---

[684] *Proof of (13.87.12):* Let $q \in \mathbb{N}$. Then, Proposition 2.3.6(iii) (applied to q instead of n) yields

$$\Delta\left(h_q\right) = \sum_{i+j=q} h_i \otimes h_j = \sum_{i \in \{0,1,\ldots,q\}} h_i \otimes h_{q-i}$$

(here, we have substituted $(i, q-i)$ for (i,j) in the sum), qed.

[685] *Proof of (13.87.13):* Let $q \in \mathbb{N}$. Applying (13.87.1) to nq instead of m, we obtain

$$\mathbf{v}_n\left(h_{nq}\right) = \begin{cases} h_{(nq)/n}, & \text{if } n \mid nq; \\ 0, & \text{if } n \nmid nq \end{cases} = h_{(nq)/n} \qquad (\text{since } n \mid nq)$$
$$= h_q,$$

qed.

[686] *Proof of (13.87.15):* Let $i \in \{0,1,\ldots,r\}$ be such that we don't have $(n \mid i$ and $n \mid r-i)$. If $n \nmid i$, then

$$\left(\mathbf{v}_n \otimes \mathbf{v}_n\right)\left(h_i \otimes h_{r-i}\right) = \underbrace{\mathbf{v}_n\left(h_i\right)}_{\substack{= \begin{cases} h_{i/n}, & \text{if } n \mid i; \\ 0, & \text{if } n \nmid i \end{cases} \\ \text{(by (13.87.1), applied to } m=i)}} \otimes \mathbf{v}_n\left(h_{r-i}\right) = \underbrace{\begin{cases} h_{i/n}, & \text{if } n \mid i; \\ 0, & \text{if } n \nmid i \end{cases}}_{\substack{=0 \\ (\text{since } n \nmid i)}} \otimes \mathbf{v}_n\left(h_{r-i}\right)$$
$$= 0 \otimes \mathbf{v}_n\left(h_{r-i}\right) = 0.$$

Hence, if $n \nmid i$, then (13.87.15) is proven. Thus, for the rest of this proof of (13.87.15), we can WLOG assume that we don't have $n \nmid i$. Assume this.

We have $n \mid i$ (since we don't have $n \nmid i$). Hence, we don't have $n \mid r-i$ (because otherwise, we would have $(n \mid i$ and $n \mid r-i)$, which would contradict the fact that we don't have $(n \mid i$ and $n \mid r-i)$). In other words, we have $n \nmid r-i$. Now,

$$\left(\mathbf{v}_n \otimes \mathbf{v}_n\right)\left(h_i \otimes h_{r-i}\right) = \mathbf{v}_n\left(h_i\right) \otimes \underbrace{\mathbf{v}_n\left(h_{r-i}\right)}_{\substack{= \begin{cases} h_{(r-i)/n}, & \text{if } n \mid r-i; \\ 0, & \text{if } n \nmid r-i \end{cases} \\ \text{(by (13.87.1), applied to } m=r-i)}} = \mathbf{v}_n\left(h_i\right) \otimes \underbrace{\begin{cases} h_{(r-i)/n}, & \text{if } n \mid r-i; \\ 0, & \text{if } n \nmid r-i \end{cases}}_{\substack{=0 \\ (\text{since } n \nmid r-i)}}$$
$$= \mathbf{v}_n\left(h_i\right) \otimes 0 = 0.$$

This proves (13.87.15).

[687] *Proof of (13.87.16):* Let $i \in \{0,1,\ldots,r\}$ be such that $n \nmid i$. Then, we don't have $n \mid i$. Hence, we don't have $(n \mid i$ and $n \mid r-i)$. Thus, (13.87.16) follows from (13.87.15), qed.

Let us now assume that $n \nmid r$. Then, (13.87.1) (applied to $m = r$) yields $\mathbf{v}_n \left(h_r\right) = \begin{cases} h_{r/n}, & \text{if } n \mid r; \\ 0, & \text{if } n \nmid r \end{cases} = 0$ (since $n \nmid r$), so that $\left(\Delta \circ \mathbf{v}_n\right)\left(h_r\right) = \Delta \left(\underbrace{\mathbf{v}_n \left(h_r\right)}_{=0} \right) = \Delta\left(0\right) = 0$. Also,

$$\left(\mathbf{v}_n \otimes \mathbf{v}_n\right)\left(h_i \otimes h_{r-i}\right) = 0 \qquad \text{for every } i \in \{0,1,\ldots,r\} \tag{13.87.17}$$

[688]. Now,

$$\begin{aligned} \left(\left(\mathbf{v}_n \otimes \mathbf{v}_n\right) \circ \Delta\right)\left(h_r\right) &= \left(\mathbf{v}_n \otimes \mathbf{v}_n\right) \left(\underbrace{\Delta\left(h_r\right)}_{\substack{=\sum_{i\in\{0,1,\ldots,r\}} h_i \otimes h_{r-i} \\ \text{(by (13.87.12), applied to } q=r\text{)}}} \right) = \left(\mathbf{v}_n \otimes \mathbf{v}_n\right) \left(\sum_{i\in\{0,1,\ldots,r\}} h_i \otimes h_{r-i} \right) \\ &= \sum_{i\in\{0,1,\ldots,r\}} \underbrace{\left(\mathbf{v}_n \otimes \mathbf{v}_n\right)\left(h_i \otimes h_{r-i}\right)}_{\substack{=0 \\ \text{(by (13.87.17))}}} \qquad \text{(since the map } \mathbf{v}_n \otimes \mathbf{v}_n \text{ is } \mathbf{k}\text{-linear)} \\ &= \sum_{i\in\{0,1,\ldots,r\}} 0 = 0. \end{aligned}$$

Compared to $\left(\Delta \circ \mathbf{v}_n\right)\left(h_r\right) = 0$, this yields $\left(\Delta \circ \mathbf{v}_n\right)\left(h_r\right) = \left(\left(\mathbf{v}_n \otimes \mathbf{v}_n\right) \circ \Delta\right)\left(h_r\right)$. In other words, (13.87.14) holds.

Now, let us forget that we assumed that $n \nmid r$. We thus have proven that (13.87.14) holds under the assumption that $n \nmid r$. Hence, for the rest of the proof of (13.87.14), we can WLOG assume that we don't have $n \nmid r$. Assume this.

We have $n \mid r$ (since we don't have $n \nmid r$). Hence, r/n is a positive integer. Denote this positive integer r/n by s. Then, $s = r/n$, so that $r = ns$. The equality (13.87.1) (applied to $m = r$) yields $\mathbf{v}_n\left(h_r\right) = \begin{cases} h_{r/n}, & \text{if } n \mid r; \\ 0, & \text{if } n \nmid r \end{cases} = h_{r/n}$ (since $n \mid r$), and we have

$$\left(\Delta \circ \mathbf{v}_n\right)\left(h_r\right) = \Delta \left(\underbrace{\mathbf{v}_n\left(h_r\right)}_{\substack{=h_{r/n}=h_s \\ \text{(since } r/n=s\text{)}}} \right) = \Delta\left(h_s\right) = \sum_{i\in\{0,1,\ldots,s\}} h_i \otimes h_{s-i} \tag{13.87.18}$$

(by (13.87.12), applied to $q = s$).

[688] *Proof of (13.87.17):* Let $i \in \{0,1,\ldots,r\}$.

Assume (for the sake of contradiction) that ($n \mid i$ and $n \mid r-i$). Then, $i \equiv 0 \mod n$ (since $n \mid i$) and $r \equiv i \mod n$ (since $n \mid r-i$). Hence, $r \equiv i \equiv 0 \mod n$, so that $n \mid r$. This contradicts $n \nmid r$. This contradiction shows that our assumption (that ($n \mid i$ and $n \mid r-i$)) was wrong. Hence, we do not have ($n \mid i$ and $n \mid r-i$). Thus, (13.87.17) follows from (13.87.15), qed.

On the other hand,

$$\begin{aligned}
\left(\left(\mathbf{v}_n\otimes\mathbf{v}_n\right)\circ\Delta\right)\left(h_r\right) &= \left(\mathbf{v}_n\otimes\mathbf{v}_n\right)\left(\underbrace{\Delta\left(h_r\right)}_{\substack{=\sum_{i\in\{0,1,\ldots,r\}} h_i\otimes h_{r-i}\\ \text{(by (13.87.12), applied to } q=r\text{)}}}\right) = \left(\mathbf{v}_n\otimes\mathbf{v}_n\right)\left(\sum_{i\in\{0,1,\ldots,r\}} h_i\otimes h_{r-i}\right)\\
&= \sum_{i\in\{0,1,\ldots,r\}} \left(\mathbf{v}_n\otimes\mathbf{v}_n\right)\left(h_i\otimes h_{r-i}\right) \qquad \text{(since the map } \mathbf{v}_n\otimes\mathbf{v}_n \text{ is } \mathbf{k}\text{-linear)}\\
&= \sum_{\substack{i\in\{0,1,\ldots,r\};\\ n\mid i}} \underbrace{\left(\mathbf{v}_n\otimes\mathbf{v}_n\right)\left(h_i\otimes h_{r-i}\right)}_{=\mathbf{v}_n\left(h_i\right)\otimes\mathbf{v}_n\left(h_{r-i}\right)} + \sum_{\substack{i\in\{0,1,\ldots,r\};\\ n\nmid i}} \underbrace{\left(\mathbf{v}_n\otimes\mathbf{v}_n\right)\left(h_i\otimes h_{r-i}\right)}_{\substack{=0\\ \text{(by (13.87.16))}}}\\
&= \sum_{\substack{i\in\{0,1,\ldots,r\};\\ n\mid i}} \mathbf{v}_n\left(h_i\right)\otimes\mathbf{v}_n\left(h_{r-i}\right) + \underbrace{\sum_{\substack{i\in\{0,1,\ldots,r\};\\ n\nmid i}} 0}_{=0} = \sum_{\substack{i\in\{0,1,\ldots,r\};\\ n\mid i}} \mathbf{v}_n\left(h_i\right)\otimes\mathbf{v}_n\left(h_{r-i}\right)\\
&= \underbrace{\sum_{i\in\{0,1,\ldots,r/n\}}}_{\substack{=\sum_{i\in\{0,1,\ldots,s\}}\\ \text{(since } r/n=s\text{)}}} \mathbf{v}_n\left(h_{ni}\right)\otimes\mathbf{v}_n\left(\underbrace{h_{r-ni}}_{\substack{=h_{n(s-i)}\\ \text{(since } r-ni=n(s-i)\\ \text{(because } r=ns=ni+n(s-i)\text{))}}}\right)\\
&\qquad \text{(here, we have substituted } ni \text{ for } i \text{ in the sum)}\\
&= \sum_{i\in\{0,1,\ldots,s\}} \underbrace{\mathbf{v}_n\left(h_{ni}\right)}_{\substack{=h_i\\ \text{(by (13.87.13), applied}\\ \text{to } q=i\text{)}}} \otimes \underbrace{\mathbf{v}_n\left(h_{n(s-i)}\right)}_{\substack{=h_{s-i}\\ \text{(by (13.87.13), applied}\\ \text{to } q=s-i\text{)}}} = \sum_{i\in\{0,1,\ldots,s\}} h_i\otimes h_{s-i}.
\end{aligned}$$

Compared with (13.87.18), this yields $\left(\Delta\circ\mathbf{v}_n\right)\left(h_r\right) = \left(\left(\mathbf{v}_n\otimes\mathbf{v}_n\right)\circ\Delta\right)\left(h_r\right)$. Thus, (13.87.14) is proven.

Now, recall that the family $\left(h_r\right)_{r\geq 1}$ generates the $\mathbf{k}$-algebra Λ (according to Proposition 2.4.1). In other words, $\left(h_r\right)_{r\geq 1}$ is a generating set of the $\mathbf{k}$-algebra Λ. The two $\mathbf{k}$-algebra homomorphisms $\Delta\circ\mathbf{v}_n$ and $\left(\mathbf{v}_n\otimes\mathbf{v}_n\right)\circ\Delta$ are equal to each other on this generating set (according to (13.87.14)), and therefore must be identical (because if two $\mathbf{k}$-algebra homomorphisms from the same domain are equal to each other on a generating set of their domain, then these two homomorphisms must be identical). In other words, $\Delta\circ\mathbf{v}_n = \left(\mathbf{v}_n\otimes\mathbf{v}_n\right)\circ\Delta$.

One can similarly check that $\epsilon\circ\mathbf{v}_n = \epsilon$. We can now conclude that the map $\mathbf{v}_n$ is a $\mathbf{k}$-coalgebra homomorphism (since it is $\mathbf{k}$-linear and satisfies $\Delta\circ\mathbf{v}_n = \left(\mathbf{v}_n\otimes\mathbf{v}_n\right)\circ\Delta$ and $\epsilon\circ\mathbf{v}_n = \epsilon$), therefore a $\mathbf{k}$-bialgebra homomorphism (since it also is a $\mathbf{k}$-algebra homomorphism), and therefore a Hopf algebra homomorphism (by Corollary 1.4.27). This solves Exercise 2.9.10(e).

(f) Recall first that $\left(h_\lambda\right)_{\lambda\in\mathrm{Par}}$ and $\left(m_\lambda\right)_{\lambda\in\mathrm{Par}}$ are mutually dual bases with respect to the Hall inner product on Λ (according to Corollary 2.5.17(a)). Thus,

$$\left(m_\lambda, h_\mu\right) = \delta_{\lambda,\mu} \qquad \text{for any two partitions } \lambda \text{ and } \mu. \tag{13.87.19}$$

Let us introduce a notation: For every weak composition λ and every positive integer s, let $\lambda\left\{s\right\}$ denote the weak composition $\left(s\lambda_1, s\lambda_2, s\lambda_3, \ldots\right)$, where λ is written in the form $\left(\lambda_1,\lambda_2,\lambda_3,\ldots\right)$. Notice that if λ is a partition and s is a positive integer, then $\lambda\left\{s\right\}$ is a partition as well, and satisfies

$$\ell\left(\lambda\left\{s\right\}\right) = \ell\left(\lambda\right). \tag{13.87.20}$$

We have

$$\mathbf{f}_n m_\lambda = m_{\lambda\{n\}} \qquad \text{for every partition } \lambda \tag{13.87.21}$$

[689].

We need to prove that the maps $\mathbf{f}_n : \Lambda \to \Lambda$ and $\mathbf{v}_n : \Lambda \to \Lambda$ are adjoint with respect to the Hall inner product on Λ. In other words, we need to prove that

$$(\mathbf{f}_n a, b) = (a, \mathbf{v}_n b) \qquad \text{for any } a \in \Lambda \text{ and } b \in \Lambda. \tag{13.87.28}$$

Proof of (13.87.28): Fix $a \in \Lambda$ and $b \in \Lambda$.

[689] *Proof of (13.87.21):* Let λ be a partition. Write λ in the form $\lambda = (\lambda_1, \lambda_2, \lambda_3, \ldots)$. By the definition of m_λ, we have $m_\lambda = \sum_{\alpha \in \mathfrak{S}_{(\infty)}\lambda} \mathbf{x}^\alpha$. But the definition of $\mathbf{f}_n$ yields

$$\mathbf{f}_n m_\lambda = \underbrace{m_\lambda}_{=\sum_{\alpha \in \mathfrak{S}_{(\infty)}\lambda} \mathbf{x}^\alpha} (x_1^n, x_2^n, x_3^n, \ldots) = \left(\sum_{\alpha \in \mathfrak{S}_{(\infty)}\lambda} \mathbf{x}^\alpha \right) (x_1^n, x_2^n, x_3^n, \ldots) = \sum_{\alpha \in \mathfrak{S}_{(\infty)}\lambda} \mathbf{x}^\alpha (x_1^n, x_2^n, x_3^n, \ldots). \tag{13.87.22}$$

But every weak composition α satisfies

$$\mathbf{x}^\alpha (x_1^n, x_2^n, x_3^n, \ldots) = \mathbf{x}^{\alpha\{n\}}. \tag{13.87.23}$$

(*Proof of (13.87.23):* Let α be a weak composition. Write α as $(\alpha_1, \alpha_2, \alpha_3, \ldots)$. Then, $\alpha\{n\} = (n\alpha_1, n\alpha_2, n\alpha_3, \ldots)$ (by the definition of $\alpha\{n\}$). Thus, $\mathbf{x}^{\alpha\{n\}} = x_1^{n\alpha_1} x_2^{n\alpha_2} x_3^{n\alpha_3} \cdots$ (by the definition of $\mathbf{x}^{\alpha\{n\}}$). On the other hand, $\mathbf{x}^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} x_3^{\alpha_3} \cdots$ (by the definition of $\mathbf{x}^\alpha$), so that

$$\underbrace{\mathbf{x}^\alpha}_{=x_1^{\alpha_1} x_2^{\alpha_2} x_3^{\alpha_3} \cdots} (x_1^n, x_2^n, x_3^n, \ldots) = \left(x_1^{\alpha_1} x_2^{\alpha_2} x_3^{\alpha_3} \cdots \right) (x_1^n, x_2^n, x_3^n, \ldots) = (x_1^n)^{\alpha_1} (x_2^n)^{\alpha_2} (x_3^n)^{\alpha_3} \cdots = x_1^{n\alpha_1} x_2^{n\alpha_2} x_3^{n\alpha_3} \cdots = \mathbf{x}^{\alpha\{n\}}.$$

Thus, (13.87.23) is proven.)

Now, (13.87.22) becomes

$$\mathbf{f}_n m_\lambda = \sum_{\alpha \in \mathfrak{S}_{(\infty)}\lambda} \underbrace{\mathbf{x}^\alpha (x_1^n, x_2^n, x_3^n, \ldots)}_{\substack{=\mathbf{x}^{\alpha\{n\}} \\ \text{(by (13.87.23))}}} = \sum_{\alpha \in \mathfrak{S}_{(\infty)}\lambda} \mathbf{x}^{\alpha\{n\}}. \tag{13.87.24}$$

On the other hand, for every weak composition α and every permutation $\sigma \in \mathfrak{S}_{(\infty)}$, we have

$$\sigma(\alpha\{n\}) = (\sigma\alpha)\{n\}. \tag{13.87.25}$$

(*Proof of (13.87.25):* Let α be a weak composition. Let $\sigma \in \mathfrak{S}_{(\infty)}$. Write α as $(\alpha_1, \alpha_2, \alpha_3, \ldots)$. Then, $\alpha\{n\} = (n\alpha_1, n\alpha_2, n\alpha_3, \ldots)$ (by the definition of $\alpha\{n\}$), so that $\sigma(\alpha\{n\}) = \left(n\alpha_{\sigma^{-1}(1)}, n\alpha_{\sigma^{-1}(2)}, n\alpha_{\sigma^{-1}(3)}, \ldots \right)$. But $\sigma\alpha = \left(\alpha_{\sigma^{-1}(1)}, \alpha_{\sigma^{-1}(2)}, \alpha_{\sigma^{-1}(3)}, \ldots \right)$ (since $\alpha = (\alpha_1, \alpha_2, \alpha_3, \ldots)$) and thus $(\sigma\alpha)\{n\} = \left(n\alpha_{\sigma^{-1}(1)}, n\alpha_{\sigma^{-1}(2)}, n\alpha_{\sigma^{-1}(3)}, \ldots \right)$ (by the definition of $(\sigma\alpha)\{n\}$). Compared with $\sigma(\alpha\{n\}) = \left(n\alpha_{\sigma^{-1}(1)}, n\alpha_{\sigma^{-1}(2)}, n\alpha_{\sigma^{-1}(3)}, \ldots \right)$, this yields $\sigma(\alpha\{n\}) = (\sigma\alpha)\{n\}$. This proves (13.87.25).)

Now,

$$\alpha\{n\} \in \mathfrak{S}_{(\infty)}(\lambda\{n\}) \qquad \text{for every } \alpha \in \mathfrak{S}_{(\infty)}\lambda. \tag{13.87.26}$$

(*Proof of (13.87.26):* Let $\alpha \in \mathfrak{S}_{(\infty)}\lambda$. Then, there exists a $\sigma \in \mathfrak{S}_{(\infty)}$ such that $\alpha = \sigma\lambda$. Consider this σ. Then, (13.87.25) (applied to λ instead of α) yields that $\sigma(\lambda\{n\}) = (\sigma\lambda)\{n\}$. Compared with $\underbrace{\alpha}_{=\sigma\lambda}\{n\} = (\sigma\lambda)\{n\}$, this yields $\alpha\{n\} = \underbrace{\sigma}_{\in \mathfrak{S}_{(\infty)}} (\lambda\{n\}) \in \mathfrak{S}_{(\infty)}(\lambda\{n\})$. Thus, (13.87.26) is proven.)

Also,

$$\text{every element of } \mathfrak{S}_{(\infty)}(\lambda\{n\}) \text{ has the form } \alpha\{n\} \text{ for some } \alpha \in \mathfrak{S}_{(\infty)}\lambda. \tag{13.87.27}$$

(*Proof of (13.87.27):* Let β be an element of $\mathfrak{S}_{(\infty)}(\lambda\{n\})$. Then, there exists a $\sigma \in \mathfrak{S}_{(\infty)}$ such that $\beta = \sigma(\lambda\{n\})$. Consider this σ. Now, $\underbrace{\sigma}_{\in \mathfrak{S}_{(\infty)}} \lambda \in \mathfrak{S}_{(\infty)}\lambda$. Also, (13.87.25) (applied to λ instead of α) yields that $\sigma(\lambda\{n\}) = (\sigma\lambda)\{n\}$. Hence, $\beta = \sigma(\lambda\{n\}) = (\sigma\lambda)\{n\}$. Therefore, β has the form $\alpha\{n\}$ for some $\alpha \in \mathfrak{S}_{(\infty)}\lambda$ (namely, for $\alpha = \sigma\lambda$). Let us now forget that we fixed β. We thus have shown that every element β of $\mathfrak{S}_{(\infty)}(\lambda\{n\})$ has the form $\alpha\{n\}$ for some $\alpha \in \mathfrak{S}_{(\infty)}\lambda$. In other words, we have proven (13.87.27).)

Now, the map

$$\mathfrak{S}_{(\infty)}\lambda \to \mathfrak{S}_{(\infty)}(\lambda\{n\}),$$
$$\alpha \mapsto \alpha\{n\}$$

is well-defined (according to (13.87.26)). This map is injective (because any weak composition α can be uniquely reconstructed from $\alpha\{n\}$) and surjective (due to (13.87.27)); therefore, this map is bijective.

Now, (13.87.24) becomes

$$\mathbf{f}_n m_\lambda = \sum_{\alpha \in \mathfrak{S}_{(\infty)}\lambda} \mathbf{x}^{\alpha\{n\}} = \sum_{\alpha \in \mathfrak{S}_{(\infty)}(\lambda\{n\})} \mathbf{x}^\alpha$$

Recall that $(m_\lambda)_{\lambda\in\operatorname{Par}}$ is a basis of the $\mathbf{k}$-module Λ. Hence, in proving (13.87.28), we can WLOG assume that a is an element of this basis $(m_\lambda)_{\lambda\in\operatorname{Par}}$ (because the equality (13.87.28) is $\mathbf{k}$-linear in a). Assume this. Thus, a is an element of the basis $(m_\lambda)_{\lambda\in\operatorname{Par}}$. In other words, there exists a $\mu \in \operatorname{Par}$ such that $a = m_\mu$. Consider this μ. We have $\mathbf{f}_n \underbrace{a}_{=m_\mu} = \mathbf{f}_n m_\mu = m_{\mu\{n\}}$ (by (13.87.21), applied to $\lambda = \mu$).

Recall that $(h_\lambda)_{\lambda\in\operatorname{Par}}$ is a basis of the $\mathbf{k}$-module Λ. Hence, in proving (13.87.28), we can WLOG assume that b is an element of this basis $(h_\lambda)_{\lambda\in\operatorname{Par}}$ (because the equality (13.87.28) is $\mathbf{k}$-linear in b). Assume this. Thus, b is an element of the basis $(h_\lambda)_{\lambda\in\operatorname{Par}}$. In other words, there exists a $\nu \in \operatorname{Par}$ such that $b = h_\nu$. Consider this ν. We have

$$\left(\underbrace{\mathbf{f}_n a}_{=m_{\mu\{n\}}}, \underbrace{b}_{=h_\nu}\right) = \left(m_{\mu\{n\}}, h_\nu\right) = \delta_{\mu\{n\},\nu} \tag{13.87.29}$$

(by (13.87.19), applied to $\mu\{n\}$ and ν instead of λ and μ).

Let us write the partition ν in the form $(\nu_1, \nu_2, \nu_3, \ldots)$. Then, $\nu = \left(\nu_1, \nu_2, \ldots, \nu_{\ell(\nu)}\right)$, so that $h_\nu = h_{\nu_1} h_{\nu_2} \cdots h_{\nu_{\ell(\nu)}}$ (by the definition of h_ν).

Let us first assume that

$$(\text{there exists an } i \in \{1,2,3,\ldots\} \text{ such that } n \nmid \nu_i)\,. \tag{13.87.30}$$

Then, $\mathbf{v}_n(h_\nu) = 0$ [690]. Also, $\mu\{n\} \neq \nu$ [691], so that $\delta_{\mu\{n\},\nu} = 0$. Thus, (13.87.29) becomes $(\mathbf{f}_n a, b) = \delta_{\mu\{n\},\nu} = 0$. Compared with $\left(a, \mathbf{v}_n \underbrace{b}_{=h_\nu}\right) = \left(a, \underbrace{\mathbf{v}_n(h_\nu)}_{=0}\right) = (a, 0) = 0$ (since the Hall inner product is $\mathbf{k}$-bilinear), this yields $(\mathbf{f}_n a, b) = (a, \mathbf{v}_n b)$. Thus, (13.87.28) holds.

Now, let us forget that we assumed that (13.87.30) holds. We thus have proven (13.87.28) under the assumption that (13.87.30) holds. Hence, for the rest of our proof of (13.87.28), we can WLOG assume that (13.87.30) does not hold. Assume this.

(here, we substituted α for $\alpha\{n\}$ in the sum, since the map

$$\mathfrak{S}_{(\infty)}\lambda \to \mathfrak{S}_{(\infty)}(\lambda\{n\}),$$
$$\alpha \mapsto \alpha\{n\}$$

is bijective). Compared with

$$m_{\lambda\{n\}} = \sum_{\alpha\in\mathfrak{S}_{(\infty)}(\lambda\{n\})} \mathbf{x}^\alpha \qquad \left(\text{by the definition of } m_{\lambda\{n\}}\right),$$

this yields $\mathbf{f}_n m_\lambda = m_{\lambda\{n\}}$. Thus, (13.87.21) is proven.

[690] *Proof.* There exists an $i \in \{1,2,3,\ldots\}$ such that $n \nmid \nu_i$. Consider this i. We have $n \nmid \nu_i$, thus $\nu_i \neq 0$, and therefore ν_i is a positive integer. Hence, $i \leq \ell(\nu)$, so that $i \in \{1,2,\ldots,\ell(\nu)\}$. Also, the definition of $\mathbf{v}_n$ yields $\mathbf{v}_n(h_{\nu_i}) = \begin{cases} h_{\nu_i/n}, & \text{if } n \mid \nu_i; \\ 0, & \text{if } n \nmid \nu_i \end{cases} = 0$ (since $n \nmid \nu_i$). But

$$h_\nu = h_{\nu_1} h_{\nu_2} \cdots h_{\nu_{\ell(\nu)}} = \left(h_{\nu_1} h_{\nu_2} \cdots h_{\nu_{i-1}}\right) h_{\nu_i} \left(h_{\nu_{i+1}} h_{\nu_{i+2}} \cdots h_{\nu_{\ell(\nu)}}\right) \qquad (\text{since } i \in \{1,2,\ldots,\ell(\nu)\})\,.$$

Applying the map $\mathbf{v}_n$ to both sides of this equality, we obtain

$$\begin{aligned}\mathbf{v}_n(h_\nu) &= \mathbf{v}_n\left(\left(h_{\nu_1} h_{\nu_2} \cdots h_{\nu_{i-1}}\right) h_{\nu_i} \left(h_{\nu_{i+1}} h_{\nu_{i+2}} \cdots h_{\nu_{\ell(\nu)}}\right)\right) \\ &= \mathbf{v}_n\left(h_{\nu_1} h_{\nu_2} \cdots h_{\nu_{i-1}}\right) \cdot \underbrace{\mathbf{v}_n(h_{\nu_i})}_{=0} \cdot \mathbf{v}_n\left(h_{\nu_{i+1}} h_{\nu_{i+2}} \cdots h_{\nu_{\ell(\nu)}}\right) \qquad (\text{since } \mathbf{v}_n \text{ is a } \mathbf{k}\text{-algebra homomorphism}) \\ &= 0,\end{aligned}$$

qed.

[691] *Proof.* Assume the contrary. Thus, $\mu\{n\} = \nu$. Let us write the partition μ in the form $(\mu_1, \mu_2, \mu_3, \ldots)$. Then, $\mu\{n\} = (n\mu_1, n\mu_2, n\mu_3, \ldots)$ (by the definition of $\mu\{n\}$). Hence, $(\nu_1, \nu_2, \nu_3, \ldots) = \nu = \mu\{n\} = (n\mu_1, n\mu_2, n\mu_3, \ldots)$. Thus, every positive integer j satisfies $\nu_j = n\mu_j$.

But (13.87.30) yields that there exists an $i \in \{1,2,3,\ldots\}$ such that $n \nmid \nu_i$. Consider this i. Now, recall that every positive integer j satisfies $\nu_j = n\mu_j$. Applied to $j = i$, this yields $\nu_i = n\mu_i$, so that $n \mid \nu_i$. This contradicts $n \nmid \nu_i$. This contradiction proves that our assumption was wrong, qed.

We have assumed that (13.87.30) does not hold. In other words, there exists no $i \in \{1, 2, 3, \ldots\}$ such that $n \nmid \nu_i$. In other words,

$$\text{every } i \in \{1, 2, 3, \ldots\} \text{ satisfies } n \mid \nu_i. \tag{13.87.31}$$

Thus, ν_i / n is a nonnegative integer for every $i \in \{1, 2, 3, \ldots\}$. These nonnegative integers satisfy $\nu_1/n \geq \nu_2/n \geq \nu_3/n \geq \cdots$ (since $\nu_1 \geq \nu_2 \geq \nu_3 \geq \cdots$ (since $(\nu_1, \nu_2, \nu_3, \ldots) = \nu$ is a partition)) and $(\nu_i/n = 0 \text{ for all sufficiently high } i)$ (since $(\nu_i = 0 \text{ for all sufficiently high } i)$ (since $(\nu_1, \nu_2, \nu_3, \ldots) = \nu$ is a partition)). Hence, $(\nu_1/n, \nu_2/n, \nu_3/n, \ldots)$ is a partition. Let us denote this partition by κ. We have $\kappa\{n\} = \nu$ [692] and thus $\ell(\kappa) = \ell(\nu)$ [693]. We have $\mathbf{v}_n(h_\nu) = h_\kappa$ [694] and thus

$$\left(\underbrace{a}_{=m_\mu}, \mathbf{v}_n \underbrace{b}_{=h_\nu}\right) = \left(m_\mu, \underbrace{\mathbf{v}_n(h_\nu)}_{=h_\kappa}\right) = (m_\mu, h_\kappa) = \delta_{\mu,\kappa} \tag{13.87.32}$$

(by (13.87.19), applied to μ and κ instead of λ and μ).

Now, let us write the partition μ in the form $(\mu_1, \mu_2, \mu_3, \ldots)$. Then, $\mu\{n\} = (n\mu_1, n\mu_2, n\mu_3, \ldots)$ (by the definition of $\mu\{n\}$).

Now, we have the following equivalence of assertions:

$$\begin{aligned}
&(\mu\{n\} = \nu) \\
\Longleftrightarrow\ & ((n\mu_1, n\mu_2, n\mu_3, \ldots) = (\nu_1, \nu_2, \nu_3, \ldots)) \\
&\qquad (\text{since } \mu\{n\} = (n\mu_1, n\mu_2, n\mu_3, \ldots) \text{ and } \nu = (\nu_1, \nu_2, \nu_3, \ldots)) \\
\Longleftrightarrow\ & (\text{every } i \in \{1, 2, 3, \ldots\} \text{ satisfies } n\mu_i = \nu_i) \\
\Longleftrightarrow\ & (\text{every } i \in \{1, 2, 3, \ldots\} \text{ satisfies } \mu_i = \nu_i/n) \\
\Longleftrightarrow\ & ((\mu_1, \mu_2, \mu_3, \ldots) = (\nu_1/n, \nu_2/n, \nu_3/n, \ldots)) \\
\Longleftrightarrow\ & (\mu = \kappa) \qquad (\text{since } (\mu_1, \mu_2, \mu_3, \ldots) = \mu \text{ and } (\nu_1/n, \nu_2/n, \nu_3/n, \ldots) = \kappa).
\end{aligned} \tag{13.87.33}$$

[692] *Proof.* We have $\kappa = (\nu_1/n, \nu_2/n, \nu_3/n, \ldots)$. Hence, the definition of $\kappa\{n\}$ yields $\kappa\{n\} = (n(\nu_1/n), n(\nu_2/n), n(\nu_3/n), \ldots) = (\nu_1, \nu_2, \nu_3, \ldots) = \nu$, qed.

[693] *Proof.* Applying (13.87.20) to κ and n instead of λ and s, we obtain $\ell(\kappa\{n\}) = \ell(\kappa)$. Since $\kappa\{n\} = \nu$, this rewrites as $\ell(\nu) = \ell(\kappa)$, qed.

[694] *Proof.* We have $\kappa = (\nu_1/n, \nu_2/n, \nu_3/n, \ldots)$ and therefore $\kappa = \left(\nu_1/n, \nu_2/n, \ldots, \nu_{\ell(\kappa)}/n\right)$. Since $\ell(\kappa) = \ell(\nu)$, this rewrites as $\kappa = \left(\nu_1/n, \nu_2/n, \ldots, \nu_{\ell(\nu)}/n\right)$. Hence, the definition of h_κ yields $h_\kappa = h_{\nu_1/n} h_{\nu_2/n} \cdots h_{\nu_{\ell(\nu)}/n} = \prod_{i=1}^{\ell(\nu)} h_{\nu_i/n}$.

On the other hand, $\nu = \left(\nu_1, \nu_2, \ldots, \nu_{\ell(\nu)}\right)$, so that the definition of h_ν yields $h_\nu = h_{\nu_1} h_{\nu_2} \cdots h_{\nu_{\ell(\nu)}} = \prod_{i=1}^{\ell(\nu)} h_{\nu_i}$. Applying the map $\mathbf{v}_n$ to both sides of this equality, we obtain

$$\begin{aligned}
\mathbf{v}_n(h_\nu) = \mathbf{v}_n\left(\prod_{i=1}^{\ell(\nu)} h_{\nu_i}\right) &= \prod_{i=1}^{\ell(\nu)} \underbrace{\mathbf{v}_n(h_{\nu_i})}_{\substack{=\begin{cases} h_{\nu_i/n}, & \text{if } n \mid \nu_i; \\ 0, & \text{if } n \nmid \nu_i \end{cases} \\ \text{(by (13.87.1), applied to } m = \nu_i)}} \qquad (\text{since } \mathbf{v}_n \text{ is a } \mathbf{k}\text{-algebra homomorphism}) \\
&= \prod_{i=1}^{\ell(\nu)} \underbrace{\begin{cases} h_{\nu_i/n}, & \text{if } n \mid \nu_i; \\ 0, & \text{if } n \nmid \nu_i \end{cases}}_{\substack{=h_{\nu_i/n} \\ \text{(since } n \mid \nu_i \text{ (by (13.87.31)))}}} = \prod_{i=1}^{\ell(\nu)} h_{\nu_i/n} = h_\kappa
\end{aligned}$$

(since $h_\kappa = \prod_{i=1}^{\ell(\nu)} h_{\nu_i/n}$), qed.

But (13.87.29) becomes

$$\begin{aligned}
(\mathbf{f}_n a, b) &= \delta_{\mu\{n\},\nu} = \begin{cases} 1, & \text{if } \mu\{n\} = \nu; \\ 0, & \text{otherwise} \end{cases} \qquad \left(\text{by the definition of } \delta_{\mu\{n\},\nu}\right) \\
&= \begin{cases} 1, & \text{if } \mu = \kappa; \\ 0, & \text{otherwise} \end{cases} \qquad \left(\text{since } \mu\{n\} = \nu \text{ is equivalent to } \mu = \kappa \text{ (by (13.87.33))}\right) \\
&= \delta_{\mu,\kappa} \qquad \left(\text{since } \delta_{\mu,\kappa} = \begin{cases} 1, & \text{if } \mu = \kappa; \\ 0, & \text{otherwise} \end{cases} \text{ (by the definition of } \delta_{\mu,\kappa})\right) \\
&= (a, \mathbf{v}_n b) \qquad (\text{by (13.87.32)}) .
\end{aligned}$$

Thus, (13.87.28) is proven. As we know, this completes the solution of Exercise 2.9.10(f).

(g) Our solution to Exercise 2.9.10(g) shall proceed in three steps:

- *Step 1:* proving that Exercise 2.9.10(g) holds if $\mathbf{k} = \mathbb{Q}$.
- *Step 2:* proving that Exercise 2.9.10(g) holds if $\mathbf{k} = \mathbb{Z}$.
- *Step 3:* proving that Exercise 2.9.10(g) holds in the general case.

Let us now get to the details of these three steps:

Step 1: We shall prove that Exercise 2.9.10(g) holds if $\mathbf{k} = \mathbb{Q}$.

Indeed, assume that $\mathbf{k} = \mathbb{Q}$. We know that $\mathbf{f}_n$ is a $\mathbf{k}$-algebra homomorphism (due to Exercise 2.9.9(a)), and that $\mathbf{v}_n$ is a $\mathbf{k}$-algebra homomorphism. Hence, $\mathbf{v}_n \circ \mathbf{f}_n$ is a $\mathbf{k}$-algebra homomorphism. Also, $\underbrace{\operatorname{id}_\Lambda \star \operatorname{id}_\Lambda \star \cdots \star \operatorname{id}_\Lambda}_{n \text{ times}}$ is a $\mathbf{k}$-algebra homomorphism (due to Exercise 1.5.11(b), applied to $H = \Lambda$, $A = \Lambda$, $k = n$ and $f_i = \operatorname{id}_\Lambda$). In other words, $\operatorname{id}_\Lambda^{\star n}$ is a $\mathbf{k}$-algebra homomorphism.

Let r be a positive integer. We shall now show that $(\mathbf{v}_n \circ \mathbf{f}_n)(p_r) = \operatorname{id}_\Lambda^{\star n}(p_r)$.

Indeed, it is easy to see that

$$\operatorname{id}_\Lambda^{\star a}(p_r) = a p_r \qquad \text{for every } a \in \mathbb{N} \tag{13.87.34}$$

[695]. Applied to $a = n$, this yields $\operatorname{id}_\Lambda^{\star n}(p_r) = np_r$. But the definition of $\mathbf{f}_n(p_r)$ yields

$$\begin{aligned}
\mathbf{f}_n(p_r) &= p_r\left(x_1^n, x_2^n, x_3^n, \ldots\right) \\
&= \left(x_1^n\right)^r + \left(x_2^n\right)^r + \left(x_3^n\right)^r + \cdots \qquad \left(\text{since } p_r = x_1^r + x_2^r + x_3^r + \cdots\right) \\
&= x_1^{rn} + x_2^{rn} + x_3^{rn} + \ldots = p_{rn}.
\end{aligned} \tag{13.87.35}$$

Hence,

$$\begin{aligned}
\left(\mathbf{v}_n \circ \mathbf{f}_n\right)(p_r) &= \mathbf{v}_n\left(\underbrace{\mathbf{f}_n(p_r)}_{=p_{rn}}\right) = \mathbf{v}_n(p_{rn}) \\
&= \begin{cases} np_{rn/n}, & \text{if } n \mid rn; \\ 0, & \text{if } n \nmid rn \end{cases} \qquad (\text{by Exercise 2.9.10(a), applied to } m = rn) \\
&= np_{rn/n} \qquad (\text{since } n \mid rn) \\
&= np_r = \operatorname{id}_\Lambda^{\star n}(p_r)
\end{aligned}$$

(since $\operatorname{id}_\Lambda^{\star n}(p_r) = np_r$).

Now, let us forget that we fixed r. We thus have proven that

$$\left(\mathbf{v}_n \circ \mathbf{f}_n\right)(p_r) = \operatorname{id}_\Lambda^{\star n}(p_r) \qquad \text{for every positive integer } r. \tag{13.87.36}$$

We have assumed that $\mathbf{k} = \mathbb{Q}$. Thus, $\mathbb{Q}$ is a subring of $\mathbf{k}$. Hence, the elements $p_1, p_2, p_3, \ldots$ of Λ generate the $\mathbf{k}$-algebra Λ (due to Proposition 2.4.1). But $\mathbf{v}_n \circ \mathbf{f}_n$ and $\operatorname{id}_\Lambda^{\star n}$ are two $\mathbf{k}$-algebra homomorphisms from Λ. These two homomorphisms $\mathbf{v}_n \circ \mathbf{f}_n$ and $\operatorname{id}_\Lambda^{\star n}$ are equal to each other on each of the elements $p_1, p_2, p_3, \ldots$ (due to (13.87.36)), and therefore are identical (because if two $\mathbf{k}$-algebra homomorphisms with one and the same domain are equal to each other on a generating set of the domain, then these homomorphisms must be identical). In other words, $\mathbf{v}_n \circ \mathbf{f}_n = \operatorname{id}_\Lambda^{\star n}$. Thus, Exercise 2.9.10(g) is solved under the assumption that $\mathbf{k} = \mathbb{Q}$. Our Step 1 is complete.

Step 2: We shall prove that Exercise 2.9.10(g) holds if $\mathbf{k} = \mathbb{Z}$.

Indeed, let us first consider the general case, without any assumptions on $\mathbf{k}$.

[695] *Proof of (13.87.34):* We shall prove (13.87.34) by induction over a:

Induction base: We have $\epsilon(p_r) = 0$ (since r is positive). Now, $\underbrace{\operatorname{id}_\Lambda^{\star 0}}_{=u\circ\epsilon}(p_r) = (u \circ \epsilon)(p_r) = u\left(\underbrace{\epsilon(p_r)}_{=0}\right) = u(0) = 0$, so that $\operatorname{id}_\Lambda^{\star 0}(p_r) = 0 = 0p_r$. In other words, (13.87.34) holds for $a = 0$. This completes the induction base.

Induction step: Let A be a positive integer. Assume that (13.87.34) holds for $a = A - 1$. We now need to show that (13.87.34) holds for $a = A$.

We know that (13.87.34) holds for $a = A-1$. In other words, $\operatorname{id}_\Lambda^{\star(A-1)}(p_r) = (A-1)p_r$. Now, recall that $\Delta p_r = 1 \otimes p_r + p_r \otimes 1$ (by Proposition 2.3.6(i), applied to r instead of n). On the other hand, $\operatorname{id}_\Lambda^{\star(A-1)}$ is a $\mathbf{k}$-algebra homomorphism (in fact, this is shown in the same way as we have shown that $\operatorname{id}_\Lambda^{\star n}$ is a $\mathbf{k}$-algebra homomorphism), so that we have $\operatorname{id}_\Lambda^{\star(A-1)}(1) = 1$. Now, $\operatorname{id}_\Lambda^{\star A} = \operatorname{id}_\Lambda^{\star(A-1)} \star \operatorname{id}_\Lambda = m \circ \left(\operatorname{id}_\Lambda^{\star(A-1)} \otimes \operatorname{id}_\Lambda\right) \circ \Delta$ (by the definition of convolution). Applying both sides of this equality to p_r, we obtain

$$\begin{aligned}
\operatorname{id}_\Lambda^{\star A}(p_r) &= \left(m \circ \left(\operatorname{id}_\Lambda^{\star(A-1)} \otimes \operatorname{id}_\Lambda\right) \circ \Delta\right)(p_r) = m\left(\left(\operatorname{id}_\Lambda^{\star(A-1)} \otimes \operatorname{id}_\Lambda\right) \underbrace{(\Delta p_r)}_{=1\otimes p_r + p_r \otimes 1}\right) \\
&= m\left(\underbrace{\left(\operatorname{id}_\Lambda^{\star(A-1)} \otimes \operatorname{id}_\Lambda\right)(1 \otimes p_r + p_r \otimes 1)}_{=\operatorname{id}_\Lambda^{\star(A-1)}(1)\otimes \operatorname{id}_\Lambda(p_r) + \operatorname{id}_\Lambda^{\star(A-1)}(p_r)\otimes \operatorname{id}_\Lambda(1)}\right) \\
&= m\left(\operatorname{id}_\Lambda^{\star(A-1)}(1) \otimes \operatorname{id}_\Lambda(p_r) + \operatorname{id}_\Lambda^{\star(A-1)}(p_r) \otimes \operatorname{id}_\Lambda(1)\right) \\
&= \underbrace{\operatorname{id}_\Lambda^{\star(A-1)}(1)}_{=1} \cdot \underbrace{\operatorname{id}_\Lambda(p_r)}_{=p_r} + \underbrace{\operatorname{id}_\Lambda^{\star(A-1)}(p_r)}_{=(A-1)p_r} \cdot \underbrace{\operatorname{id}_\Lambda(1)}_{=1} = p_r + (A-1)p_r = Ap_r.
\end{aligned}$$

Thus, (13.87.34) holds for $a = A$. This completes the induction step, so that (13.87.34) is proven.

We shall now formalize the intuitive concept that the constructions relevant to Exercise 2.9.10(g) (the Hopf algebra Λ, its elements m_λ, p_λ, e_λ etc., and the maps $\mathbf{v}_n$, $\mathbf{f}_n$ and $\operatorname{id}_\Lambda^{\star n}$) are "functorial" with respect to the base ring $\mathbf{k}$. This is chiefly a matter of introducing notations:

- We denote the $\mathbf{k}$-algebra $\Lambda = \Lambda_\mathbf{k}$ by $\Lambda^{[\mathbf{k}]}$. (This is just a minor change of notation that serves to make it more similar to the notations $\mathbf{f}_n^{[\mathbf{k}]}$, $\mathbf{v}_n^{[\mathbf{k}]}$ and $m_\lambda^{[\mathbf{k}]}$ further below.)
- If $\mathbf{m}$ and $\mathbf{n}$ are two commutative rings, and $\varphi : \mathbf{m} \to \mathbf{n}$ is a ring homomorphism, then φ canonically induces a ring homomorphism $\Lambda^{[\mathbf{m}]} \to \Lambda^{[\mathbf{n}]}$. [696] We denote this latter homomorphism by $\Lambda^{[\varphi]}$.

  Notice that while $\Lambda^{[\mathbf{k}]}$ denotes a ring, $\Lambda^{[\varphi]}$ denotes a map. This might look confusing, but it makes sense from the viewpoint of category theory: There is a functor from the category of commutative rings to itself, which sends every commutative ring $\mathbf{k}$ to $\Lambda^{[\mathbf{k}]}$, and every morphism $\varphi : \mathbf{m} \to \mathbf{n}$ of commutative rings to $\Lambda^{[\varphi]}$. Unsurprisingly, this functor is denoted by Λ.
- A ring homomorphism $\varphi : \mathbf{m} \to \mathbf{n}$ also induces a ring homomorphism $\Lambda^{[\mathbf{m}]} \otimes_\mathbf{m} \Lambda^{[\mathbf{m}]} \to \Lambda^{[\mathbf{n}]} \otimes_\mathbf{n} \Lambda^{[\mathbf{n}]}$ [697]. This makes $\Lambda \otimes \Lambda$ into a functor of $\mathbf{k}$.
- For every partition λ, we shall denote the monomial symmetric function $m_\lambda \in \Lambda$ by $m_\lambda^{[\mathbf{k}]}$. This notation makes the dependency of m_λ on the base ring $\mathbf{k}$ more explicit, and thus allows us to talk about these elements defined over several different base rings at the same time (without the danger of confusing them). Of course, the element m_λ does not "really" depend on $\mathbf{k}$, in the sense that it is defined in the same way for every $\mathbf{k}$. This entails that for any two commutative rings $\mathbf{m}$ and $\mathbf{n}$ and any ring homomorphism $\varphi : \mathbf{m} \to \mathbf{n}$, we have
$$m_\lambda^{[\mathbf{n}]} = \Lambda^{[\varphi]} \left(m_\lambda^{[\mathbf{m}]} \right) \tag{13.87.37}$$
  for every partition λ.
- We shall denote the homomorphism $\mathbf{f}_n : \Lambda \to \Lambda$ by $\mathbf{f}_n^{[\mathbf{k}]}$. This notation makes the dependency of $\mathbf{f}_n$ on the base ring $\mathbf{k}$ more explicit.

  The definition of $\mathbf{f}_n$ was functorial in $\mathbf{k}$. Thus, for any two commutative rings $\mathbf{m}$ and $\mathbf{n}$ and any ring homomorphism $\varphi : \mathbf{m} \to \mathbf{n}$, we have
$$\mathbf{f}_n^{[\mathbf{n}]} \circ \Lambda^{[\varphi]} = \Lambda^{[\varphi]} \circ \mathbf{f}_n^{[\mathbf{m}]} \tag{13.87.38}$$
  as maps from $\Lambda^{[\mathbf{m}]}$ to $\Lambda^{[\mathbf{n}]}$.
- Similarly, we shall denote the homomorphism $\mathbf{v}_n : \Lambda \to \Lambda$ by $\mathbf{v}_n^{[\mathbf{k}]}$. This homomorphism is again defined in a way that is functorial in $\mathbf{k}$. Thus, for any two commutative rings $\mathbf{m}$ and $\mathbf{n}$ and any ring homomorphism $\varphi : \mathbf{m} \to \mathbf{n}$, we have
$$\mathbf{v}_n^{[\mathbf{n}]} \circ \Lambda^{[\varphi]} = \Lambda^{[\varphi]} \circ \mathbf{v}_n^{[\mathbf{m}]} \tag{13.87.39}$$
  as maps from $\Lambda^{[\mathbf{m}]}$ to $\Lambda^{[\mathbf{n}]}$.
- Notice that the Hopf algebra structure on Λ is functorial in $\mathbf{k}$. [698] As a consequence, for any two commutative rings $\mathbf{m}$ and $\mathbf{n}$ and any ring homomorphism $\varphi : \mathbf{m} \to \mathbf{n}$, we have
$$\operatorname{id}_{\Lambda_\mathbf{n}}^{\star n} \circ \Lambda^{[\varphi]} = \Lambda^{[\varphi]} \circ \operatorname{id}_{\Lambda_\mathbf{m}}^{\star n} \tag{13.87.40}$$
  as maps from $\Lambda^{[\mathbf{m}]}$ to $\Lambda^{[\mathbf{n}]}$.
- For every commutative ring $\mathbf{k}$, there exists a canonical (and unique) ring homomorphism $\mathbb{Z} \to \mathbf{k}$. Denote this homomorphism by $\rho_\mathbf{k}$.

[696] In fact, this ring homomorphism $\Lambda^{[\mathbf{m}]} \to \Lambda^{[\mathbf{n}]}$ can be defined by taking the ring homomorphism $\varphi[[\mathbf{x}]] : \mathbf{m}[[\mathbf{x}]] \to \mathbf{n}[[\mathbf{x}]]$ canonically induced by φ and restricting it to the subring $\Lambda^{[\mathbf{m}]}$ of $\mathbf{m}[[\mathbf{x}]]$.

[697] There are several ways to define this homomorphism. One of them is to define it as the $\mathbf{m}$-module homomorphism $\Lambda^{[\mathbf{m}]} \otimes_\mathbf{m} \Lambda^{[\mathbf{m}]} \to \Lambda^{[\mathbf{n}]} \otimes_\mathbf{n} \Lambda^{[\mathbf{n}]}$ canonically induced by the $\mathbf{m}$-bilinear map
$$\Lambda^{[\mathbf{m}]} \times \Lambda^{[\mathbf{m}]} \to \Lambda^{[\mathbf{n}]} \otimes_\mathbf{n} \Lambda^{[\mathbf{n}]},$$
$$(a, b) \mapsto \left(\Lambda^{[\varphi]}(a) \right) \otimes_\mathbf{n} \left(\Lambda^{[\varphi]}(b) \right).$$
(Here, $\Lambda^{[\mathbf{n}]} \otimes_\mathbf{n} \Lambda^{[\mathbf{n}]}$ is an $\mathbf{m}$-module because the ring homomorphism $\varphi : \mathbf{m} \to \mathbf{n}$ makes $\mathbf{n}$ into an $\mathbf{m}$-algebra.)

[698] To make sense of this statement, we need to recall that $\Lambda \otimes \Lambda$ has been made into a functor of $\mathbf{k}$.

In Step 1, we have shown that Exercise 2.9.10(g) holds if $\mathbf{k} = \mathbb{Q}$. In other words, we have $\mathbf{v}_n \circ \mathbf{f}_n = \mathrm{id}_\Lambda^{\star n}$ if $\mathbf{k} = \mathbb{Q}$. In other words,

$$\mathbf{v}_n^{[\mathbb{Q}]} \circ \mathbf{f}_n^{[\mathbb{Q}]} = \mathrm{id}_{\Lambda_\mathbb{Q}}^{\star n}. \tag{13.87.41}$$

Recall that $\left(m_\lambda^{[\mathbf{k}]}\right)_{\lambda \in \mathrm{Par}} = (m_\lambda)_{\lambda \in \mathrm{Par}}$ is a basis of the $\mathbf{k}$-module $\Lambda_\mathbf{k}$ for every commutative ring $\mathbf{k}$. Applying this to $\mathbf{k} = \mathbb{Z}$, we conclude that $\left(m_\lambda^{[\mathbb{Z}]}\right)_{\lambda \in \mathrm{Par}}$ is a basis of the $\mathbb{Z}$-module $\Lambda_\mathbb{Z}$.

The ring homomorphism $\rho_\mathbb{Q} : \mathbb{Z} \to \mathbb{Q}$ is just the canonical inclusion of $\mathbb{Z}$ into $\mathbb{Q}$, and thus is injective. Hence, the map $\Lambda^{[\rho_\mathbb{Q}]}$ is injective (because $\Lambda^{[\varphi]}$ is injective for every injective ring homomorphism φ).

Now, let us recall that our goal (in Step 2) is to show that Exercise 2.9.10(g) holds if $\mathbf{k} = \mathbb{Z}$. In other words, our goal is to show that

$$\mathbf{v}_n^{[\mathbb{Z}]} \circ \mathbf{f}_n^{[\mathbb{Z}]} = \mathrm{id}_{\Lambda_\mathbb{Z}}^{\star n}.$$

This is an equality between $\mathbb{Z}$-module homomorphisms (indeed, it is clear that both $\mathbf{v}_n^{[\mathbb{Z}]} \circ \mathbf{f}_n^{[\mathbb{Z}]}$ and $\mathrm{id}_{\Lambda_\mathbb{Z}}^{\star n}$ are $\mathbb{Z}$-module homomorphisms). Hence, in order to prove it, it is sufficient to verify it on the basis $\left(m_\lambda^{[\mathbb{Z}]}\right)_{\lambda \in \mathrm{Par}}$ of the $\mathbb{Z}$-module $\Lambda_\mathbb{Z}$. In other words, it is sufficient to prove that

$$\left(\mathbf{v}_n^{[\mathbb{Z}]} \circ \mathbf{f}_n^{[\mathbb{Z}]}\right)\left(m_\lambda^{[\mathbb{Z}]}\right) = \mathrm{id}_{\Lambda_\mathbb{Z}}^{\star n}\left(m_\lambda^{[\mathbb{Z}]}\right) \qquad \text{for every } \lambda \in \mathrm{Par}. \tag{13.87.42}$$

So let us prove (13.87.42) now:

Proof of (13.87.42): Let $\lambda \in \mathrm{Par}$. We have

$$\begin{aligned}
&\Lambda^{[\rho_\mathbb{Q}]}\left(\left(\mathbf{v}_n^{[\mathbb{Z}]} \circ \mathbf{f}_n^{[\mathbb{Z}]}\right)\left(m_\lambda^{[\mathbb{Z}]}\right)\right) \\
&= \left(\underbrace{\Lambda^{[\rho_\mathbb{Q}]} \circ \mathbf{v}_n^{[\mathbb{Z}]}}_{\substack{= \mathbf{v}_n^{[\mathbb{Q}]} \circ \Lambda^{[\rho_\mathbb{Q}]} \\ \text{(because of } \mathbf{v}_n^{[\mathbb{Q}]} \circ \Lambda^{[\rho_\mathbb{Q}]} = \Lambda^{[\rho_\mathbb{Q}]} \circ \mathbf{v}_n^{[\mathbb{Z}]} \\ \text{(by (13.87.39), applied} \\ \text{to } \mathbf{m} = \mathbb{Z},\ \mathbf{n} = \mathbb{Q} \text{ and } \varphi = \rho_\mathbb{Q}))}} \circ \mathbf{f}_n^{[\mathbb{Z}]}\right)\left(m_\lambda^{[\mathbb{Z}]}\right) = \left(\mathbf{v}_n^{[\mathbb{Q}]} \circ \underbrace{\Lambda^{[\rho_\mathbb{Q}]} \circ \mathbf{f}_n^{[\mathbb{Z}]}}_{\substack{= \mathbf{f}_n^{[\mathbb{Q}]} \circ \Lambda^{[\rho_\mathbb{Q}]} \\ \text{(because of } \mathbf{f}_n^{[\mathbb{Q}]} \circ \Lambda^{[\rho_\mathbb{Q}]} = \Lambda^{[\rho_\mathbb{Q}]} \circ \mathbf{f}_n^{[\mathbb{Z}]} \\ \text{(by (13.87.38), applied} \\ \text{to } \mathbf{m} = \mathbb{Z},\ \mathbf{n} = \mathbb{Q} \text{ and } \varphi = \rho_\mathbb{Q}))}}\right)\left(m_\lambda^{[\mathbb{Z}]}\right) \\
&= \left(\underbrace{\mathbf{v}_n^{[\mathbb{Q}]} \circ \mathbf{f}_n^{[\mathbb{Q}]}}_{\substack{= \mathrm{id}_{\Lambda_\mathbb{Q}}^{\star n} \\ \text{(by (13.87.41))}}} \circ \Lambda^{[\rho_\mathbb{Q}]}\right)\left(m_\lambda^{[\mathbb{Z}]}\right) = \underbrace{\left(\mathrm{id}_{\Lambda_\mathbb{Q}}^{\star n} \circ \Lambda^{[\rho_\mathbb{Q}]}\right)}_{\substack{= \Lambda^{[\rho_\mathbb{Q}]} \circ \mathrm{id}_{\Lambda_\mathbb{Z}}^{\star n} \\ \text{(by (13.87.40), applied} \\ \text{to } \mathbf{m} = \mathbb{Z},\ \mathbf{n} = \mathbb{Q} \text{ and } \varphi = \rho_\mathbb{Q})}} \left(m_\lambda^{[\mathbb{Z}]}\right) \\
&= \left(\Lambda^{[\rho_\mathbb{Q}]} \circ \mathrm{id}_{\Lambda_\mathbb{Z}}^{\star n}\right)\left(m_\lambda^{[\mathbb{Z}]}\right) = \Lambda^{[\rho_\mathbb{Q}]}\left(\mathrm{id}_{\Lambda_\mathbb{Z}}^{\star n}\left(m_\lambda^{[\mathbb{Z}]}\right)\right).
\end{aligned}$$

Since the map $\Lambda^{[\rho_\mathbb{Q}]}$ is injective, this yields $\left(\mathbf{v}_n^{[\mathbb{Z}]} \circ \mathbf{f}_n^{[\mathbb{Z}]}\right)\left(m_\lambda^{[\mathbb{Z}]}\right) = \mathrm{id}_{\Lambda_\mathbb{Z}}^{\star n}\left(m_\lambda^{[\mathbb{Z}]}\right)$. Thus, (13.87.42) is proven.

As we said, proving (13.87.42) is sufficient to showing that $\mathbf{v}_n^{[\mathbb{Z}]} \circ \mathbf{f}_n^{[\mathbb{Z}]} = \mathrm{id}_{\Lambda_\mathbb{Z}}^{\star n}$. Hence, $\mathbf{v}_n^{[\mathbb{Z}]} \circ \mathbf{f}_n^{[\mathbb{Z}]} = \mathrm{id}_{\Lambda_\mathbb{Z}}^{\star n}$ is shown (since (13.87.42) is proven). In other words, Exercise 2.9.10(g) holds if $\mathbf{k} = \mathbb{Z}$. This completes Step 2.

Step 3: We shall now prove that Exercise 2.9.10(g) holds in the general case.

Let us use all the notations that we introduced in Step 2.

In Step 2, we have shown that Exercise 2.9.10(g) holds if $\mathbf{k} = \mathbb{Z}$. In other words, we have $\mathbf{v}_n \circ \mathbf{f}_n = \mathrm{id}_\Lambda^{\star n}$ if $\mathbf{k} = \mathbb{Z}$. In other words,

$$\mathbf{v}_n^{[\mathbb{Z}]} \circ \mathbf{f}_n^{[\mathbb{Z}]} = \mathrm{id}_{\Lambda_\mathbb{Z}}^{\star n}. \tag{13.87.43}$$

Recall that $(m_\lambda)_{\lambda \in \mathrm{Par}}$ is a basis of the $\mathbf{k}$-module $\Lambda_\mathbf{k}$.

Now, let us recall that our goal (in Step 3) is to show that Exercise 2.9.10(g) holds. In other words, our goal is to show that

$$\mathbf{v}_n \circ \mathbf{f}_n = \mathrm{id}^{\star n}.$$

This is an equality between $\mathbf{k}$-module homomorphisms (indeed, it is clear that both $\mathbf{v}_n \circ \mathbf{f}_n$ and $\operatorname{id}^{\star n}$ are $\mathbf{k}$-module homomorphisms). Hence, in order to prove it, it is sufficient to verify it on the basis $(m_\lambda)_{\lambda \in \operatorname{Par}}$ of the $\mathbf{k}$-module $\Lambda_{\mathbf{k}}$. In other words, it is sufficient to prove that

$$(\mathbf{v}_n \circ \mathbf{f}_n)(m_\lambda) = \operatorname{id}^{\star n}(m_\lambda) \qquad \text{for every } \lambda \in \operatorname{Par}. \tag{13.87.44}$$

So let us prove (13.87.44) now:

Proof of (13.87.44): Let $\lambda \in \operatorname{Par}$. Then,

$$m_\lambda = m_\lambda^{[\mathbf{k}]} = \Lambda^{[\rho_{\mathbf{k}}]}\left(m_\lambda^{[\mathbb{Z}]}\right) \qquad (\text{by (13.87.37) (applied to } \mathbf{m} = \mathbb{Z},\ \mathbf{n} = \mathbf{k} \text{ and } \varphi = \rho_{\mathbf{k}})). \tag{13.87.45}$$

Hence,

$$\left(\underbrace{\mathbf{v}_n}_{=\mathbf{v}_n^{[\mathbf{k}]}} \circ \underbrace{\mathbf{f}_n}_{=\mathbf{f}_n^{[\mathbf{k}]}}\right)\left(\underbrace{m_\lambda}_{=\Lambda^{[\rho_{\mathbf{k}}]}\left(m_\lambda^{[\mathbb{Z}]}\right)}\right) = \left(\mathbf{v}_n^{[\mathbf{k}]} \circ \mathbf{f}_n^{[\mathbf{k}]}\right)\left(\Lambda^{[\rho_{\mathbf{k}}]}\left(m_\lambda^{[\mathbb{Z}]}\right)\right) = \left(\mathbf{v}_n^{[\mathbf{k}]} \circ \underbrace{\mathbf{f}_n^{[\mathbf{k}]} \circ \Lambda^{[\rho_{\mathbf{k}}]}}_{\substack{=\Lambda^{[\rho_{\mathbf{k}}]} \circ \mathbf{f}_n^{[\mathbb{Z}]} \\ \text{(by (13.87.38), applied} \\ \text{to } \mathbf{m}=\mathbb{Z},\ \mathbf{n}=\mathbf{k} \text{ and } \varphi=\rho_{\mathbf{k}})}}\right)\left(m_\lambda^{[\mathbb{Z}]}\right)$$

$$= \left(\underbrace{\mathbf{v}_n^{[\mathbf{k}]} \circ \Lambda^{[\rho_{\mathbf{k}}]}}_{\substack{=\Lambda^{[\rho_{\mathbf{k}}]} \circ \mathbf{v}_n^{[\mathbb{Z}]} \\ \text{(by (13.87.39), applied} \\ \text{to } \mathbf{m}=\mathbb{Z},\ \mathbf{n}=\mathbf{k} \text{ and } \varphi=\rho_{\mathbf{k}})}} \circ \mathbf{f}_n^{[\mathbb{Z}]}\right)\left(m_\lambda^{[\mathbb{Z}]}\right)$$

$$= \left(\Lambda^{[\rho_{\mathbf{k}}]} \circ \underbrace{\mathbf{v}_n^{[\mathbb{Z}]} \circ \mathbf{f}_n^{[\mathbb{Z}]}}_{\substack{=\operatorname{id}_{\Lambda_{\mathbb{Z}}}^{\star n} \\ \text{(by (13.87.43))}}}\right)\left(m_\lambda^{[\mathbb{Z}]}\right) = \underbrace{\left(\Lambda^{[\rho_{\mathbf{k}}]} \circ \operatorname{id}_{\Lambda_{\mathbb{Z}}}^{\star n}\right)}_{\substack{=\operatorname{id}_{\Lambda_{\mathbf{k}}}^{\star n} \circ \Lambda^{[\rho_{\mathbf{k}}]} \\ \text{(because of } \operatorname{id}_{\Lambda_{\mathbf{k}}}^{\star n} \circ \Lambda^{[\rho_{\mathbf{k}}]} = \Lambda^{[\rho_{\mathbf{k}}]} \circ \operatorname{id}_{\Lambda_{\mathbb{Z}}}^{\star n} \\ \text{(by (13.87.40), applied} \\ \text{to } \mathbf{m}=\mathbb{Z},\ \mathbf{n}=\mathbf{k} \text{ and } \varphi=\rho_{\mathbf{k}}))}}\left(m_\lambda^{[\mathbb{Z}]}\right)$$

$$= \left(\operatorname{id}_{\Lambda_{\mathbf{k}}}^{\star n} \circ \Lambda^{[\rho_{\mathbf{k}}]}\right)\left(m_\lambda^{[\mathbb{Z}]}\right) = \underbrace{\operatorname{id}_{\Lambda_{\mathbf{k}}}^{\star n}}_{=\operatorname{id}^{\star n}}\left(\underbrace{\Lambda^{[\rho_{\mathbf{k}}]}\left(m_\lambda^{[\mathbb{Z}]}\right)}_{\substack{=m_\lambda \\ \text{(by (13.87.45))}}}\right) = \operatorname{id}^{\star n}(m_\lambda).$$

Thus, (13.87.44) is proven.

As we said, proving (13.87.44) is sufficient to showing that $\mathbf{v}_n \circ \mathbf{f}_n = \operatorname{id}^{\star n}$. Hence, $\mathbf{v}_n \circ \mathbf{f}_n = \operatorname{id}^{\star n}$ is shown (since (13.87.44) is proven). In other words, Exercise 2.9.10(g) holds in the general case. This completes Step 3 and, with it, the solution of Exercise 2.9.10(g).

(h) The structure of our solution of Exercise 2.9.10(h) is similar to that of our solution of Exercise 2.9.10(g) above. It proceeds in three steps:

- *Step 1:* proving that Exercise 2.9.10(h) holds if $\mathbf{k} = \mathbb{Q}$.
- *Step 2:* proving that Exercise 2.9.10(h) holds if $\mathbf{k} = \mathbb{Z}$.
- *Step 3:* proving that Exercise 2.9.10(h) holds in the general case.

The details of Steps 2 and 3 are very similar to the details of the corresponding steps in the solution of Exercise 2.9.10(g), and so we will not dwell on these details. However, we need to give the details of Step 1:

Step 1: We shall prove that Exercise 2.9.10(h) holds if $\mathbf{k} = \mathbb{Q}$.

Indeed, assume that $\mathbf{k} = \mathbb{Q}$. We know that $\mathbf{f}_n$ is a $\mathbf{k}$-algebra homomorphism (due to Exercise 2.9.9(a)), and that $\mathbf{v}_m$ is a $\mathbf{k}$-algebra homomorphism. Hence, $\mathbf{f}_n \circ \mathbf{v}_m$ and $\mathbf{v}_m \circ \mathbf{f}_n$ are $\mathbf{k}$-algebra homomorphisms.

Let r be a positive integer. We shall show that $(\mathbf{f}_n \circ \mathbf{v}_m)(p_r) = (\mathbf{v}_m \circ \mathbf{f}_n)(p_r)$.

Indeed, we first notice that

$$\mathbf{f}_n(p_r) = p_{rn}. \tag{13.87.46}$$

(This can be proven just as in (13.87.35).)

Now, let us first assume that $m \nmid r$. Then, Exercise 2.9.10(a) (applied to m and r instead of n and m) yields

$$\mathbf{v}_m(p_r) = \begin{cases} mp_{r/m}, & \text{if } m \mid r; \\ 0, & \text{if } m \nmid r \end{cases} = 0 \qquad (\text{since } m \nmid r),$$

so that $(\mathbf{f}_n \circ \mathbf{v}_m)(p_r) = \mathbf{f}_n\left(\underbrace{\mathbf{v}_m(p_r)}_{=0}\right) = \mathbf{f}_n(0) = 0$ (since the map $\mathbf{f}_n$ is $\mathbf{k}$-linear). On the other hand, if we had $m \mid rn$, then we would have $m \mid r$ (since m is coprime to n), which would contradict $m \nmid r$. Hence, we cannot have $m \mid rn$. Thus, we have $m \nmid rn$. Now,

$$\begin{aligned}(\mathbf{v}_m \circ \mathbf{f}_n)(p_r) &= \mathbf{v}_m\left(\underbrace{\mathbf{f}_n(p_r)}_{\substack{=p_{rn} \\ (\text{by } (13.87.46))}}\right) = \mathbf{v}_m(p_{rn}) = \begin{cases} mp_{rn/m}, & \text{if } m \mid rn; \\ 0, & \text{if } m \nmid rn \end{cases} \\ &\qquad (\text{by Exercise 2.9.10(a), applied to } m \text{ and } rn \text{ instead of } n \text{ and } m) \\ &= 0 \qquad (\text{since } m \nmid rn).\end{aligned}$$

Compared with $(\mathbf{f}_n \circ \mathbf{v}_m)(p_r) = 0$, this yields $(\mathbf{f}_n \circ \mathbf{v}_m)(p_r) = (\mathbf{v}_m \circ \mathbf{f}_n)(p_r)$.

Now, let us forget our assumption that $m \nmid r$. Hence, we have proven that $(\mathbf{f}_n \circ \mathbf{v}_m)(p_r) = (\mathbf{v}_m \circ \mathbf{f}_n)(p_r)$ under the assumption that $m \nmid r$. As a consequence, for the rest of the proof of $(\mathbf{f}_n \circ \mathbf{v}_m)(p_r) = (\mathbf{v}_m \circ \mathbf{f}_n)(p_r)$, we can WLOG assume that we don't have $m \nmid r$. Assume this.

We have $m \mid r$ (since we don't have $m \nmid r$). Now, Exercise 2.9.10(a) (applied to m and r instead of n and m) yields

$$\mathbf{v}_m(p_r) = \begin{cases} mp_{r/m}, & \text{if } m \mid r; \\ 0, & \text{if } m \nmid r \end{cases} = mp_{r/m} \qquad (\text{since } m \mid r),$$

so that

$$\begin{aligned}(\mathbf{f}_n \circ \mathbf{v}_m)(p_r) &= \mathbf{f}_n\left(\underbrace{\mathbf{v}_m(pr)}_{=mp_{r/m}}\right) = \mathbf{f}_n\left(mp_{r/m}\right) = m \underbrace{\mathbf{f}_n\left(p_{r/m}\right)}_{\substack{=p_{(r/m)n} \\ (\text{by } (13.87.46), \text{ applied to} \\ r/m \text{ instead of } r)}} \qquad (\text{since the map } \mathbf{f}_n \text{ is } \mathbf{k}\text{-linear}) \\ &= mp_{(r/m)n} = mp_{rn/m} \qquad (\text{since } (r/m)\,n = rn/m).\end{aligned}$$

Compared with

$$\begin{aligned}(\mathbf{v}_m \circ \mathbf{f}_n)(p_r) &= \mathbf{v}_m\left(\underbrace{\mathbf{f}_n(p_r)}_{\substack{=p_{rn} \\ (\text{by } (13.87.46))}}\right) = \mathbf{v}_m(p_{rn}) = \begin{cases} mp_{rn/m}, & \text{if } m \mid rn; \\ 0, & \text{if } m \nmid rn \end{cases} \\ &\qquad (\text{by Exercise 2.9.10(a), applied to } m \text{ and } rn \text{ instead of } n \text{ and } m) \\ &= mp_{rn/m} \qquad (\text{since } m \mid rn \text{ (since } m \mid r \mid rn)),\end{aligned}$$

this yields $(\mathbf{f}_n \circ \mathbf{v}_m)(p_r) = (\mathbf{v}_m \circ \mathbf{f}_n)(p_r)$. Hence, $(\mathbf{f}_n \circ \mathbf{v}_m)(p_r) = (\mathbf{v}_m \circ \mathbf{f}_n)(p_r)$ is proven.

Now, let us forget that we fixed r. We thus have proven that

$$(\mathbf{f}_n \circ \mathbf{v}_m)(p_r) = (\mathbf{v}_m \circ \mathbf{f}_n)(p_r) \qquad \text{for every positive integer } r. \tag{13.87.47}$$

We have assumed that $\mathbf{k} = \mathbb{Q}$. Thus, $\mathbb{Q}$ is a subring of $\mathbf{k}$. Hence, the elements $p_1, p_2, p_3, \ldots$ of Λ generate the $\mathbf{k}$-algebra Λ (due to Proposition 2.4.1). But $\mathbf{f}_n \circ \mathbf{v}_m$ and $\mathbf{v}_m \circ \mathbf{f}_n$ are two $\mathbf{k}$-algebra homomorphisms from Λ. These two homomorphisms $\mathbf{f}_n \circ \mathbf{v}_m$ and $\mathbf{v}_m \circ \mathbf{f}_n$ are equal to each other on each of the elements $p_1, p_2, p_3, \ldots$ (due to (13.87.47)), and therefore are identical (because if two $\mathbf{k}$-algebra homomorphisms with one and the same domain are equal to each other on a generating set of the domain, then these homomorphisms must be identical). In other words, $\mathbf{f}_n \circ \mathbf{v}_m = \mathbf{v}_m \circ \mathbf{f}_n$. Thus, Exercise 2.9.10(h) is solved under the assumption that $\mathbf{k} = \mathbb{Q}$. Our Step 1 is complete.

As already mentioned, Steps 2 and Steps 3 are very similar to the corresponding steps in our above solution of Exercise 2.9.10(g). Thus, we forego showing these steps. Exercise 2.9.10(h) is thus solved.

(i) Our solution of Exercise 2.9.10(i) will be somewhat similar to that of Exercise 2.9.10(g), but simpler. We will only need two steps:

- *Step 1:* proving that Exercise 2.9.10(i) holds if $\mathbf{k} = \mathbb{Z}$.
- *Step 2:* proving that Exercise 2.9.10(i) holds in the general case.

Let us go through the details of each step:

Step 1: We shall prove that Exercise 2.9.10(i) holds if $\mathbf{k} = \mathbb{Z}$.

Indeed, assume that $\mathbf{k} = \mathbb{Z}$. Every positive integer r satisfies

$$p_r = \sum_{d \mid r} d w_d^{r/d} \tag{13.87.48}$$

(according to Exercise 2.9.3(e), applied to r instead of n).

Now, for every positive integer m, we define an element $\widetilde{w}_m$ of Λ by

$$\widetilde{w}_m = \begin{cases} w_{m/n}, & \text{if } n \mid m; \\ 0, & \text{if } n \nmid m \end{cases} . \tag{13.87.49}$$

It is now easy to see that every positive integer r satisfies

$$\mathbf{v}_n \left(p_r\right) = \sum_{d \mid r} d \widetilde{w}_d^{r/d}. \tag{13.87.50}$$

[699]

We shall now prove that every positive integer m satisfies

$$\mathbf{v}_n \left(w_m\right) = \widetilde{w}_m. \tag{13.87.52}$$

[699] *Proof of (13.87.50):* Let r be a positive integer. Let us first assume that $n \nmid r$. Then, every positive divisor d of r satisfies

$$\begin{aligned} \widetilde{w}_d &= \begin{cases} w_{d/n}, & \text{if } n \mid d; \\ 0, & \text{if } n \nmid d \end{cases} \qquad \left(\text{by the definition of } \widetilde{w}_d\right) \\ &= 0 \qquad \left(\begin{array}{c} \text{since } n \nmid d \text{ (because otherwise, we would have } n \mid d \\ \text{and thus } n \mid d \mid r \text{, which would contradict } n \nmid r) \end{array}\right). \end{aligned} \tag{13.87.51}$$

But Exercise 2.9.10(a) (applied to r instead of m) yields

$$\mathbf{v}_n \left(p_r\right) = \begin{cases} n p_{r/n}, & \text{if } n \mid r; \\ 0, & \text{if } n \nmid r \end{cases} = 0 \qquad \left(\text{since } n \nmid r\right).$$

Compared with

$$\sum_{d \mid r} d \widetilde{w}_d^{r/d} = \sum_{d \mid r} d \left(\underbrace{\widetilde{w}_d}_{\substack{=0 \\ \text{(by (13.87.51))}}} \right)^{r/d} = \sum_{d \mid r} d \underbrace{0^{r/d}}_{\substack{=0 \\ \text{(since } r/d>0)}} = \sum_{d \mid r} d 0 = 0,$$

this yields $\mathbf{v}_n \left(p_r\right) = \sum_{d \mid r} d \widetilde{w}_d^{r/d}$. Thus, (13.87.50) holds.

Let us now forget that we assumed that $n \nmid r$. We thus have proven (13.87.50) under the assumption that $n \nmid r$. Hence, for the rest of the proof of (13.87.50), we can WLOG assume that we don't have $n \nmid r$. Assume this.

We have $n \mid r$ (since we don't have $n \nmid r$). Hence, r/n is a positive integer. Exercise 2.9.10(a) (applied to r instead of m) yields

$$\begin{aligned} \mathbf{v}_n \left(p_r\right) &= \begin{cases} n p_{r/n}, & \text{if } n \mid r; \\ 0, & \text{if } n \nmid r \end{cases} = n \underbrace{p_{r/n}}_{\substack{=\sum_{d \mid r/n} d w_d^{(r/n)/d} \\ \text{(by (13.87.48), applied to} \\ r/n \text{ instead of } r)}} \qquad \left(\text{since } n \mid r\right) \\ &= n \sum_{d \mid r/n} d w_d^{(r/n)/d} = \sum_{d \mid r/n} n d \underbrace{w_d^{(r/n)/d}}_{\substack{=w_d^{r/(nd)} \\ \text{(since } (r/n)/d = r/(nd))}} = \sum_{d \mid r/n} n d w_d^{r/(nd)}. \end{aligned}$$

Proof of (13.87.52): We proceed by strong induction over m:

Induction step: Fix a positive integer M. Assume that (13.87.52) holds for every positive integer $m < M$. We now must show that (13.87.52) holds for $m = M$.

The $\mathbf{k}$-module Λ is free, and thus torsionfree. Hence,

$$\text{every element } a \text{ of } \Lambda \text{ satisfying } Ma = 0 \text{ satisfies } a = 0 \tag{13.87.53}$$

(since $\mathbf{k} = \mathbb{Z}$).

We have assumed that (13.87.52) holds for every positive integer $m < M$. In other words,

$$\mathbf{v}_n \left(w_m \right) = \widetilde{w}_m \qquad \text{for every positive integer } m < M. \tag{13.87.54}$$

Compared with

$$\begin{aligned}
\sum_{d|r} d\widetilde{w}_d^{r/d} &= \sum_{d|r} d \left(\underbrace{\widetilde{w}_d}_{\substack{= \begin{cases} w_{d/n}, & \text{if } n \mid d; \\ 0, & \text{if } n \nmid d \end{cases} \\ \text{(by the definition of } \widetilde{w}_d)}} \right)^{r/d} = \sum_{d|r} d \left(\begin{cases} w_{d/n}, & \text{if } n \mid d; \\ 0, & \text{if } n \nmid d \end{cases} \right)^{r/d} \\
&= \sum_{\substack{d|r; \\ n|d}} d \left(\underbrace{\begin{cases} w_{d/n}, & \text{if } n \mid d; \\ 0, & \text{if } n \nmid d \end{cases}}_{\substack{= w_{d/n} \\ \text{(since } n|d)}} \right)^{r/d} + \sum_{\substack{d|r; \\ n\nmid d}} d \left(\underbrace{\begin{cases} w_{d/n}, & \text{if } n \mid d; \\ 0, & \text{if } n \nmid d \end{cases}}_{\substack{=0 \\ \text{(since } n\nmid d)}} \right)^{r/d} \\
&= \sum_{\substack{d|r; \\ n|d}} dw_{d/n}^{r/d} + \sum_{\substack{d|r; \\ n\nmid d}} d \underbrace{0^{r/d}}_{\substack{=0 \\ \text{(since } r/d>0)}} = \sum_{\substack{d|r; \\ n|d}} dw_{d/n}^{r/d} + \underbrace{\sum_{\substack{d|r; \\ n\nmid d}} d0}_{=0} = \sum_{\substack{d|r; \\ n|d}} dw_{d/n}^{r/d} \\
&= \sum_{d|r/n} nd \underbrace{w_{nd/n}^{r/(nd)}}_{\substack{= w_d^{r/(nd)} \\ \text{(since } nd/n=d)}} \qquad \text{(here, we have substituted } nd \text{ for } d \text{ in the sum)} \\
&= \sum_{d|r/n} ndw_d^{r/(nd)},
\end{aligned}$$

this yields $\mathbf{v}_n \left(p_r \right) = \sum_{d|r} d\widetilde{w}_d^{r/d}$. Thus, (13.87.50) is proven.

Applying (13.87.48) to $r = M$, we obtain $p_M = \sum_{d|M} d w_d^{M/d}$. Applying the map $\mathbf{v}_n$ to both sides of this equality, we obtain

$$\begin{aligned}
\mathbf{v}_n(p_M) &= \mathbf{v}_n\left(\sum_{d|M} d w_d^{M/d}\right) = \sum_{d|M} d\left(\mathbf{v}_n(w_d)\right)^{M/d} \qquad (\text{since } \mathbf{v}_n \text{ is a } \mathbf{k}\text{-algebra homomorphism}) \\
&= M \underbrace{\left(\mathbf{v}_n(w_M)\right)^{M/M}}_{=(\mathbf{v}_n(w_M))^1=\mathbf{v}_n(w_M)} + \underbrace{\sum_{\substack{d|M;\\ d\neq M}}}_{=\sum_{\substack{d|M;\\ d<M}}} d\left(\mathbf{v}_n(w_d)\right)^{M/d} \\
&\qquad (\text{here, we have split off the addend for } d = M \text{ from the sum}) \\
&= M\mathbf{v}_n(w_M) + \sum_{\substack{d|M;\\ d<M}} d\left(\underbrace{\mathbf{v}_n(w_d)}_{\substack{=\widetilde{w}_d \\ (\text{by (13.87.54), applied} \\ \text{to } m=d)}}\right)^{M/d} = M\mathbf{v}_n(w_M) + \sum_{\substack{d|M;\\ d<M}} d\widetilde{w}_d^{M/d}.
\end{aligned}$$

Compared with

$$\begin{aligned}
\mathbf{v}_n(p_M) &= \sum_{d|M} d\widetilde{w}_d^{M/d} \qquad (\text{by (13.87.50), applied to } r = M) \\
&= M \underbrace{\widetilde{w}_M^{M/M}}_{=\widetilde{w}_M^1=\widetilde{w}_M} + \underbrace{\sum_{\substack{d|M;\\ d\neq M}}}_{=\sum_{\substack{d|M;\\ d<M}}} d\widetilde{w}_d^{M/d} \qquad (\text{here, we have split off the addend for } d = M \text{ from the sum}) \\
&= M\widetilde{w}_M + \sum_{\substack{d|M;\\ d<M}} d\widetilde{w}_d^{M/d},
\end{aligned}$$

this yields

$$M\mathbf{v}_n(w_M) + \sum_{\substack{d|M;\\ d<M}} d\widetilde{w}_d^{M/d} = M\widetilde{w}_M + \sum_{\substack{d|M;\\ d<M}} d\widetilde{w}_d^{M/d}.$$

Subtracting $\sum_{\substack{d|M;\\ d<M}} d\widetilde{w}_d^{M/d}$ from both sides of this equality, we obtain $M\mathbf{v}_n(w_M) = M\widetilde{w}_M$. Hence,

$$M\left(\mathbf{v}_n(w_M) - \widetilde{w}_M\right) = \underbrace{M\mathbf{v}_n(w_M)}_{=M\widetilde{w}_M} - M\widetilde{w}_M = M\widetilde{w}_M - M\widetilde{w}_M = 0.$$

Thus, (13.87.53) (applied to $a = \mathbf{v}_n(w_M) - \widetilde{w}_M$) yields $\mathbf{v}_n(w_M) - \widetilde{w}_M = 0$, so that $\mathbf{v}_n(w_M) = \widetilde{w}_M$. In other words, (13.87.52) holds for $m = M$. This completes the induction step. Thus, (13.87.52) is proven.

Now, every positive integer m satisfies

$$\begin{aligned}
\mathbf{v}_n(w_m) &= \widetilde{w}_m \qquad (\text{by (13.87.52)}) \\
&= \begin{cases} w_{m/n}, & \text{if } n \mid m; \\ 0, & \text{if } n \nmid m \end{cases} \qquad (\text{by the definition of } \widetilde{w}_m).
\end{aligned}$$

Thus, Exercise 2.9.10(i) is solved in the case when $\mathbf{k} = \mathbb{Z}$. In other words, Step 1 is finished.

Step 2: We shall now show that Exercise 2.9.10(i) holds in the general case.

Indeed, let us use all the notations that we introduced in Step 2 of the solution of Exercise 2.9.10(g). Let us furthermore introduce one more notation:

- For every positive integer m, we shall denote the element w_m of Λ by $w_m^{[\mathbf{k}]}$. This notation makes the dependency of w_m on the base ring $\mathbf{k}$ more explicit. Of course, the element w_m does not "really" depend on $\mathbf{k}$, in the sense that it is defined in the same way for every $\mathbf{k}$. This entails that for any two commutative rings $\mathbf{m}$ and $\mathbf{n}$ and any ring homomorphism $\varphi : \mathbf{m} \to \mathbf{n}$, we have
$$w_m^{[\mathbf{n}]} = \Lambda^{[\varphi]} \left(w_m^{[\mathbf{m}]} \right) \tag{13.87.55}$$
for every positive integer m.

Now, fix a positive integer m. In Step 1, we have shown that Exercise 2.9.10(i) holds if $\mathbf{k} = \mathbb{Z}$. In other words, we have
$$\mathbf{v}_n \left(w_m \right) = \begin{cases} w_{m/n}, & \text{if } n \mid m; \\ 0, & \text{if } n \nmid m \end{cases} \qquad \text{if } \mathbf{k} = \mathbb{Z}.$$
In other words,
$$\mathbf{v}_n^{[\mathbb{Z}]} \left(w_m^{[\mathbb{Z}]} \right) = \begin{cases} w_{m/n}^{[\mathbb{Z}]}, & \text{if } n \mid m; \\ 0, & \text{if } n \nmid m \end{cases} . \tag{13.87.56}$$

Now,
$$w_m = w_m^{[\mathbf{k}]} = \Lambda^{[\rho_{\mathbf{k}}]} \left(w_m^{[\mathbb{Z}]} \right) \tag{13.87.57}$$
(by (13.87.55) (applied to $\mathbf{m} = \mathbb{Z}$, $\mathbf{n} = \mathbf{k}$ and $\varphi = \rho_{\mathbf{k}}$)). Thus,
$$\begin{aligned}
\underbrace{\mathbf{v}_n}_{=\mathbf{v}_n^{[\mathbf{k}]}} \left(\underbrace{w_m}_{=\Lambda^{[\rho_{\mathbf{k}}]}\left(w_m^{[\mathbb{Z}]}\right)} \right) &= \mathbf{v}_n^{[\mathbf{k}]} \left(\Lambda^{[\rho_{\mathbf{k}}]} \left(w_m^{[\mathbb{Z}]} \right) \right) = \underbrace{\left(\mathbf{v}_n^{[\mathbf{k}]} \circ \Lambda^{[\rho_{\mathbf{k}}]} \right)}_{\substack{=\Lambda^{[\rho_{\mathbf{k}}]} \circ \mathbf{v}_n^{[\mathbb{Z}]} \\ \text{(by (13.87.39), applied to} \\ \mathbf{m}=\mathbb{Z},\ \mathbf{n}=\mathbf{k} \text{ and } \varphi=\rho_{\mathbf{k}})}} \left(w_m^{[\mathbb{Z}]} \right) \\
&= \left(\Lambda^{[\rho_{\mathbf{k}}]} \circ \mathbf{v}_n^{[\mathbb{Z}]} \right) \left(w_m^{[\mathbb{Z}]} \right) = \Lambda^{[\rho_{\mathbf{k}}]} \left(\underbrace{\mathbf{v}_n^{[\mathbb{Z}]} \left(w_m^{[\mathbb{Z}]} \right)}_{\substack{= \begin{cases} w_{m/n}^{[\mathbb{Z}]}, & \text{if } n \mid m; \\ 0, & \text{if } n \nmid m \end{cases} \\ \text{(by (13.87.56))}}} \right) \\
&= \Lambda^{[\rho_{\mathbf{k}}]} \left(\begin{cases} w_{m/n}^{[\mathbb{Z}]}, & \text{if } n \mid m; \\ 0, & \text{if } n \nmid m \end{cases} \right) = \begin{cases} \Lambda^{[\rho_{\mathbf{k}}]} \left(w_{m/n}^{[\mathbb{Z}]} \right), & \text{if } n \mid m; \\ \Lambda^{[\rho_{\mathbf{k}}]} \left(0 \right), & \text{if } n \nmid m \end{cases} \\
&= \begin{cases} \Lambda^{[\rho_{\mathbf{k}}]} \left(w_{m/n}^{[\mathbb{Z}]} \right), & \text{if } n \mid m; \\ 0, & \text{if } n \nmid m \end{cases} \\
&\qquad \left(\begin{array}{c} \text{since } \Lambda^{[\rho_{\mathbf{k}}]} \left(0 \right) = 0 \text{ (because } \Lambda^{[\rho_{\mathbf{k}}]} \text{ is a ring homomorphism)} \\ \text{in the case when } n \nmid m \end{array} \right) .
\end{aligned} \tag{13.87.58}$$

On the other hand, we have
$$\Lambda^{[\rho_{\mathbf{k}}]} \left(w_{m/n}^{[\mathbb{Z}]} \right) = w_{m/n} \qquad \text{in the case when } n \mid m \tag{13.87.59}$$

[700]. Now, (13.87.58) becomes

$$\mathbf{v}_n (w_m) = \begin{cases} \Lambda^{[\rho_{\mathbf{k}}]} \left(w^{[\mathbb{Z}]}_{m/n} \right), & \text{if } n \mid m; \\ 0, & \text{if } n \nmid m \end{cases} = \begin{cases} w_{m/n}, & \text{if } n \mid m; \\ 0, & \text{if } n \nmid m \end{cases}$$

$$\left(\begin{array}{c} \text{since } \Lambda^{[\rho_{\mathbf{k}}]} \left(w^{[\mathbb{Z}]}_{m/n} \right) = w_{m/n} \text{ (according to (13.87.59))} \\ \text{in the case when } n \mid m \end{array} \right).$$

Thus, Exercise 2.9.10(i) holds. This completes Step 2, and thus Exercise 2.9.10(i) is solved.

13.88. **Solution to Exercise 2.9.11.** *Solution to Exercise 2.9.11.* (b) Let us first notice that

$$X_{n,d,s} = \sum_{\substack{w=(w_1,w_2,\ldots,w_n)\in\{1,2,3,\ldots\}^n; \\ |\operatorname{Des}(w)|=d;\ |\operatorname{Stag}(w)|=s}} x_{w_1} x_{w_2} \cdots x_{w_n} \qquad \text{for all } d \in \mathbb{N} \text{ and } s \in \mathbb{N}. \tag{13.88.1}$$

(This is just the definition of $X_{n,d,s}$, written as a formula.)

Let us first show that if n is a positive integer, then any $d \in \mathbb{N}$ and $s \in \mathbb{N}$ satisfy

$$\begin{aligned} &(d+1) X_{n,d+1,s} + (s+1) X_{n,d,s+1} + (n-1-d-s) X_{n,d,s} \\ &= \sum_{i=1}^{n-1} \sum_{e=0}^{d} \sum_{t=0}^{s} X_{i,e,t} X_{n-i,d-e,s-t}. \end{aligned} \tag{13.88.2}$$

Proof of (13.88.2): Let n be a positive integer. Let $d \in \mathbb{N}$ and $s \in \mathbb{N}$. Let $i \in \{1, 2, ..., n-1\}$ be arbitrary. We make some more definitions:

- Define a power series $\mathcal{D}_i \in \mathbf{k}[[\mathbf{x}]]$ as the sum of the monomials $x_{w_1} x_{w_2} \cdots x_{w_n}$ over all n-tuples $w = (w_1, w_2, ..., w_n) \in \{1, 2, 3, \ldots\}^n$ satisfying $|\operatorname{Des}(w)| = d+1$, $|\operatorname{Stag}(w)| = s$ and $w_i > w_{i+1}$.
- Define a power series $\mathcal{S}_i \in \mathbf{k}[[\mathbf{x}]]$ as the sum of the monomials $x_{w_1} x_{w_2} \cdots x_{w_n}$ over all n-tuples $w = (w_1, w_2, ..., w_n) \in \{1, 2, 3, \ldots\}^n$ satisfying $|\operatorname{Des}(w)| = d$, $|\operatorname{Stag}(w)| = s+1$ and $w_i = w_{i+1}$.
- Define a power series $\mathcal{A}_i \in \mathbf{k}[[\mathbf{x}]]$ as the sum of the monomials $x_{w_1} x_{w_2} \cdots x_{w_n}$ over all n-tuples $w = (w_1, w_2, ..., w_n) \in \{1, 2, 3, \ldots\}^n$ satisfying $|\operatorname{Des}(w)| = d$, $|\operatorname{Stag}(w)| = s$ and $w_i < w_{i+1}$.

These three power series $\mathcal{D}_i$, $\mathcal{S}_i$ and $\mathcal{A}_i$ are not necessarily symmetric (but will nevertheless come useful). Let us notice that the definitions of $\mathcal{D}_i$, $\mathcal{S}_i$ and $\mathcal{A}_i$ can be rewritten as follows:

- The power series $\mathcal{D}_i \in \mathbf{k}[[\mathbf{x}]]$ is the sum of the monomials $x_{w_1} x_{w_2} \cdots x_{w_n}$ over all n-tuples $w = (w_1, w_2, ..., w_n) \in \{1, 2, 3, \ldots\}^n$ satisfying $|(\operatorname{Des}(w)) \setminus \{i\}| = d$, $|(\operatorname{Stag}(w)) \setminus \{i\}| = s$ and $w_i > w_{i+1}$.
- The power series $\mathcal{S}_i \in \mathbf{k}[[\mathbf{x}]]$ is the sum of the monomials $x_{w_1} x_{w_2} \cdots x_{w_n}$ over all n-tuples $w = (w_1, w_2, ..., w_n) \in \{1, 2, 3, \ldots\}^n$ satisfying $|(\operatorname{Des}(w)) \setminus \{i\}| = d$, $|(\operatorname{Stag}(w)) \setminus \{i\}| = s$ and $w_i = w_{i+1}$.
- The power series $\mathcal{A}_i \in \mathbf{k}[[\mathbf{x}]]$ is the sum of the monomials $x_{w_1} x_{w_2} \cdots x_{w_n}$ over all n-tuples $w = (w_1, w_2, ..., w_n) \in \{1, 2, 3, \ldots\}^n$ satisfying $|(\operatorname{Des}(w)) \setminus \{i\}| = d$, $|(\operatorname{Stag}(w)) \setminus \{i\}| = s$ and $w_i < w_{i+1}$.

These reformulations make it obvious that the sum $\mathcal{D}_i + \mathcal{S}_i + \mathcal{A}_i$ is precisely the sum of the monomials $x_{w_1} x_{w_2} \cdots x_{w_n}$ over all n-tuples $w = (w_1, w_2, ..., w_n) \in \{1, 2, 3, \ldots\}^n$ satisfying $|(\operatorname{Des}(w)) \setminus \{i\}| = d$ and $|(\operatorname{Stag}(w)) \setminus \{i\}| = s$. In other words,

$$\mathcal{D}_i + \mathcal{S}_i + \mathcal{A}_i = \sum_{\substack{w=(w_1,w_2,\ldots,w_n)\in\{1,2,3,\ldots\}^n; \\ |(\operatorname{Des}(w))\setminus\{i\}|=d;\ |(\operatorname{Stag}(w))\setminus\{i\}|=s}} x_{w_1} x_{w_2} \cdots x_{w_n}. \tag{13.88.3}$$

But the sum on the right hand side of (13.88.3) can be rewritten in terms of more familiar sums. In fact, the n-tuples $w \in \{1, 2, 3, \ldots\}^n$ are in bijection with the pairs (u, v) consisting of an i-tuple $u \in \{1, 2, 3, \ldots\}^i$ and an $(n-i)$-tuple $v \in \{1, 2, 3, \ldots\}^{n-i}$. This bijection sends an n-tuple $(w_1, w_2, ..., w_n)$ to the

[700] *Proof of (13.87.59):* Assume that $n \mid m$. Then, m/n is a positive integer. Hence, $w_{m/n} = \Lambda^{[\rho_{\mathbf{k}}]} \left(w^{[\mathbb{Z}]}_{m/n} \right)$ (in fact, this follows from the same argument that was used to prove (13.87.57), but with m replaced by m/n). In other words, $\Lambda^{[\rho_{\mathbf{k}}]} \left(w^{[\mathbb{Z}]}_{m/n} \right) = w_{m/n}$. This proves (13.87.59).

pair $((w_1, w_2, ..., w_i), (w_{i+1}, w_{i+2}, ..., w_n))$, and has the property that $|(\operatorname{Des}(w)) \setminus \{i\}| = |\operatorname{Des}(u)| + |\operatorname{Des}(v)|$ [701] and $|(\operatorname{Stag}(w)) \setminus \{i\}| = |\operatorname{Stag}(u)| + |\operatorname{Stag}(v)|$ [702]. Hence,

$$\begin{aligned}
&\sum_{\substack{w=(w_1,w_2,...,w_n)\in\{1,2,3,...\}^n;\\ |(\operatorname{Des}(w))\setminus\{i\}|=d;\ |(\operatorname{Stag}(w))\setminus\{i\}|=s}} x_{w_1}x_{w_2}\cdots x_{w_n} \\
&= \underbrace{\sum_{\substack{u=(u_1,u_2,...,u_i)\in\{1,2,3,...\}^i;\\ v=(v_1,v_2,...,v_{n-i})\in\{1,2,3,...\}^{n-i};\\ |\operatorname{Des}(u)|+|\operatorname{Des}(v)|=d;\\ |\operatorname{Stag}(u)|+|\operatorname{Stag}(v)|=s}}}_{=\sum_{e=0}^{d}\sum_{t=0}^{s}\sum_{\substack{u=(u_1,u_2,...,u_i)\in\{1,2,3,...\}^i;\\ |\operatorname{Des}(u)|=e;\ |\operatorname{Stag}(u)|=t}}\sum_{\substack{v=(v_1,v_2,...,v_{n-i})\in\{1,2,3,...\}^{n-i};\\ |\operatorname{Des}(v)|=d-e;\ |\operatorname{Stag}(v)|=s-t}}} x_{u_1}x_{u_2}\cdots x_{u_i}x_{v_1}x_{v_2}\cdots x_{v_{n-i}} \\
&= \sum_{e=0}^{d}\sum_{t=0}^{s} \sum_{\substack{u=(u_1,u_2,...,u_i)\in\{1,2,3,...\}^i;\\ |\operatorname{Des}(u)|=e;\ |\operatorname{Stag}(u)|=t}} \sum_{\substack{v=(v_1,v_2,...,v_{n-i})\in\{1,2,3,...\}^{n-i};\\ |\operatorname{Des}(v)|=d-e;\ |\operatorname{Stag}(v)|=s-t}} x_{u_1}x_{u_2}\cdots x_{u_i}x_{v_1}x_{v_2}\cdots x_{v_{n-i}} \\
&= \sum_{e=0}^{d}\sum_{t=0}^{s} \left(\sum_{\substack{u=(u_1,u_2,...,u_i)\in\{1,2,3,...\}^i;\\ |\operatorname{Des}(u)|=e;\ |\operatorname{Stag}(u)|=t}} x_{u_1}x_{u_2}\cdots x_{u_i} \right) \left(\sum_{\substack{v=(v_1,v_2,...,v_{n-i})\in\{1,2,3,...\}^{n-i};\\ |\operatorname{Des}(v)|=d-e;\ |\operatorname{Stag}(v)|=s-t}} x_{v_1}x_{v_2}\cdots x_{v_{n-i}} \right). \qquad (13.88.4)
\end{aligned}$$

However, for every $e \in \{0, 1, ..., d\}$ and $t \in \{0, 1, ..., s\}$, the equality (13.88.1) (applied to i, e and t instead of n, d and s) yields

$$X_{i,e,t} = \sum_{\substack{w=(w_1,w_2,...,w_i)\in\{1,2,3,...\}^i;\\ |\operatorname{Des}(w)|=e;\ |\operatorname{Stag}(w)|=t}} x_{w_1}x_{w_2}\cdots x_{w_i} = \sum_{\substack{u=(u_1,u_2,...,u_i)\in\{1,2,3,...\}^i;\\ |\operatorname{Des}(u)|=e;\ |\operatorname{Stag}(u)|=t}} x_{u_1}x_{u_2}\cdots x_{u_i}. \tag{13.88.5}$$

Also, for every $e \in \{0, 1, ..., d\}$ and $t \in \{0, 1, ..., s\}$, the equality (13.88.1) (applied to $n - i$, $d - e$ and $s - t$ instead of n, d and s) yields

$$X_{n-i,d-e,s-t} = \sum_{\substack{w=(w_1,w_2,...,w_{n-i})\in\{1,2,3,...\}^{n-i};\\ |\operatorname{Des}(w)|=d-e;\ |\operatorname{Stag}(w)|=s-t}} x_{w_1}x_{w_2}\cdots x_{w_{n-i}} = \sum_{\substack{v=(v_1,v_2,...,v_{n-i})\in\{1,2,3,...\}^{n-i};\\ |\operatorname{Des}(v)|=d-e;\ |\operatorname{Stag}(v)|=s-t}} x_{v_1}x_{v_2}\cdots x_{v_{n-i}}. \tag{13.88.6}$$

Now, (13.88.4) becomes

$$\begin{aligned}
&\sum_{\substack{w=(w_1,w_2,...,w_n)\in\{1,2,3,...\}^n;\\ |(\operatorname{Des}(w))\setminus\{i\}|=d;\ |(\operatorname{Stag}(w))\setminus\{i\}|=s}} x_{w_1}x_{w_2}\cdots x_{w_n} \\
&= \sum_{e=0}^{d}\sum_{t=0}^{s} \underbrace{\left(\sum_{\substack{u=(u_1,u_2,...,u_i)\in\{1,2,3,...\}^i;\\ |\operatorname{Des}(u)|=e;\ |\operatorname{Stag}(u)|=t}} x_{u_1}x_{u_2}\cdots x_{u_i} \right)}_{\substack{=X_{i,e,t}\\ \text{(by (13.88.5))}}} \underbrace{\left(\sum_{\substack{v=(v_1,v_2,...,v_{n-i})\in\{1,2,3,...\}^{n-i};\\ |\operatorname{Des}(v)|=d-e;\ |\operatorname{Stag}(v)|=s-t}} x_{v_1}x_{v_2}\cdots x_{v_{n-i}} \right)}_{\substack{=X_{n-i,d-e,s-t}\\ \text{(by (13.88.6))}}} \\
&= \sum_{e=0}^{d}\sum_{t=0}^{s} X_{i,e,t}X_{n-i,d-e,s-t}.
\end{aligned}$$

[701] In fact, $(\operatorname{Des}(w))\setminus\{i\}$ is the union of the set $\operatorname{Des}(u)$ with the set $\operatorname{Des}(v)$ shifted by i (that is, the set $\{p + i \mid p \in \operatorname{Des}(v)\}$). This is a disjoint union, and thus we find $|(\operatorname{Des}(w)) \setminus \{i\}| = |\operatorname{Des}(u)| + |\operatorname{Des}(v)|$ (since the set $\operatorname{Des}(v)$ shifted by i has cardinality $|\operatorname{Des}(v)|$).

[702] for similar reasons

Hence, (13.88.3) becomes

$$\mathcal{D}_i + \mathcal{S}_i + \mathcal{A}_i = \sum_{\substack{w=(w_1,w_2,\ldots,w_n)\in\{1,2,3,\ldots\}^n;\\ |(\operatorname{Des}(w))\setminus\{i\}|=d;\ |(\operatorname{Stag}(w))\setminus\{i\}|=s}} x_{w_1}x_{w_2}\cdots x_{w_n}$$

$$= \sum_{e=0}^{d}\sum_{t=0}^{s} X_{i,e,t}X_{n-i,d-e,s-t}. \qquad (13.88.7)$$

Now, let us forget that we fixed i. We thus have defined $\mathcal{D}_i$, $\mathcal{S}_i$ and $\mathcal{A}_i$ and proven the equality (13.88.7) for all $i \in \{1, 2, ..., n-1\}$.

Now, let us take a closer look at the sums $\sum_{i=1}^{n-1}\mathcal{D}_i$, $\sum_{i=1}^{n-1}\mathcal{S}_i$ and $\sum_{i=1}^{n-1}\mathcal{A}_i$:

- The definition of $\mathcal{D}_i$ can be rewritten as follows:

$$\mathcal{D}_i = \sum_{\substack{w=(w_1,w_2,\ldots,w_n)\in\{1,2,3,\ldots\}^n;\\ |\operatorname{Des}(w)|=d+1;\ |\operatorname{Stag}(w)|=s;\ w_i>w_{i+1}}} x_{w_1}x_{w_2}\cdots x_{w_n} \qquad \text{for every } i\in\{1,2,...,n-1\}.$$

Summing up these equations over all $i \in \{1, 2, ..., n-1\}$, we obtain

$$\sum_{i=1}^{n-1}\mathcal{D}_i = \sum_{i=1}^{n-1}\sum_{\substack{w=(w_1,w_2,\ldots,w_n)\in\{1,2,3,\ldots\}^n;\\ |\operatorname{Des}(w)|=d+1;\ |\operatorname{Stag}(w)|=s;\ w_i>w_{i+1}}} x_{w_1}x_{w_2}\cdots x_{w_n}$$

$$= \sum_{\substack{w=(w_1,w_2,\ldots,w_n)\in\{1,2,3,\ldots\}^n;\\ |\operatorname{Des}(w)|=d+1;\ |\operatorname{Stag}(w)|=s}} \underbrace{\sum_{\substack{i\in\{1,2,\ldots,n-1\};\\ w_i>w_{i+1}}} x_{w_1}x_{w_2}\cdots x_{w_n}}_{=|\{i\in\{1,2,\ldots,n-1\}\ |\ w_i>w_{i+1}\}|x_{w_1}x_{w_2}\cdots x_{w_n}}$$

$$= \sum_{\substack{w=(w_1,w_2,\ldots,w_n)\in\{1,2,3,\ldots\}^n;\\ |\operatorname{Des}(w)|=d+1;\ |\operatorname{Stag}(w)|=s}} \underbrace{|\{i\in\{1,2,...,n-1\}\ \mid\ w_i>w_{i+1}\}|}_{=|\operatorname{Des} w|=d+1} x_{w_1}x_{w_2}\cdots x_{w_n}$$

$$= (d+1)\underbrace{\sum_{\substack{w=(w_1,w_2,\ldots,w_n)\in\{1,2,3,\ldots\}^n;\\ |\operatorname{Des}(w)|=d+1;\ |\operatorname{Stag}(w)|=s}} x_{w_1}x_{w_2}\cdots x_{w_n}}_{\substack{=X_{n,d+1,s}\\ \text{(by (13.88.1), applied to } d+1 \text{ instead of } d)}}$$

$$= (d+1)\,X_{n,d+1,s}. \qquad (13.88.8)$$

- Similarly, we can see that

$$\sum_{i=1}^{n-1}\mathcal{S}_i = (s+1)\,X_{n,d,s+1}. \qquad (13.88.9)$$

- Similarly, we can see that

$$\sum_{i=1}^{n-1}\mathcal{A}_i = (n-1-d-s)\,X_{n,d,s}. \qquad (13.88.10)$$

[703]

Now, summing the equality (13.88.7) over all $i \in \{1, 2, ..., n-1\}$ yields

$$\sum_{i=1}^{n-1}(\mathcal{D}_i + \mathcal{S}_i + \mathcal{A}_i) = \sum_{i=1}^{n-1}\sum_{e=0}^{d}\sum_{t=0}^{s} X_{i,e,t}X_{n-i,d-e,s-t}.$$

[703] In proving this, we have to observe that $|\{i \in \{1, 2, ..., n-1\} \mid w_i < w_{i+1}\}| = n - 1 - |\operatorname{Des} w| - |\operatorname{Stag} w|$ for every $w = (w_1, w_2, \ldots, w_n) \in \{1, 2, 3, ...\}^n$. This is clear from realizing that the sets $\{i \in \{1, 2, ..., n-1\} \mid w_i < w_{i+1}\}$, $\operatorname{Des} w$ and $\operatorname{Stag} w$ are disjoint and their union is $\{1, 2, ..., n-1\}$.

Hence,

$$\begin{aligned}
&\sum_{i=1}^{n-1}\sum_{e=0}^{d}\sum_{t=0}^{s} X_{i,e,t}X_{n-i,d-e,s-t} \\
&= \sum_{i=1}^{n-1}\left(\mathcal{D}_i+\mathcal{S}_i+\mathcal{A}_i\right) = \underbrace{\sum_{i=1}^{n-1}\mathcal{D}_i}_{\substack{=(d+1)X_{n,d+1,s} \\ \text{(by (13.88.8))}}} + \underbrace{\sum_{i=1}^{n-1}\mathcal{S}_i}_{\substack{=(s+1)X_{n,d,s+1} \\ \text{(by (13.88.9))}}} + \underbrace{\sum_{i=1}^{n-1}\mathcal{A}_i}_{\substack{=(n-1-d-s)X_{n,d,s} \\ \text{(by (13.88.10))}}} \\
&= (d+1)\,X_{n,d+1,s} + (s+1)\,X_{n,d,s+1} + (n-1-d-s)\,X_{n,d,s}.
\end{aligned}$$

This proves (13.88.2).

In solving Exercise 2.9.11(b), we shall use the following trick to simplify our life. Recall that there is a canonical ring homomorphism $\varphi : \mathbb{Z} \to \mathbf{k}$. This homomorphism gives rise to a ring homomorphism $\varphi[[\mathbf{x}]] : \mathbb{Z}[[\mathbf{x}]] \to \mathbf{k}[[\mathbf{x}]]$, and this latter homomorphism $\varphi[[\mathbf{x}]]$ sends $\Lambda_{\mathbb{Z}}$ to $\Lambda_{\mathbf{k}}$; that is, we have $(\varphi[[\mathbf{x}]])(\Lambda_{\mathbb{Z}}) \subset \Lambda_{\mathbf{k}}$. Moreover, it is clear that (for any nonnegative integers d and s) the ring homomorphism $\varphi[[\mathbf{x}]]$ sends the element $X_{n,d,s}$ of $\mathbb{Z}[[\mathbf{x}]]$ to the element $X_{n,d,s}$ of $\mathbf{k}[[\mathbf{x}]]$ (because the definition of $X_{n,d,s}$ is functorial in the base ring $\mathbf{k}$). Therefore, if we can prove (for given nonnegative integers d and s) that the element $X_{n,d,s}$ of $\mathbb{Z}[[\mathbf{x}]]$ belongs to $\Lambda_{\mathbb{Z}}$, then it will automatically follow that the element $X_{n,d,s}$ of $\mathbf{k}[[\mathbf{x}]]$ belongs to $(\varphi[[\mathbf{x}]])(\Lambda_{\mathbb{Z}}) \subset \Lambda_{\mathbf{k}}$; this will complete the solution to Exercise 2.9.11(b). Hence, in order to solve Exercise 2.9.11(b), it only remains to prove (for any nonnegative integers d and s) that the element $X_{n,d,s}$ of $\mathbb{Z}[[\mathbf{x}]]$ belongs to $\Lambda_{\mathbb{Z}}$. In other words, it only remains to solve Exercise 2.9.11(b) in the case of $\mathbf{k} = \mathbb{Z}$. Hence, in solving Exercise 2.9.11(b), we can WLOG assume that $\mathbf{k} = \mathbb{Z}$. Assume this.

Now, we are going to solve Exercise 2.9.11(b) by strong induction over $n + d$. So (for the induction step) we need to show that $X_{n,d,s} \in \Lambda$, and we can assume (as the induction hypothesis) that $X_{n',d',s'} \in \Lambda$ is already known to hold for any nonnegative integers n', d' and s' satisfying $n' + d' < n + d$.

We must be in one of the following two cases:

Case 1: We have $d > 0$.

Case 2: We have $d = 0$.

Let us first consider Case 1. In this case, we have $d > 0$. Hence, $d - 1 \in \mathbb{N}$. We also WLOG assume that n is positive (otherwise, $X_{n,d,s}$ is a constant and thus lies in Λ for sure). Applying (13.88.2) to $d - 1$ instead of d, we obtain

$$dX_{n,d,s} + (s+1)\,X_{n,d-1,s+1} + (n-d-s)\,X_{n,d-1,s} = \sum_{i=1}^{n-1}\sum_{e=0}^{d-1}\sum_{t=0}^{s} X_{i,e,t}X_{n-i,d-1-e,s-t}.$$

Thus,

$$dX_{n,d,s} = \sum_{i=1}^{n-1}\sum_{e=0}^{d-1}\sum_{t=0}^{s} X_{i,e,t}X_{n-i,d-1-e,s-t} - (s+1)\,X_{n,d-1,s+1} + (n-d-s)\,X_{n,d-1,s}. \tag{13.88.11}$$

Each of the power series $X_{i,e,t}$, $X_{n-i,d-1-e,s-t}$, $X_{n,d-1,s+1}$ and $X_{n,d-1,s}$ on the right hand side of this equality is already known to lie in Λ (by the induction hypothesis). Hence, the right hand side of (13.88.11) lies in Λ (since Λ is a ring), and thus (13.88.11) shows that $dX_{n,d,s} \in \Lambda$. Since $d > 0$, this yields $X_{n,d,s} \in \Lambda$ (because if a power series $Q \in \mathbb{Z}[[\mathbf{x}]]$ satisfies $dQ \in \Lambda$, then Q must belong to Λ itself[704]). Hence, the induction step is complete in Case 1.

Let us now consider Case 2. In this case, we have $d = 0$. Hence, $X_{n,d,s} = X_{n,0,s}$.

We shall now show that

$$X_{n,0,s} = \sum_{\substack{\lambda\in\operatorname{Par}_n; \\ \ell(\lambda)=n-s}} m_\lambda. \tag{13.88.12}$$

[704] It is here that we are using our assumption that $\mathbf{k} = \mathbb{Z}$.

Proof of (13.88.12): Summing up the equality (2.1.1) over all $\lambda \in \operatorname{Par}_n$ satisfying $\ell(\lambda) = n - s$, we obtain

$$\sum_{\substack{\lambda \in \operatorname{Par}_n; \\ \ell(\lambda) = n-s}} m_\lambda = \sum_{\substack{\lambda \in \operatorname{Par}_n; \\ \ell(\lambda) = n-s}} \sum_{\alpha \in \mathfrak{S}_{(\infty)}\lambda} \mathbf{x}^\alpha . \tag{13.88.13}$$

On the other hand, the equality (13.88.1) (applied to 0 instead of d) yields

$$X_{n,0,s} = \sum_{\substack{w=(w_1,w_2,\ldots,w_n) \in \{1,2,3,\ldots\}^n; \\ |\operatorname{Des}(w)|=0; \ |\operatorname{Stag}(w)|=s}} x_{w_1} x_{w_2} \cdots x_{w_n} . \tag{13.88.14}$$

The condition $|\operatorname{Des}(w)| = 0$ on an n-tuple $w = (w_1, w_2, \ldots, w_n) \in \{1, 2, 3, \ldots\}^n$ is equivalent to $w_1 \leq w_2 \leq \ldots \leq w_n$, and therefore (13.88.14) rewrites as

$$X_{n,0,s} = \sum_{\substack{w=(w_1,w_2,\ldots,w_n) \in \{1,2,3,\ldots\}^n; \\ w_1 \leq w_2 \leq \ldots \leq w_n; \ |\operatorname{Stag}(w)|=s}} x_{w_1} x_{w_2} \cdots x_{w_n} . \tag{13.88.15}$$

For an n-tuple $w = (w_1, w_2, \ldots, w_n) \in \{1, 2, 3, \ldots\}^n$ satisfying $w_1 \leq w_2 \leq \ldots \leq w_n$, the condition $|\operatorname{Stag}(w)| = s$ is equivalent to the condition that $|\{w_1, w_2, \ldots, w_n\}| = n - s$. Hence, (13.88.15) rewrites as

$$X_{n,0,s} = \sum_{\substack{w=(w_1,w_2,\ldots,w_n) \in \{1,2,3,\ldots\}^n; \\ w_1 \leq w_2 \leq \ldots \leq w_n; \ |\{w_1,w_2,\ldots,w_n\}|=n-s}} x_{w_1} x_{w_2} \cdots x_{w_n} . \tag{13.88.16}$$

For any fixed weak composition $\alpha = (\alpha_1, \alpha_2, \alpha_3, \ldots)$, the monomial $\mathbf{x}^\alpha$ occurs **at most once** on the right hand side of (13.88.16) (because there is at most one way to write $\mathbf{x}^\alpha$ in the form $x_{w_1} x_{w_2} \cdots x_{w_n}$ for a $w = (w_1, w_2, \ldots, w_n) \in \{1, 2, 3, \ldots\}^n$ satisfying $w_1 \leq w_2 \leq \ldots \leq w_n$). We can easily tell whether it occurs or not by looking at the size $|\alpha|$ of α and the number of positive integers i satisfying $\alpha_i \neq 0$: Namely, the monomial $\mathbf{x}^\alpha$ for a fixed weak composition $\alpha = (\alpha_1, \alpha_2, \alpha_3, \ldots)$ occurs on the right hand side of (13.88.16) if and only if it satisfies the following two properties:

- We have $|\alpha| = n$.
- There are precisely $n - s$ positive integers i satisfying $\alpha_i \neq 0$.

Thus, the monomials satisfying these two properties occur exactly once on the right hand side of (13.88.16), while all other monomials don't occur there at all. But the same conclusion can be reached for the right hand side of (13.88.13). Hence, for any fixed weak composition $\alpha = (\alpha_1, \alpha_2, \alpha_3, \ldots)$, the monomial $\mathbf{x}^\alpha$ occurs on the right hand side of (13.88.16) precisely as often as it occurs on the right hand side of (13.88.13). Hence, the right hand side of (13.88.16) equals the right hand side of (13.88.13). Therefore, the left hand side of (13.88.16) also equals the left hand side of (13.88.13). In other words, $X_{n,0,s} = \sum_{\substack{\lambda \in \operatorname{Par}_n; \\ \ell(\lambda)=n-s}} m_\lambda$. This proves (13.88.12).

Now, (13.88.12) yields $X_{n,0,s} = \sum_{\substack{\lambda \in \operatorname{Par}_n; \\ \ell(\lambda)=n-s}} \underbrace{m_\lambda}_{\in \Lambda} \in \Lambda$, so that $X_{n,d,s} = X_{n,0,s} \in \Lambda$. This completes the induction step in Case 2.

Thus, the induction step is complete in both Cases 1 and 2. This finally completes the induction step. Exercise 2.9.11(b) is thus solved.

(a) *First solution of Exercise 2.9.11(a):* Recall the power series $X_{n,d,s}$ defined in Exercise 2.9.11(b) for any $d \in \mathbb{N}$ and $s \in \mathbb{N}$. Exercise 2.9.11(b) yields that $X_{n,k,0} \in \Lambda$. But an n-tuple w is Smirnov if and only if its stagnation set $\operatorname{Stag}(w)$ is empty, i.e., if and only if $|\operatorname{Stag}(w)| = 0$. Hence, $X_{n,k,0}$ is the sum of the monomials $x_{w_1} x_{w_2} \cdots x_{w_n}$ over all Smirnov n-tuples $w \in \{1, 2, 3, \ldots\}^n$ satisfying $|\operatorname{Des}(w)| = k$. This shows that $X_{n,k,0} = X_{n,k}$. Thus, $X_{n,k} = X_{n,k,0} \in \Lambda$, so that Exercise 2.9.11(a) is solved.

Second solution of Exercise 2.9.11(a): Here is an alternative solution to Exercise 2.9.11(a), which avoids using part (b). It is an application of [199, proof of Theorem 4.5] to our special setting.

We proceed similarly to the proof of Proposition 2.2.4: It suffices to show that for every positive integer p, the power series $X_{n,k}$ is invariant under swapping the variables x_p and x_{p+1}. So let us fix a positive integer p.

Let $S_{n,k}$ denote the set of all Smirnov n-tuples $w \in \{1,2,3,\ldots\}^n$ satisfying $|\operatorname{Des}(w)| = k$. Then, the definition of $X_{n,k}$ rewrites as follows:

$$X_{n,k} = \sum_{w=(w_1,w_2,\ldots,w_n)\in S_{n,k}} x_{w_1}x_{w_2}\ldots x_{w_n}. \tag{13.88.17}$$

For every $w \in S_{n,k}$, let $\mathbf{x}_w$ denote the monomial $x_{w_1}x_{w_2}\ldots x_{w_n}$, where w is written in the form $w = (w_1, w_2, \ldots, w_n)$. Then, (13.88.17) rewrites as

$$X_{n,k} = \sum_{w\in S_{n,k}} \mathbf{x}_w. \tag{13.88.18}$$

We need to show that this power series $X_{n,k}$ is invariant under swapping the variables x_p and x_{p+1}. In order to do so, it is clearly enough to provide an involution $J : S_{n,k} \to S_{n,k}$ which has the property that for every $w \in S_{n,k}$, the monomial $\mathbf{x}_{J(w)}$ is obtained from the monomial $\mathbf{x}_w$ by swapping the variables x_p and x_{p+1}.

To define the involution J, we introduce some notations. If $w = (w_1, w_2, \ldots, w_n) \in \{1,2,3,\ldots\}^n$ is any n-tuple, then a nonempty interval I of $\{1,2,\ldots,n\}$ will be called a $(p,p+1)$-*interval* of w if every $i \in I$ satisfies $w_i \in \{p,p+1\}$. A $(p,p+1)$-*run* of w will mean a $(p,p+1)$-interval of w maximal with respect to inclusion. For instance, if $n = 9$, $w = (3,1,2,5,4,2,3,2,4)$ and $p = 2$, then the $(p,p+1)$-intervals of w are $\{1\}$, $\{3\}$, $\{6\}$, $\{7\}$, $\{8\}$, $\{6,7\}$, $\{7,8\}$ and $\{6,7,8\}$ (since the letters of w belonging to $\{p,p+1\}$ are the 1-st, 3-rd, the 6-th, the 7-th and the 8-th letter), while only $\{1\}$, $\{3\}$ and $\{6,7,8\}$ are $(p,p+1)$-runs of w.

An interval of $\{1,2,\ldots,n\}$ is said to be *even* if it has even size, and *odd* if it has odd size.

It is easy to see that if $w = (w_1, w_2, \ldots, w_n) \in \{1,2,3,\ldots\}^n$ is any Smirnov n-tuple, and $\{a+1, a+2, \ldots, b\}$ is any $(p,p+1)$-interval of w, then the $(b-a)$-tuple $(w_{a+1}, w_{a+2}, \ldots, w_b)$ alternates between p's and $(p+1)$'s, that is, has one of the forms

$$(p, p+1, p, p+1, p, \ldots, p+1) \qquad \text{and} \qquad (p+1, p, p+1, p, p+1, \ldots, p)$$

(if $b-a$ is even) or one of the forms

$$(p, p+1, p, p+1, p, \ldots, p) \qquad \text{and} \qquad (p+1, p, p+1, p, p+1, \ldots, p+1)$$

(if $b-a$ is odd).

Now, we define a map $J : S_{n,k} \to S_{n,k}$ as follows: Let $w = (w_1, w_2, \ldots, w_n) \in S_{n,k}$. We know that w is Smirnov and satisfies $|\operatorname{Des}(w)| = k$. It is clear that the $(p,p+1)$-runs of w are disjoint and even separated from each other by at least 1 (this means that if α and β are elements of two distinct $(p,p+1)$-runs of w, then $|\alpha - \beta| > 1$). For every $i \in \{1,2,\ldots,n\}$, define a positive integer w_i' as follows:

- If i belongs to an odd $(p,p+1)$-run of w, then set $w_i' = \begin{cases} p+1, & \text{if } w_i = p; \\ p, & \text{if } w_i = p+1 \end{cases}$.
- Otherwise, set $w_i' = w_i$.

Thus, we have defined an n-tuple $(w_1', w_2', \ldots, w_n')$ of positive integers.[705] Denote this n-tuple $(w_1', w_2', \ldots, w_n')$ by w'. The reader can easily check that this new n-tuple w' is again Smirnov and satisfies $|\operatorname{Des}(w')| = k$. In other words, $w' \in S_{n,k}$. Now, set $J(w) = w'$. We have thus defined a map $J : S_{n,k} \to S_{n,k}$. It is easy to see that, for every $w \in S_{n,k}$, the $(p,p+1)$-runs of $J(w)$ are exactly the $(p,p+1)$-runs of w, and applying the map J to $J(w)$ precisely reverts the changes made by the map J to w. In other words, $J \circ J = \mathrm{id}$, so that the map J is an involution. Finally, it is straightforward to see that for every $w \in S_{n,k}$, we have the following facts:

1. The number of entries equal to $p+1$ in $J(w)$ equals the number of entries equal to p in w.
2. The number of entries equal to p in $J(w)$ equals the number of entries equal to $p+1$ in w.
3. For every $j \in \{1,2,3,\ldots\} \setminus \{p,p+1\}$, the number of entries equal to j in $J(w)$ equals the number of entries equal to j in w.

[705] Informally speaking, $(w_1', w_2', \ldots, w_n')$ is simply obtained by changing those p's and $p+1$'s in w whose positions belong to odd $(p,p+1)$-runs of w into $p+1$'s and p's, respectively, while leaving all other entries of w intact. For instance, in our above example of $n = 9$, $w = (3,1,2,5,4,2,3,2,4)$ and $p = 2$, we would have $(w_1', w_2', \ldots, w_n') = (2,1,3,5,4,3,2,3,4)$. (All letters 2 and 3 have been changed here because all $(p,p+1)$-runs of this w were odd.) For another example, if $n = 5$, $w = (2,3,1,5,2)$ and $p = 2$, then $(w_1', w_2', \ldots, w_n') = (2,3,1,5,3)$ (the first two letters are unchanged since the $(p,p+1)$-run $\{1,2\}$ is even).

[706] These three statements, combined, show that for every $w \in S_{n,k}$, the monomial $\mathbf{x}_{J(w)}$ is obtained from the monomial $\mathbf{x}_w$ by swapping the variables x_p and x_{p+1}. This concludes our solution of Exercise 2.9.11(a).

Remark: We could have given an alternative solution to Exercise 2.9.11(b) that would still rely on (13.88.2), but proceed by induction on $n+s$ (rather than $n+d$) and handle the case $s=0$ separately (rather than the case $d=0$ as we did). In the case $s=0$, the assertion of Exercise 2.9.11(b) follows from Exercise 2.9.11(a). (But this only works combined with a solution to Exercise 2.9.11(a) that does not rely on Exercise 2.9.11(b).)

(c) For every $d \in \mathbb{N}$, $s \in \mathbb{N}$ and $i \in \{1,2,...,n-1\}$, we define the power series $\mathcal{D}_i$, $\mathcal{S}_i$ and $\mathcal{A}_i$ as in the solution to Exercise 2.9.11(b) above.

Let us first show that if n is a positive integer, then any $d \in \mathbb{N}$ and $s \in \mathbb{N}$ satisfy

$$(d+1)\,U_{n,d+1,s} + (s+1)\,U_{n,d,s+1} + (n-1-d-s)\,U_{n,d,s} = (d+1)\,X_{n,d+1,s}. \tag{13.88.19}$$

Proof of (13.88.19): Let n be a positive integer. Let $d \in \mathbb{N}$ and $s \in \mathbb{N}$. Let $i \in \{1,2,...,n-1\}$ be arbitrary.

Before we make any new definitions, let us recall a statement that we have shown during our proof of (13.88.2) (back when we were solving Exercise 2.9.11(b)): The power series $\mathcal{D}_i \in \mathbf{k}[[\mathbf{x}]]$ is the sum of the monomials $x_{w_1}x_{w_2}\cdots x_{w_n}$ over all n-tuples $w=(w_1,w_2,...,w_n) \in \{1,2,3,\ldots\}^n$ satisfying $|(\operatorname{Des}(w)) \setminus \{i\}| = d$, $|(\operatorname{Stag}(w)) \setminus \{i\}| = s$ and $w_i > w_{i+1}$. Written as a formula, this yields

$$\mathcal{D}_i = \sum_{\substack{w=(w_1,w_2,...,w_n)\in\{1,2,3,\ldots\}^n;\\ |(\operatorname{Des}(w))\setminus\{i\}|=d;\ |(\operatorname{Stag}(w))\setminus\{i\}|=s;\\ w_i>w_{i+1}}} x_{w_1}x_{w_2}\cdots x_{w_n}.$$

Applying this to $n-i$ instead of i, we obtain

$$\mathcal{D}_{n-i} = \sum_{\substack{w=(w_1,w_2,...,w_n)\in\{1,2,3,\ldots\}^n;\\ |(\operatorname{Des}(w))\setminus\{n-i\}|=d;\ |(\operatorname{Stag}(w))\setminus\{n-i\}|=s;\\ w_{n-i}>w_{n-i+1}}} x_{w_1}x_{w_2}\cdots x_{w_n}. \tag{13.88.20}$$

Now, let us make some more definitions:

- Define a power series $\mathcal{D}'_i \in \mathbf{k}[[\mathbf{x}]]$ as the sum of the monomials $x_{w_1}x_{w_2}\cdots x_{w_n}$ over all n-tuples $w=(w_1,w_2,...,w_n) \in \{1,2,3,\ldots\}^n$ satisfying $|\operatorname{Des}(w)| = d+1$, $|\operatorname{Stag}(w)| = s$, $w_1 < w_n$ and $w_i > w_{i+1}$.
- Define a power series $\mathcal{S}'_i \in \mathbf{k}[[\mathbf{x}]]$ as the sum of the monomials $x_{w_1}x_{w_2}\cdots x_{w_n}$ over all n-tuples $w=(w_1,w_2,...,w_n) \in \{1,2,3,\ldots\}^n$ satisfying $|\operatorname{Des}(w)| = d$, $|\operatorname{Stag}(w)| = s+1$, $w_1 < w_n$ and $w_i = w_{i+1}$.
- Define a power series $\mathcal{A}'_i \in \mathbf{k}[[\mathbf{x}]]$ as the sum of the monomials $x_{w_1}x_{w_2}\cdots x_{w_n}$ over all n-tuples $w=(w_1,w_2,...,w_n) \in \{1,2,3,\ldots\}^n$ satisfying $|\operatorname{Des}(w)| = d$, $|\operatorname{Stag}(w)| = s$, $w_1 < w_n$ and $w_i < w_{i+1}$.

These three power series $\mathcal{D}'_i$, $\mathcal{S}'_i$ and $\mathcal{A}'_i$ were defined in obvious analogy to the power series $\mathcal{D}_i$, $\mathcal{S}_i$ and $\mathcal{A}_i$ defined in our solution to Exercise 2.9.11(b) above. In the same way as we have proved (13.88.3) back there, we can see that

$$\mathcal{D}'_i + \mathcal{S}'_i + \mathcal{A}'_i = \sum_{\substack{w=(w_1,w_2,...,w_n)\in\{1,2,3,\ldots\}^n;\\ |(\operatorname{Des}(w))\setminus\{i\}|=d;\ |(\operatorname{Stag}(w))\setminus\{i\}|=s;\\ w_1<w_n}} x_{w_1}x_{w_2}\cdots x_{w_n}. \tag{13.88.21}$$

But the sum on the right hand side of (13.88.21) can be rewritten in terms of more familiar sums. In fact, the n-tuples $w=(w_1,w_2,...,w_n) \in \{1,2,3,\ldots\}^n$ are in bijection with the pairs (u,v) consisting of an i-tuple $u=(u_1,u_2,...,u_i) \in \{1,2,3,\ldots\}^i$ and an $(n-i)$-tuple $v=(v_1,v_2,...,v_{n-i}) \in \{1,2,3,\ldots\}^{n-i}$. This bijection sends an n-tuple $(w_1,w_2,...,w_n)$ to the pair $((w_1,w_2,...,w_i),(w_{i+1},w_{i+2},...,w_n))$, and has the property that $|(\operatorname{Des}(w)) \setminus \{i\}| = |\operatorname{Des}(u)| + |\operatorname{Des}(v)|$ [707], $|(\operatorname{Stag}(w)) \setminus \{i\}| = |\operatorname{Stag}(u)| + |\operatorname{Stag}(v)|$ [708],

[706] The idea is that the map J switches the number of p's with the number of $(p+1)$'s in any odd $(p,p+1)$-run, while the numbers in an even $(p,p+1)$-run are already equal to begin with.

[707] In fact, $(\operatorname{Des}(w))\setminus\{i\}$ is the union of the set $\operatorname{Des}(u)$ with the set $\operatorname{Des}(v)$ shifted by i (that is, the set $\{p+i \mid p \in \operatorname{Des}(v)\}$). This is a disjoint union, and thus we find that $|(\operatorname{Des}(w)) \setminus \{i\}| = |\operatorname{Des}(u)| + |\operatorname{Des}(v)|$ (since the set $\operatorname{Des}(v)$ shifted by i has cardinality $|\operatorname{Des}(v)|$).

[708] for similar reasons

$x_{w_1}x_{w_2}\cdots x_{w_n} = x_{u_1}x_{u_2}\cdots x_{u_i}x_{v_1}x_{v_2}\cdots x_{v_{n-i}}$, $w_1 = u_1$ and $w_n = v_{n-i}$. Hence,

$$\begin{aligned}
&\sum_{\substack{w=(w_1,w_2,\dots,w_n)\in\{1,2,3,\dots\}^n;\\ |(\operatorname{Des}(w))\setminus\{i\}|=d;\ |(\operatorname{Stag}(w))\setminus\{i\}|=s;\\ w_1<w_n}} x_{w_1}x_{w_2}\cdots x_{w_n}\\
&= \underbrace{\sum_{\substack{u=(u_1,u_2,\dots,u_i)\in\{1,2,3,\dots\}^i;\\ v=(v_1,v_2,\dots,v_{n-i})\in\{1,2,3,\dots\}^{n-i};\\ |\operatorname{Des}(u)|+|\operatorname{Des}(v)|=d;\\ |\operatorname{Stag}(u)|+|\operatorname{Stag}(v)|=s;\\ u_1<v_{n-i}}}}_{=\sum_{\substack{v=(v_1,v_2,\dots,v_{n-i})\in\{1,2,3,\dots\}^{n-i};\\ u=(u_1,u_2,\dots,u_i)\in\{1,2,3,\dots\}^i;\\ |\operatorname{Des}(u)|+|\operatorname{Des}(v)|=d;\\ |\operatorname{Stag}(u)|+|\operatorname{Stag}(v)|=s;\\ u_1<v_{n-i}}}} \underbrace{x_{u_1}x_{u_2}\cdots x_{u_i}x_{v_1}x_{v_2}\cdots x_{v_{n-i}}}_{=x_{v_1}x_{v_2}\cdots x_{v_{n-i}}x_{u_1}x_{u_2}\cdots x_{u_i}}\\
&= \sum_{\substack{v=(v_1,v_2,\dots,v_{n-i})\in\{1,2,3,\dots\}^{n-i};\\ u=(u_1,u_2,\dots,u_i)\in\{1,2,3,\dots\}^i;\\ |\operatorname{Des}(u)|+|\operatorname{Des}(v)|=d;\\ |\operatorname{Stag}(u)|+|\operatorname{Stag}(v)|=s;\\ u_1<v_{n-i}}} x_{v_1}x_{v_2}\cdots x_{v_{n-i}}x_{u_1}x_{u_2}\cdots x_{u_i}.
\end{aligned}\tag{13.88.22}$$

On the other hand, the pairs (v,u) consisting of an $(n-i)$-tuple $v=(v_1,v_2,...,v_{n-i})\in\{1,2,3,\ldots\}^{n-i}$ and an i-tuple $u=(u_1,u_2,...,u_i)\in\{1,2,3,\ldots\}^i$ are in bijection with the n-tuples $w=(w_1,w_2,...,w_n)\in\{1,2,3,\ldots\}^n$. This bijection sends a pair (v,u) with $v=(v_1,v_2,...,v_{n-i})$ and $u=(u_1,u_2,...,u_i)$ to the n-tuple $(v_1,v_2,...,v_{n-i},u_1,u_2,...,u_i)\in\{1,2,3,\ldots\}^n$, and has the property that $|(\operatorname{Des}(w))\setminus\{n-i\}| = |\operatorname{Des}(u)| + |\operatorname{Des}(v)|$ [709], $|(\operatorname{Stag}(w))\setminus\{n-i\}| = |\operatorname{Stag}(u)|+|\operatorname{Stag}(v)|$ [710], $x_{w_1}x_{w_2}\cdots x_{w_n} = x_{v_1}x_{v_2}\cdots x_{v_{n-i}}x_{u_1}x_{u_2}\cdots x_{u_i}$, $w_{n-i+1}=u_1$ and $w_{n-i}=v_{n-i}$. We can use this bijection to transform the sum on the right hand side of (13.88.22), and obtain

$$\begin{aligned}
&\sum_{\substack{v=(v_1,v_2,\dots,v_{n-i})\in\{1,2,3,\dots\}^{n-i};\\ u=(u_1,u_2,\dots,u_i)\in\{1,2,3,\dots\}^i;\\ |\operatorname{Des}(u)|+|\operatorname{Des}(v)|=d;\\ |\operatorname{Stag}(u)|+|\operatorname{Stag}(v)|=s;\\ u_1<v_{n-i}}} x_{u_1}x_{u_2}\cdots x_{u_i}x_{v_1}x_{v_2}\cdots x_{v_{n-i}}\\
&= \sum_{\substack{w=(w_1,w_2,\dots,w_n)\in\{1,2,3,\dots\}^n;\\ |(\operatorname{Des}(w))\setminus\{n-i\}|=d;\ |(\operatorname{Stag}(w))\setminus\{n-i\}|=s;\\ w_{n-i+1}<w_{n-i}}} x_{w_1}x_{w_2}\cdots x_{w_n}\\
&= \sum_{\substack{w=(w_1,w_2,\dots,w_n)\in\{1,2,3,\dots\}^n;\\ |(\operatorname{Des}(w))\setminus\{n-i\}|=d;\ |(\operatorname{Stag}(w))\setminus\{n-i\}|=s;\\ w_{n-i}>w_{n-i+1}}} x_{w_1}x_{w_2}\cdots x_{w_n} = \mathcal{D}_{n-i} \qquad (\text{by (13.88.20)}).
\end{aligned}\tag{13.88.23}$$

[709] In fact, $(\operatorname{Des}(w))\setminus\{n-i\}$ is the union of the set $\operatorname{Des}(v)$ with the set $\operatorname{Des}(u)$ shifted by $n-i$ (that is, the set $\{p+(n-i)\ \mid\ p\in\operatorname{Des}(u)\}$). This is a disjoint union; thus, we find $|(\operatorname{Des}(w))\setminus\{n-i\}| = |\operatorname{Des}(u)| + |\operatorname{Des}(v)|$ (since the set $\operatorname{Des}(u)$ shifted by $n-i$ has cardinality $|\operatorname{Des}(u)|$).

[710] for similar reasons

Now, (13.88.21) becomes

$$\begin{aligned}\mathcal{D}'_i + \mathcal{S}'_i + \mathcal{A}'_i &= \sum_{\substack{w=(w_1,w_2,\ldots,w_n)\in\{1,2,3,\ldots\}^n;\\ |(\operatorname{Des}(w))\setminus\{i\}|=d;\ |(\operatorname{Stag}(w))\setminus\{i\}|=s;\\ w_1<w_n}} x_{w_1}x_{w_2}\cdots x_{w_n} \\ &= \sum_{\substack{v=(v_1,v_2,\ldots,v_{n-i})\in\{1,2,3,\ldots\}^{n-i};\\ u=(u_1,u_2,\ldots,u_i)\in\{1,2,3,\ldots\}^{i};\\ |\operatorname{Des}(u)|+|\operatorname{Des}(v)|=d;\\ |\operatorname{Stag}(u)|+|\operatorname{Stag}(v)|=s;\\ u_1<v_{n-i}}} x_{v_1}x_{v_2}\cdots x_{v_{n-i}}x_{u_1}x_{u_2}\cdots x_{u_i} \qquad (\text{by } (13.88.22)) \\ &= \mathcal{D}_{n-i}. \end{aligned} \tag{13.88.24}$$

Now, let us forget that we fixed i. We thus have defined $\mathcal{D}'_i$, $\mathcal{S}'_i$ and $\mathcal{A}'_i$ and proven the equality (13.88.24) for all $i \in \{1,2,...,n-1\}$.

Now, let us take a closer look at the sums $\sum_{i=1}^{n-1}\mathcal{D}'_i$, $\sum_{i=1}^{n-1}\mathcal{S}'_i$ and $\sum_{i=1}^{n-1}\mathcal{A}'_i$:

- Similarly to how we proved (13.88.8), we can show that

$$\sum_{i=1}^{n-1}\mathcal{D}'_i = (d+1)\,U_{n,d+1,s}. \tag{13.88.25}$$

- Similarly to how we proved (13.88.9), we can see that

$$\sum_{i=1}^{n-1}\mathcal{S}'_i = (s+1)\,U_{n,d,s+1}. \tag{13.88.26}$$

- Similarly to how we proved (13.88.10), we can see that

$$\sum_{i=1}^{n-1}\mathcal{A}'_i = (n-1-d-s)\,U_{n,d,s}. \tag{13.88.27}$$

Now, summing the equality (13.88.24) over all $i \in \{1,2,...,n-1\}$ yields

$$\sum_{i=1}^{n-1}\left(\mathcal{D}'_i + \mathcal{S}'_i + \mathcal{A}'_i\right) = \sum_{i=1}^{n-1}\mathcal{D}_{n-i} = \sum_{i=1}^{n-1}\mathcal{D}_i = (d+1)\,X_{n,d+1,s} \qquad (\text{by } (13.88.8)).$$

Hence,

$$\begin{aligned}(d+1)\,X_{n,d+1,s} = \sum_{i=1}^{n-1}\left(\mathcal{D}'_i + \mathcal{S}'_i + \mathcal{A}'_i\right) &= \underbrace{\sum_{i=1}^{n-1}\mathcal{D}'_i}_{\substack{=(d+1)U_{n,d+1,s}\\ (\text{by } (13.88.25))}} + \underbrace{\sum_{i=1}^{n-1}\mathcal{S}'_i}_{\substack{=(s+1)U_{n,d,s+1}\\ (\text{by } (13.88.26))}} + \underbrace{\sum_{i=1}^{n-1}\mathcal{A}'_i}_{\substack{=(n-1-d-s)U_{n,d,s}\\ (\text{by } (13.88.27))}} \\ &= (d+1)\,U_{n,d+1,s} + (s+1)\,U_{n,d,s+1} + (n-1-d-s)\,U_{n,d,s}.\end{aligned}$$

This proves (13.88.19).

In solving Exercise 2.9.11(c), we WLOG assume that $\mathbf{k} = \mathbb{Z}$ (for the same reason why we could assume $\mathbf{k} = \mathbb{Z}$ in solving Exercise 2.9.11(b)).

Now, we are going to prove $U_{n,d,s} \in \Lambda$ by strong induction over $n+d$. So (for the induction step) we need to show that $U_{n,d,s} \in \Lambda$, and we can assume (as the induction hypothesis) that $U_{n',d',s'} \in \Lambda$ is already known to hold for any nonnegative integers n', d' and s' satisfying $n'+d' < n+d$.

We must be in one of the following two cases:

Case 1: We have $d > 0$.

Case 2: We have $d = 0$.

In Case 1, we can proceed in the same way as in the corresponding case of the solution of Exercise 2.9.11(b) (with the only difference that we now have to use $X_{n,d,s} \in \Lambda$, but this follows from the already solved Exercise 2.9.11(b)).

Let us now consider Case 2. In this case, we have $d = 0$. We now distinguish between two subcases:

Subcase 2.1: We have $s = n-1$.

Subcase 2.2: We have $s \neq n-1$.

Let us consider Subcase 2.1 first. In this subcase, we have $s = n-1$. We will show that $U_{n,d,s} = 0$ in this subcase.

Indeed, let $w = (w_1, w_2, \ldots, w_n) \in \{1,2,3,\ldots\}^n$ be such that $|\operatorname{Des}(w)| = d$, $|\operatorname{Stag}(w)| = s$ and $w_1 < w_n$. Then, $\operatorname{Stag}(w)$ is a subset of $\{1,2,...,n-1\}$ whose cardinality is $|\operatorname{Stag}(w)| = s = n-1 = |\{1,2,...,n-1\}|$. Obviously, the only such subset is $\{1,2,...,n-1\}$ itself, and so $\operatorname{Stag}(w)$ must be $\{1,2,...,n-1\}$. Hence, every $j \in \{1,2,...,n-1\}$ satisfies $j \in \{1,2,...,n-1\} = \operatorname{Stag}(w) = \{i \in \{1,2,\ldots,n-1\} : w_i = w_{i+1}\}$ and thus $w_j = w_{j+1}$. In other words, $w_1 = w_2 = ... = w_n$. This contradicts $w_1 < w_n$.

Now, forget that we fixed w. We thus have obtained a contradiction for every $w = (w_1, w_2, \ldots, w_n) \in \{1,2,3,\ldots\}^n$ satisfying $|\operatorname{Des}(w)| = d$, $|\operatorname{Stag}(w)| = s$ and $w_1 < w_n$. Therefore, there exists no such w. Hence, the sum on the right hand side of (2.9.10) is empty and thus equals 0. Thus, (2.9.10) rewrites as $U_{n,d,s} = 0 \in \Lambda$. Hence, the induction step is complete in Subcase 2.1.

Let us now consider Subcase 2.2. In this case, $s \neq n-1$. We will show that $U_{n,d,s} = X_{n,d,s}$ in this subcase.

Indeed, let $w = (w_1, w_2, \ldots, w_n) \in \{1,2,3,\ldots\}^n$ be such that $|\operatorname{Des}(w)| = d$ and $|\operatorname{Stag}(w)| = s$. We will show that $w_1 < w_n$.

We have $|\operatorname{Des}(w)| = d = 0$, so that the set $\operatorname{Des}(w)$ is empty. In other words, every $j \in \{1,2,...,n-1\}$ satisfies $w_j \leq w_{j+1}$. Hence, we have the chain of inequalities $w_1 \leq w_2 \leq ... \leq w_n$. At least one inequality in this chain must be strict (because otherwise, we would have $w_1 = w_2 = ... = w_n$, thus $w_j = w_{j+1}$ for every $j \in \{1,2,...,n-1\}$, thus $j \in \operatorname{Stag}(w)$ for every $j \in \{1,2,...,n-1\}$, which would lead to $\operatorname{Stag}(w) = \{1,2,...,n-1\}$ and thus $|\operatorname{Stag}(w)| = |\{1,2,...,n-1\}| = n-1$, in contradiction to $\operatorname{Stag}(w) = s \neq n-1$), and thus we have $w_1 < w_n$.

Now, let us forget that we fixed w. We thus have proven that every $w = (w_1, w_2, \ldots, w_n) \in \{1,2,3,\ldots\}^n$ satisfying $|\operatorname{Des}(w)| = d$ and $|\operatorname{Stag}(w)| = s$ automatically satisfies $w_1 < w_n$. Hence, the condition $w_1 < w_n$ under the summation sign "$\sum\limits_{\substack{w=(w_1,w_2,\ldots,w_n)\in\{1,2,3,\ldots\}^n;\\ |\operatorname{Des}(w)|=d;\ |\operatorname{Stag}(w)|=s;\\ w_1<w_n}}$" is redundant (i.e., can be removed without changing the range of the summation). Thus,

$$\sum_{\substack{w=(w_1,w_2,\ldots,w_n)\in\{1,2,3,\ldots\}^n;\\ |\operatorname{Des}(w)|=d;\ |\operatorname{Stag}(w)|=s;\\ w_1<w_n}} x_{w_1}x_{w_2}\cdots x_{w_n} = \sum_{\substack{w=(w_1,w_2,\ldots,w_n)\in\{1,2,3,\ldots\}^n;\\ |\operatorname{Des}(w)|=d;\ |\operatorname{Stag}(w)|=s}} x_{w_1}x_{w_2}\cdots x_{w_n} = X_{n,d,s}$$

(by (13.88.1)). Now, (2.9.10) becomes

$$U_{n,d,s} = \sum_{\substack{w=(w_1,w_2,\ldots,w_n)\in\{1,2,3,\ldots\}^n;\\ |\operatorname{Des}(w)|=d;\ |\operatorname{Stag}(w)|=s;\\ w_1<w_n}} x_{w_1}x_{w_2}\cdots x_{w_n} = X_{n,d,s} \in \Lambda$$

(by Exercise 2.9.11(b)). Hence, the induction step is complete in Subcase 2.2.

The induction step is thus complete in Case 1 and in each of the two Subcases 2.1 and 2.2. These are all cases, and so the induction step is finally complete. We have thus proven that

$$U_{n,d,s} \in \Lambda \qquad \text{for all positive integers } n, \text{ all } d \in \mathbb{N} \text{ and all } s \in \mathbb{N}. \tag{13.88.28}$$

In order to complete the solution of Exercise 2.9.11(c), we still need to prove that $V_{n,d,s}$ and $W_{n,d,s}$ belong to Λ for all positive integers n, all $d \in \mathbb{N}$ and all $s \in \mathbb{N}$.

Let n be a positive integer, let $d \in \mathbb{N}$ and let $s \in \mathbb{N}$. The equality (2.9.12) becomes

$$\begin{aligned}
W_{n,d,s} &= \sum_{\substack{w=(w_1,w_2,\ldots,w_n)\in\{1,2,3,\ldots\}^n;\\ |\operatorname{Des}(w)|=d;\ |\operatorname{Stag}(w)|=s;\\ w_1>w_n}} x_{w_1}x_{w_2}\cdots x_{w_n} \\
&= \sum_{\substack{(w_1,w_2,\ldots,w_n)\in\{1,2,3,\ldots\}^n;\\ |\operatorname{Des}((w_n,w_{n-1},\ldots,w_1))|=d;\ |\operatorname{Stag}((w_n,w_{n-1},\ldots,w_1))|=s;\\ w_n>w_1}} \underbrace{x_{w_n}x_{w_{n-1}}\cdots x_{w_1}}_{=x_{w_1}x_{w_2}\cdots x_{w_n}} \\
&\qquad \left(\text{here, we substituted } (w_n,w_{n-1},\ldots,w_1) \text{ for } w=(w_1,w_2,\ldots,w_n) \text{ in the sum}\right) \\
&= \sum_{\substack{(w_1,w_2,\ldots,w_n)\in\{1,2,3,\ldots\}^n;\\ |\operatorname{Des}((w_n,w_{n-1},\ldots,w_1))|=d;\ |\operatorname{Stag}((w_n,w_{n-1},\ldots,w_1))|=s;\\ w_n>w_1}} x_{w_1}x_{w_2}\cdots x_{w_n} \\
&= \sum_{\substack{w=(w_1,w_2,\ldots,w_n)\in\{1,2,3,\ldots\}^n;\\ |\operatorname{Des}((w_n,w_{n-1},\ldots,w_1))|=d;\ |\operatorname{Stag}((w_n,w_{n-1},\ldots,w_1))|=s;\\ w_n>w_1}} x_{w_1}x_{w_2}\cdots x_{w_n} \\
&= \sum_{\substack{w=(w_1,w_2,\ldots,w_n)\in\{1,2,3,\ldots\}^n;\\ n-1-|\operatorname{Des}(w)|-|\operatorname{Stag}(w)|=d;\ |\operatorname{Stag}(w)|=s;\\ w_n>w_1}} x_{w_1}x_{w_2}\cdots x_{w_n} \\
&\qquad \left(\begin{array}{c}\text{since every } w=(w_1,w_2,\ldots,w_n)\in\{1,2,3,\ldots\}^n \text{ satisfies}\\ |\operatorname{Des}((w_n,w_{n-1},\ldots,w_1))| = n-1-|\operatorname{Des}(w)|-|\operatorname{Stag}(w)|\\ \text{and } |\operatorname{Stag}((w_n,w_{n-1},\ldots,w_1))| = |\operatorname{Stag}(w)|\\ \text{(this is easily shown by observing that the set } \{1,2,...,n-1\}\\ \text{is the union of its three disjoint subsets}\\ \{i\in\{1,2,...,n-1\} \mid w_i<w_{i+1}\},\ \operatorname{Des}w \text{ and } \operatorname{Stag}w)\end{array}\right) \\
&= \sum_{\substack{w=(w_1,w_2,\ldots,w_n)\in\{1,2,3,\ldots\}^n;\\ |\operatorname{Des}(w)|=n-1-d-s;\ |\operatorname{Stag}(w)|=s;\\ w_1<w_n}} x_{w_1}x_{w_2}\cdots x_{w_n} \\
&\qquad \left(\begin{array}{c}\text{because for a } w=(w_1,w_2,\ldots,w_n)\in\{1,2,3,\ldots\}^n \text{ satisfying}\\ |\operatorname{Stag}(w)|=s, \text{ the assertions } n-1-|\operatorname{Des}(w)|-|\operatorname{Stag}(w)|=d\\ \text{and } w_n>w_1 \text{ are equivalent to } |\operatorname{Des}(w)|=n-1-d-s \text{ and } w_1<w_n\end{array}\right) \\
&= U_{n,n-1-d-s,s} \qquad \left(\text{by (2.9.10), applied to } n-1-d-s \text{ instead of } d\right) \\
&\in \Lambda \qquad \left(\text{by (13.88.28), applied to } n-1-d-s \text{ instead of } d\right).
\end{aligned}$$

Finally, (13.88.1) becomes

$$\begin{aligned}
X_{n,d,s} &= \sum_{\substack{w=(w_1,w_2,\ldots,w_n)\in\{1,2,3,\ldots\}^n;\\ |\operatorname{Des}(w)|=d;\ |\operatorname{Stag}(w)|=s}} x_{w_1}x_{w_2}\cdots x_{w_n} \\
&= \underbrace{\sum_{\substack{w=(w_1,w_2,\ldots,w_n)\in\{1,2,3,\ldots\}^n;\\ |\operatorname{Des}(w)|=d;\ |\operatorname{Stag}(w)|=s;\\ w_1<w_n}} x_{w_1}x_{w_2}\cdots x_{w_n}}_{\substack{=U_{n,d,s}\\ \text{(by (2.9.10))}}} + \underbrace{\sum_{\substack{w=(w_1,w_2,\ldots,w_n)\in\{1,2,3,\ldots\}^n;\\ |\operatorname{Des}(w)|=d;\ |\operatorname{Stag}(w)|=s;\\ w_1=w_n}} x_{w_1}x_{w_2}\cdots x_{w_n}}_{\substack{=V_{n,d,s}\\ \text{(by (2.9.11))}}} \\
&\qquad + \underbrace{\sum_{\substack{w=(w_1,w_2,\ldots,w_n)\in\{1,2,3,\ldots\}^n;\\ |\operatorname{Des}(w)|=d;\ |\operatorname{Stag}(w)|=s;\\ w_1>w_n}} x_{w_1}x_{w_2}\cdots x_{w_n}}_{\substack{=W_{n,d,s}\\ \text{(by (2.9.12))}}} \\
&= U_{n,d,s} + V_{n,d,s} + W_{n,d,s},
\end{aligned}$$

so that $V_{n,d,s} = \underbrace{X_{n,d,s}}_{\substack{\in\Lambda \\ \text{(by Exercise 2.9.11(b))}}} - \underbrace{U_{n,d,s}}_{\in\Lambda} - \underbrace{W_{n,d,s}}_{\in\Lambda} \in \Lambda - \Lambda - \Lambda \subset \Lambda$.

We have thus shown that $V_{n,d,s} \in \Lambda$ and $W_{n,d,s} \in \Lambda$. Combined with $U_{n,d,s} \in \Lambda$ (this follows from (13.88.28)), this completes the solution of Exercise 2.9.11(c).

13.89. **Solution to Exercise 2.9.13.** *Solution to Exercise 2.9.13.* We start out with two definitions:

- If m and q are integers satisfying $0 \leq q \leq m$, and if $U = (u_{i,j})_{i,j=1,2,...,m} \in \mathbf{k}^{m\times m}$ is an $m \times m$-matrix, then $\operatorname{NWsm}_q U$ will mean the matrix $(u_{i,j})_{i,j=1,2,...,q} \in \mathbf{k}^{q\times q}$. This is the submatrix of U obtained by removing all rows other than the first q rows and then removing all columns other than the first q columns.[711]
- A square matrix $(u_{i,j})_{i,j=1,2,...,m} \in \mathbf{k}^{m\times m}$ is said to be *nearly lower-triangular* if we have

$$\left(u_{i,j} = 0 \text{ for every } (i,j) \in \{1,2,...,m\}^2 \text{ satisfying } j > i+1\right).$$

  (Thus, informally, a square matrix is nearly lower-triangular if and only if all its entries above the superdiagonal are 0, where the *superdiagonal* is the set of all cells which lie just one step north of a cell on the diagonal.)

We will now show a lemma:

Lemma 13.89.1. *Let $m \in \mathbb{N}$. Let $U = (u_{i,j})_{i,j=1,2,...,m} \in \mathbf{k}^{m\times m}$ be a nearly lower-triangular $m \times m$-matrix. Then,*

$$\det U = \sum_{r=1}^{m} (-1)^{m-r} u_{m,r} \det\left(\operatorname{NWsm}_{r-1} U\right) \cdot \prod_{k=r}^{m-1} u_{k,k+1}.$$

Proof of Lemma 13.89.1. Fix some $r \in \{1,2,...,m\}$. We define the following four matrices:

- the $(r-1)\times(r-1)$-matrix $P = (u_{i,j})_{i,j=1,2,...,r-1}$, which is obtained from the matrix U by removing all rows other than the first $r-1$ rows and then removing all columns other than the first $r-1$ columns;
- the $(r-1)\times(m-r)$-matrix $Q = (u_{i,r+j})_{i=1,2,...,r-1;\ j=1,2,...,m-r}$, which is obtained from the matrix U by removing all rows other than the first $r-1$ rows and then removing all columns other than the last $m-r$ columns;
- the $(m-r)\times(r-1)$-matrix $R = (u_{r-1+i,j})_{i=1,2,...,m-r;\ j=1,2,...,r-1}$, which is obtained from the matrix U by removing all rows other than the r-th, the $(r+1)$-st, etc., the $(m-1)$-st row, and then removing all columns other than the first $r-1$ columns;
- the $(m-r)\times(m-r)$-matrix $S = (u_{r-1+i,r+j})_{i,j=1,2,...,m-r}$, which is obtained from the matrix U by removing all rows other than the r-th, the $(r+1)$-st, etc., the $(m-1)$-st row, and then removing all columns other than the last $m-r$ columns.

Now, the matrix U can be written as a block matrix as follows: $U = \begin{pmatrix} P & v & Q \\ R & w & S \\ x & y & z \end{pmatrix}$, where $\begin{pmatrix} v \\ w \\ y \end{pmatrix}$ is the r-th column of U, and where $\begin{pmatrix} x & y & z \end{pmatrix}$ is the m-th row of U. [712] Hence,

$$\begin{aligned} &(\text{the matrix obtained from } U \text{ by removing the } m\text{-th row and the } r\text{-th column}) \\ &= \begin{pmatrix} P & Q \\ R & S \end{pmatrix}. \end{aligned} \tag{13.89.1}$$

[711] The notation $\operatorname{NWsm}_q U$ stands short for "q-th northwest submatrix of U". It is the kind of submatrices whose determinants usually figure in the Sylvester criterion for the positive definiteness of a matrix.

[712] Despite being labelled with lowercase letters, v, w, x, y and z are still blocks, although each has (at least) one of its dimensions equal to 1 (and y is a 1×1-block).

But the matrix $Q = (u_{i,r+j})_{i=1,2,\ldots,r-1;\ j=1,2,\ldots,m-r}$ is the zero matrix (since U is nearly lower-triangular), and the matrices P and S are square matrices. Hence, $\begin{pmatrix} P & Q \\ R & S \end{pmatrix}$ is a block lower-triangular matrix with diagonal blocks P and S. Thus, its determinant is

$$\det \begin{pmatrix} P & Q \\ R & S \end{pmatrix} = \det P \cdot \det S \tag{13.89.2}$$

(since the determinant of a block lower-triangular matrix is known to equal the product of the determinants of its diagonal blocks).

But the matrix $S = (u_{r-1+i,r+j})_{i,j=1,2,\ldots,m-r}$ is lower-triangular (since U is nearly lower-triangular). Since it is well-known that the determinant of a lower-triangular matrix equals the product of its diagonal entries, we can therefore compute the determinant $\det S$ of S as follows:

$$\det S = \prod_{i=1}^{m-r} u_{r-1+i,r+i} = \prod_{k=r}^{m-1} u_{k,k+1} \qquad \text{(here, we have substituted } k \text{ for } r-1+i \text{ in the product)}. \tag{13.89.3}$$

Moreover, $P = \operatorname{NWsm}_{r-1} U$ (since comparing the definitions of P and of $\operatorname{NWsm}_{r-1} U$ shows that these two matrices are the same). Now, applying the map det to both sides of (13.89.1), we obtain

$$\begin{aligned} &\det (\text{the matrix obtained from } U \text{ by removing the } m\text{-th row and the } r\text{-th column}) \\ &= \det \begin{pmatrix} P & Q \\ R & S \end{pmatrix} = \det \underbrace{P}_{= \operatorname{NWsm}_{r-1} U} \cdot \underbrace{\det S}_{\substack{= \prod_{k=r}^{m-1} u_{k,k+1} \\ \text{(by (13.89.3))}}} \qquad \text{(by (13.89.2))} \\ &= \det (\operatorname{NWsm}_{r-1} U) \cdot \prod_{k=r}^{m-1} u_{k,k+1}. \end{aligned} \tag{13.89.4}$$

Now, forget that we fixed r. We can compute the determinant $\det U$ by Laplace expansion along the m-th row, thus obtaining

$$\begin{aligned} &\det U \\ &= \sum_{r=1}^{m} u_{m,r} \cdot \underbrace{(\text{the } (m,r)\text{-th cofactor of the matrix } U)}_{= (-1)^{m+r} \det(\text{the matrix obtained from } U \text{ by removing the } m\text{-th row and the } r\text{-th column})} \\ &= \sum_{r=1}^{m} u_{m,r} \cdot \underbrace{(-1)^{m+r}}_{=(-1)^{m-r}} \underbrace{\det (\text{the matrix obtained from } U \text{ by removing the } m\text{-th row and the } r\text{-th column})}_{\substack{= \det(\operatorname{NWsm}_{r-1} U) \cdot \prod_{k=r}^{m-1} u_{k,k+1} \\ \text{(by (13.89.4))}}} \\ &= \sum_{r=1}^{m} u_{m,r} \cdot (-1)^{m-r} \cdot \det (\operatorname{NWsm}_{r-1} U) \cdot \prod_{k=r}^{m-1} u_{k,k+1} = \sum_{r=1}^{m} (-1)^{m-r} u_{m,r} \det (\operatorname{NWsm}_{r-1} U) \cdot \prod_{k=r}^{m-1} u_{k,k+1}. \end{aligned}$$

This proves Lemma 13.89.1. □

Before we solve the actual exercise, we record one further identity that we will be using twice. Namely, we claim that

$$m e_m = \sum_{i=1}^{m} (-1)^{i-1} e_{m-i} p_i \qquad \text{for every } m \in \mathbb{N}. \tag{13.89.5}$$

Proof of (13.89.5): Recall the power series $H(t)$ defined in (2.4.1), and the power series $E(t)$ defined in (2.4.2). From (2.4.3), we know that $E(-t) H(t) = 1$. Differentiating both sides of this equation with respect to t, we obtain $(E(-t) H(t))' = 1' = 0$, so that

$$\begin{aligned} 0 = (E(-t) H(t))' &= \underbrace{(E(-t))'}_{=-E'(-t)} H(t) + E(-t) H'(t) \qquad \text{(by the Leibniz rule)} \\ &= -E'(-t) H(t) + E(-t) H'(t), \end{aligned}$$

and thus $E'(-t)H(t) = E(-t)H'(t)$. Hence, $\dfrac{E'(-t)}{E(-t)} = \dfrac{H'(t)}{H(t)}$.

But Exercise 2.5.21 yields $\sum_{m\geq 0} p_{m+1}t^m = \dfrac{H'(t)}{H(t)}$. Compared with $\dfrac{E'(-t)}{E(-t)} = \dfrac{H'(t)}{H(t)}$, this yields $\dfrac{E'(-t)}{E(-t)} = \sum_{m\geq 0} p_{m+1}t^m$. Substituting $-t$ for t in this equality, we obtain

$$\frac{E'(t)}{E(t)} = \sum_{m\geq 0} p_{m+1}(-t)^m = \sum_{m\geq 0} p_{m+1}(-1)^m t^m.$$

Thus,

$$\begin{aligned}
E'(t) &= \underbrace{E(t)}_{\substack{=\sum_{m\geq 0} e_m t^m \\ \text{(by the definition of } E(t))}} \cdot \left(\sum_{m\geq 0} p_{m+1}(-1)^m t^m\right) = \left(\sum_{m\geq 0} e_m t^m\right)\left(\sum_{m\geq 0} p_{m+1}(-1)^m t^m\right) \\
&= \left(\sum_{m\geq 0} p_{m+1}(-1)^m t^m\right)\left(\sum_{m\geq 0} e_m t^m\right) \\
&= \sum_{m\geq 0}\left(\sum_{i=0}^{m} p_{i+1}(-1)^i e_{m-i}\right) t^m \qquad \text{(by the definition of the product of two power series)} \\
&= \sum_{m\geq 1}\left(\sum_{i=0}^{m-1} p_{i+1}(-1)^i e_{m-1-i}\right) t^{m-1} \qquad \text{(here, we substituted } m-1 \text{ for } m \text{ in the first sum)} \\
&= \sum_{m\geq 0} \underbrace{\left(\sum_{i=0}^{m-1} p_{i+1}(-1)^i e_{m-1-i}\right)}_{\substack{=\sum_{i=1}^{m} p_i(-1)^{i-1} e_{m-i} \\ \text{(here we substituted } i-1 \text{ for } i \text{ in the sum)}}} t^{m-1} \\
&\qquad \left(\begin{array}{c}\text{here, we have added an } m=0 \text{ addend to the first sum;} \\ \text{this did not change the sum since this addend is } 0\end{array}\right) \\
&= \sum_{m\geq 0}\left(\sum_{i=1}^{m} p_i(-1)^{i-1} e_{m-i}\right) t^{m-1} = \sum_{m\geq 0}\left(\sum_{i=1}^{m} (-1)^{i-1} e_{m-i} p_i\right) t^{m-1}.
\end{aligned}$$

Compared with

$$\begin{aligned}
E'(t) &= \left(\sum_{m\geq 0} e_m t^m\right)' \qquad \left(\text{since } E(t) = \sum_{m\geq 0} e_m t^m\right) \\
&= \sum_{m\geq 0} e_m \cdot m t^{m-1} = \sum_{m\geq 0} m e_m t^{m-1},
\end{aligned}$$

this yields

$$\sum_{m\geq 0} m e_m t^{m-1} = \sum_{m\geq 0}\left(\sum_{i=1}^{m} (-1)^{i-1} e_{m-i} p_i\right) t^{m-1}.$$

Multiplying both sides of this equality by t, we obtain

$$\sum_{m\geq 0} m e_m t^m = \sum_{m\geq 0}\left(\sum_{i=1}^{m} (-1)^{i-1} e_{m-i} p_i\right) t^m.$$

Comparing coefficients in this equality of power series, we conclude that every $m \in \mathbb{N}$ satisfies

$$m e_m = \sum_{i=1}^{m} (-1)^{i-1} e_{m-i} p_i.$$

This proves (13.89.5).

Now, let us solve the exercise.

(a) We will solve Exercise 2.9.13(a) by strong induction over n. Thus, we assume (as the induction hypothesis) that

$$\det\left(A_k\right) = k!e_k \qquad \text{for all } k \in \mathbb{N} \text{ satisfying } k < n. \tag{13.89.6}$$

We now need to prove that $\det\left(A_n\right) = n!e_n$.

The matrix $A_n = \left(a_{i,j}\right)_{i,j=1,2,...,n}$ is nearly lower-triangular[713]. Hence, Lemma 13.89.1 (applied to $m = n$, $U = A_n$ and $u_{i,j} = a_{i,j}$) yields

$$\begin{aligned}
\det\left(A_n\right) &= \sum_{r=1}^{n} (-1)^{n-r} \underbrace{a_{n,r}}_{\substack{=p_{n-r+1} \\ \text{(by the definition of } a_{n,r}, \\ \text{since } n \geq r)}} \det \underbrace{\left(\operatorname{NWsm}_{r-1}\left(A_n\right)\right)}_{\substack{=A_{r-1} \\ \text{(this is easy to see by} \\ \text{the definitions of } A_n \text{ and } A_{r-1})}} \cdot \prod_{k=r}^{n-1} \underbrace{a_{k,k+1}}_{\substack{=k \\ \text{(by the definition of } a_{k,k+1}, \\ \text{since } k=(k+1)-1)}} \\
&= \sum_{r=1}^{n} (-1)^{n-r} p_{n-r+1} \underbrace{\det\left(A_{r-1}\right)}_{\substack{=(r-1)!e_{r-1} \\ \text{(by (13.89.6), applied to } k=r-1)}} \cdot \prod_{k=r}^{n-1} k \\
&= \sum_{r=1}^{n} (-1)^{n-r} p_{n-r+1} (r-1)!e_{r-1} \prod_{k=r}^{n-1} k = \sum_{r=1}^{n} (-1)^{n-r} \underbrace{(r-1)! \left(\prod_{k=r}^{n-1} k\right)}_{=(n-1)!} p_{n-r+1} e_{r-1} \\
&= (n-1)! \sum_{r=1}^{n} (-1)^{n-r} p_{n-r+1} e_{r-1} = (n-1)! \underbrace{\sum_{i=1}^{n} (-1)^{i-1} p_i e_{n-i}}_{\substack{=\sum_{i=1}^{n} (-1)^{i-1} e_{n-i} p_i = n e_n \\ \text{(because (13.89.5) (applied to } m=n) \\ \text{yields } n e_n = \sum_{i=1}^{n} (-1)^{i-1} e_{n-i} p_i)}} \\
&\qquad \text{(here, we have substituted } n-i+1 \text{ for } r \text{ in the sum)} \\
&= \underbrace{(n-1)!n}_{=n!} e_n = n!e_n.
\end{aligned}$$

This completes the induction step, and so Exercise 2.9.13(a) is solved.

(b) We will solve Exercise 2.9.13(b) by strong induction over n. Thus, we assume (as the induction hypothesis) that

$$\det\left(B_k\right) = p_k \qquad \text{for all positive integers } k \text{ satisfying } k < n. \tag{13.89.7}$$

We now need to prove that $\det\left(B_n\right) = p_n$.

[713] because for every $(i,j) \in \{1,2,...,n\}^2$ satisfying $j > i+1$, we have

$$a_{i,j} = \begin{cases} p_{i-j+1}, & \text{if } i \geq j; \\ i, & \text{if } i = j-1; \\ 0, & \text{if } i < j-1 \end{cases} = 0 \qquad (\text{since } i < j-1 \text{ (because } j > i+1)).$$

The matrix $B_n = (b_{i,j})_{i,j=1,2,\ldots,n}$ is nearly lower-triangular[714]. Hence, Lemma 13.89.1 (applied to $m = n$, $U = B_n$ and $u_{i,j} = b_{i,j}$) yields

$$\det(B_n) = \sum_{r=1}^{n} (-1)^{n-r} b_{n,r} \det \underbrace{(\mathrm{NWsm}_{r-1}(B_n))}_{\substack{=B_{r-1} \\ \text{(this is easy to see by} \\ \text{the definitions of } B_n \text{ and } B_{r-1})}} \cdot \prod_{k=r}^{n-1} \underbrace{b_{k,k+1}}_{\substack{=e_{k-(k+1)+1} \\ \text{(by the definition of } b_{k,k+1}, \\ \text{since } k+1>k\geq r\geq 1)}}$$

$$= \sum_{r=1}^{n} (-1)^{n-r} b_{n,r} \det(B_{r-1}) \cdot \prod_{k=r}^{n-1} \underbrace{e_{k-(k+1)+1}}_{=e_0=1}$$

$$= \sum_{r=1}^{n} (-1)^{n-r} b_{n,r} \det(B_{r-1}) \cdot \underbrace{\prod_{k=r}^{n-1} 1}_{=1} = \sum_{r=1}^{n} (-1)^{n-r} b_{n,r} \det(B_{r-1})$$

$$= (-1)^{n-1} \underbrace{b_{n,1}}_{\substack{=ne_n \\ \text{(by the definition} \\ \text{of } b_{n,1})}} \underbrace{\det(B_{1-1})}_{\substack{=1 \\ \text{(since } B_{1-1} \text{ is a} \\ 0\times 0\text{-matrix)}}} + \sum_{r=2}^{n} (-1)^{n-r} \underbrace{b_{n,r}}_{\substack{=e_{n-r+1} \\ \text{(by the definition of } b_{n,r}, \\ \text{since } r>1)}} \underbrace{\det(B_{r-1})}_{\substack{=p_{r-1} \\ \text{(by (13.89.7),} \\ \text{applied to } k=r-1)}}$$

$$= (-1)^{n-1} \underbrace{ne_n}_{\substack{=\sum_{i=1}^{n}(-1)^{i-1}e_{n-i}p_i \\ \text{(by (13.89.5), applied} \\ \text{to } m=n)}} + \underbrace{\sum_{r=2}^{n} (-1)^{n-r} e_{n-r+1} p_{r-1}}_{\substack{=\sum_{i=1}^{n-1}(-1)^{n-(i+1)}e_{n-i}p_i \\ \text{(here, we substituted } i+1 \text{ for } r \text{ in the sum)}}}$$

$$= (-1)^{n-1} \sum_{i=1}^{n} (-1)^{i-1} e_{n-i} p_i + \sum_{i=1}^{n-1} (-1)^{n-(i+1)} e_{n-i} p_i$$

$$= \sum_{i=1}^{n} \underbrace{(-1)^{n-1}(-1)^{i-1}}_{=(-1)^{n-i}} e_{n-i} p_i + \sum_{i=1}^{n-1} \underbrace{(-1)^{n-(i+1)}}_{=-(-1)^{n-i}} e_{n-i} p_i$$

$$= \underbrace{\sum_{i=1}^{n} (-1)^{n-i} e_{n-i} p_i}_{=\sum_{i=1}^{n-1}(-1)^{n-i}e_{n-i}p_i + (-1)^{n-n}e_{n-n}p_n} - \sum_{i=1}^{n-1} (-1)^{n-i} e_{n-i} p_i$$

$$= \sum_{i=1}^{n-1} (-1)^{n-i} e_{n-i} p_i + (-1)^{n-n} e_{n-n} p_n - \sum_{i=1}^{n-1} (-1)^{n-i} e_{n-i} p_i = \underbrace{(-1)^{n-n}}_{=1} \underbrace{e_{n-n}}_{=e_0=1} p_n = p_n.$$

This completes the induction step, and so Exercise 2.9.13(b) is solved.

Remark: The above solution follows closely the solution of Exercise 9.3 in the first author's "λ-rings: Definitions and basic properties" (`https://github.com/darijgr/lambda` , version 0.0.21), which is more or less a restatement of this exercise (since the elements of $\Lambda_{\mathbb{Z}}$ are in a 1-to-1 correspondence with unary functorial operations defined on every λ-ring).

Whenever $\ell \in \mathbb{N}$ and two partitions λ and μ of length $\leq \ell$ have the property that the transpose of the matrix $\left(h_{\lambda_i - \mu_j - i + j}\right)_{i,j=1,2,\ldots,\ell}$ is nearly lower-triangular, we can use Lemma 13.89.1 to obtain a recursive formula for the determinant of this matrix, which is $s_{\lambda/\mu}$ according to (2.4.16). This does not seem to be of much use, however.

[714] because for every $(i,j) \in \{1,2,\ldots,n\}^2$ satisfying $j > i+1$, we have

$$b_{i,j} = \begin{cases} ie_i, & \text{if } j = 1; \\ e_{i-j+1}, & \text{if } j > 1 \end{cases} = e_{i-j+1} \qquad (\text{since } j > 1 \text{ (because } j > i+1 \geq 1))$$

$$= 0 \qquad (\text{since } i-j+1 < 0 \text{ (since } j > i+1))$$

13.90. **Solution to Exercise 2.9.14.** *Solution to Exercise 2.9.14.* For any integers a and b, we define an element $\mathbf{s}(a,b)$ of Λ by

$$\mathbf{s}(a,b) = \begin{cases} s_{(a+1,1^b)}, & \text{if } a \geq 0 \text{ and } b \geq 0; \\ (-1)^b \delta_{a+b,-1}, & \text{if } a < 0 \text{ and } b \geq 0; \\ 0, & \text{if } b < 0 \end{cases} .$$

(We are introducing this $\mathbf{s}(a,b)$ chiefly for reasons of convenience: it will allow us to unify parts (b) and (c) of the exercise.)

Using the definition of $\mathbf{s}(a,b)$ and straightforward case analysis, we can see that:

$$\mathbf{s}(a,b) = s_{(a+1,1^b)} \qquad \text{for any } a \in \mathbb{N} \text{ and } b \in \mathbb{N}; \tag{13.90.1}$$

$$\mathbf{s}(a,b) = (-1)^b \delta_{a+b,-1} \qquad \text{for every negative integer } a \text{ and every } b \in \mathbb{Z}; \tag{13.90.2}$$

$$\mathbf{s}(a,0) = h_{a+1} \qquad \text{for every } a \in \mathbb{Z}; \tag{13.90.3}$$

$$\mathbf{s}(a,b) = 0 \qquad \text{for any } a \in \mathbb{Z} \text{ and any negative } b \in \mathbb{Z}. \tag{13.90.4}$$

Let us now show that

$$e_n h_m = \mathbf{s}(m, n-1) + \mathbf{s}(m-1, n) \qquad \text{for every } n \in \mathbb{Z} \text{ and every } m \in \mathbb{Z}. \tag{13.90.5}$$

Proof of (13.90.5): Let $n \in \mathbb{Z}$ and $m \in \mathbb{Z}$. It is easy to prove (13.90.5) in the case $n < 0$ (since both sides of (13.90.5) vanish in this case) and in the case $n = 0$ (here, it follows from (13.90.3)). We can therefore WLOG assume that $n > 0$. Assume this.

It is easy to prove (13.90.5) in the case $m < 0$ (in which case both sides of (13.90.5) vanish) and in the case $m = 0$ (in which case both sides of (13.90.5) equal e_n). Thus, we can WLOG assume that $m > 0$. Assume this. Since $m > 0$, we have $h_m = s_{(m)}$, so that

$$e_n \underbrace{h_m}_{=s_{(m)}} = e_n s_{(m)} = s_{(m)} e_n = \sum_{\substack{\lambda^+ : \lambda^+/(m) \text{ is a} \\ \text{vertical } n\text{-strip}}} s_{\lambda^+} \qquad (\text{by (2.7.2), applied to } \lambda = (m)). \tag{13.90.6}$$

It is easy to see that there are exactly two partitions λ^+ for which $\lambda^+/(m)$ is a vertical n-strip, namely the partitions $(m, 1^n)$ and $(m+1, 1^{n-1})$. Hence, the sum $\sum_{\substack{\lambda^+ : \lambda^+/(m) \text{ is a} \\ \text{vertical } n\text{-strip}}} s_{\lambda^+}$ has exactly two addends: that for $\lambda^+ = (m, 1^n)$ and that for $\lambda^+ = (m+1, 1^{n-1})$. Thus, this sum simplifies to $s_{(m,1^n)} + s_{(m+1,1^{n-1})}$. Hence, (13.90.6) rewrites as

$$e_n h_m = s_{(m,1^n)} + s_{(m+1,1^{n-1})}. \tag{13.90.7}$$

But (13.90.1) (applied to $a = m-1$ and $b = n$) yields $\mathbf{s}(m-1,n) = s_{(m,1^n)}$. Also, (13.90.1) (applied to $a = m$ and $b = n-1$) yields $\mathbf{s}(m, n-1) = s_{(m+1,1^{n-1})}$. Thus,

$$e_n h_m = \underbrace{s_{(m,1^n)}}_{=\mathbf{s}(m-1,n)} + \underbrace{s_{(m+1,1^{n-1})}}_{=\mathbf{s}(m,n-1)} = \mathbf{s}(m-1,n) + \mathbf{s}(m,n-1) = \mathbf{s}(m,n-1) + \mathbf{s}(m-1,n).$$

This proves (13.90.5).

Let us now prove a statement which encompasses both parts (b) and (c) of Exercise 2.9.14. Namely, we are going to prove that

$$\sum_{i=0}^{b} (-1)^i h_{a+i+1} e_{b-i} = \mathbf{s}(a,b) \qquad \text{for any } a \in \mathbb{Z} \text{ and } b \in \mathbb{Z}. \tag{13.90.8}$$

Proof of (13.90.8): Let $a \in \mathbb{Z}$ and $b \in \mathbb{Z}$. We WLOG assume that $b \in \mathbb{N}$ (since otherwise, the left hand side of (13.90.8) is an empty sum, and the right hand side is 0 by definition). For every $i \in \{0, 1, ..., b\}$, we have

$$h_{a+i+1} e_{b-i} = e_{b-i} h_{a+i+1} = \mathbf{s}\left(\underbrace{a+i+1}_{=a+(i+1)}, \underbrace{b-i-1}_{=b-(i+1)}\right) + \mathbf{s}\left(\underbrace{(a+i+1)-1}_{=a+i}, b-i\right)$$
$$\text{(by (13.90.5), applied to } n = b-i \text{ and } m = a+i+1\text{)}$$
$$= \mathbf{s}\left(a+(i+1), b-(i+1)\right) + \mathbf{s}\left(a+i, b-i\right). \tag{13.90.9}$$

Now,

$$\sum_{i=0}^{b} (-1)^i \underbrace{h_{a+i+1} e_{b-i}}_{\substack{=\mathbf{s}(a+(i+1),b-(i+1))+\mathbf{s}(a+i,b-i) \\ \text{(by (13.90.9))}}} = \sum_{i=0}^{b} \underbrace{(-1)^i \left(\mathbf{s}\left(a+(i+1), b-(i+1)\right) + \mathbf{s}\left(a+i, b-i\right)\right)}_{\substack{=(-1)^i \mathbf{s}(a+(i+1),b-(i+1))+(-1)^i \mathbf{s}(a+i,b-i) \\ =(-1)^i \mathbf{s}(a+(i+1),b-(i+1))-(-1)^{i-1} \mathbf{s}(a+i,b-i)}}$$
$$= \sum_{i=0}^{b} \left((-1)^i \mathbf{s}\left(a+(i+1), b-(i+1)\right) - (-1)^{i-1} \mathbf{s}\left(a+i, b-i\right)\right)$$
$$= (-1)^b \mathbf{s}\left(a+(b+1), \underbrace{b-(b+1)}_{=-1}\right) - \underbrace{(-1)^{0-1}}_{=-1} \mathbf{s}\left(\underbrace{a+0}_{=a}, \underbrace{b-0}_{=b}\right)$$
$$\text{(by the telescope principle)}$$
$$= (-1)^b \underbrace{\mathbf{s}(a,-1)}_{\substack{=0 \\ \text{(by (13.90.4))}}} - (-1)\mathbf{s}(a,b) = -(-1)\mathbf{s}(a,b) = \mathbf{s}(a,b).$$

This proves (13.90.8).

Here is a modified version of (13.90.8) which will be used in our solution of Exercise 2.9.14(d) further below: For any $a \in \mathbb{Z}$, $b \in \mathbb{Z}$ and $c \in \mathbb{Z}$ satisfying $c \geq b$, we have

$$\sum_{k=0}^{c} (-1)^k h_{a+k+1} e_{b-k} = \mathbf{s}(a,b). \tag{13.90.10}$$

[715]

Now, the first three parts of Exercise 2.9.14 are as good as solved: The claim of Exercise 2.9.14(a) has already been proven in (13.90.7), and the claims of Exercise 2.9.14(b) and Exercise 2.9.14(c) follow from (13.90.8).

Before we come to the solution of Exercise 2.9.14(d), we simplify our life by introducing another definition. Namely, let us define a $\mathbf{k}$-module endomorphism $\overline{\operatorname{id}}$ of Λ by $\overline{\operatorname{id}} = \operatorname{id}_\Lambda - u\epsilon$. Then, it is easy to see that

[715] *Proof of (13.90.10):* Let $a \in \mathbb{Z}$, $b \in \mathbb{Z}$ and $c \in \mathbb{Z}$ be such that $c \geq b$. We WLOG assume that $b \geq 0$ (since otherwise, both sides of (13.90.10) vanish). Then,

$$\sum_{k=0}^{c} (-1)^k h_{a+k+1} e_{b-k} = \sum_{k=0}^{b} (-1)^k h_{a+k+1} e_{b-k} + \sum_{k=b+1}^{c} (-1)^k h_{a+k+1} \underbrace{e_{b-k}}_{\substack{=0 \\ \text{(since } b-k<0 \\ \text{(since } k>b))}}$$
$$= \sum_{k=0}^{b} (-1)^k h_{a+k+1} e_{b-k} + \underbrace{\sum_{k=b+1}^{c} (-1)^k h_{a+k+1} 0}_{=0}$$
$$= \sum_{k=0}^{b} (-1)^k h_{a+k+1} e_{b-k} = \sum_{i=0}^{b} (-1)^i h_{a+i+1} e_{b-i} = \mathbf{s}(a,b) \qquad \text{(by (13.90.8))}.$$

This proves (13.90.10).

$\overline{\operatorname{id}}(s_\lambda) = s_\lambda$ for every nonempty partition λ, whereas $\overline{\operatorname{id}}(c) = 0$ for every $c \in \mathbf{k}$. Now, straightforward computation shows that

$$\overline{\operatorname{id}}(\mathbf{s}(a,b)) = s_{(a+1,1^b)} \qquad \text{for all } a \in \mathbb{N} \text{ and } b \in \mathbb{N}; \tag{13.90.11}$$

$$\overline{\operatorname{id}}(\mathbf{s}(a,b)) = 0 \qquad \text{for every } (a,b) \in \mathbb{Z}^2 \text{ satisfying } (a,b) \notin \mathbb{N}^2. \tag{13.90.12}$$

But every $x \in \Lambda$ satisfies

$$\Delta x - 1 \otimes x - x \otimes 1 = \left(\overline{\operatorname{id}} \otimes \overline{\operatorname{id}}\right)(\Delta x) - \epsilon(x) \cdot 1 \otimes 1 \qquad \text{in } \Lambda \otimes \Lambda. \tag{13.90.13}$$

(This is an identity which holds not only in Λ but in every $\mathbf{k}$-bialgebra, and which is proven by applying the bialgebra axioms.)

(d) Recall that Λ is a $\mathbf{k}$-bialgebra. Hence, $\Delta : \Lambda \to \Lambda \otimes \Lambda$ is a $\mathbf{k}$-algebra homomorphism.

We can rewrite Proposition 2.3.6(ii) as follows: Every $n \in \mathbb{N}$ satisfies

$$\Delta e_n = \sum_{i=0}^{n} e_i \otimes e_{n-i}. \tag{13.90.14}$$

Thus, every $n \in \mathbb{N}$ satisfies

$$\Delta e_n = \sum_{i \in \mathbb{N}} e_i \otimes e_{n-i}. \tag{13.90.15}$$

[716] Similarly, every $n \in \mathbb{N}$ satisfies

$$\Delta h_n = \sum_{i \in \mathbb{N}} h_i \otimes h_{n-i}. \tag{13.90.16}$$

Let $a \in \mathbb{N}$ and $b \in \mathbb{N}$. Exercise 2.9.14(b) yields $\sum_{i=0}^{b} (-1)^i h_{a+i+1} e_{b-i} = s_{(a+1,1^b)}$, so that

$$s_{(a+1,1^b)} = \sum_{i=0}^{b} (-1)^i h_{a+i+1} e_{b-i} = \sum_{k=0}^{b} (-1)^k h_{a+k+1} e_{b-k}$$

[716] Indeed, (13.90.15) follows from (13.90.14), because

$$\sum_{i \in \mathbb{N}} e_i \otimes e_{n-i} = \sum_{i=0}^{n} e_i \otimes e_{n-i} + \sum_{i=n+1}^{\infty} e_i \otimes \underbrace{e_{n-i}}_{\substack{=0 \\ (\text{since } n-i<0)}} = \sum_{i=0}^{n} e_i \otimes e_{n-i} + \underbrace{\sum_{i=n+1}^{\infty} e_i \otimes 0}_{=0} = \sum_{i=0}^{n} e_i \otimes e_{n-i}.$$

(here, we have renamed the summation index i as k). Applying the map Δ to both sides of this equality, we obtain

$$\begin{aligned}
\Delta s_{(a+1,1^b)} &= \Delta\left(\sum_{k=0}^{b}(-1)^k h_{a+k+1}e_{b-k}\right) \\
&= \sum_{k=0}^{b}(-1)^k \underbrace{\Delta\left(h_{a+k+1}\right)}_{\substack{=\sum_{i\in\mathbb{N}} h_i\otimes h_{a+k+1-i} \\ \text{(by (13.90.16), applied to } n=a+k+1\text{)}}} \cdot \underbrace{\Delta\left(e_{b-k}\right)}_{\substack{=\sum_{i\in\mathbb{N}} e_i\otimes e_{b-k-i} \\ \text{(by (13.90.15), applied to } n=b-k\text{)}}} \\
&\qquad \left(\text{since } \Delta \text{ is a } \mathbf{k}\text{-algebra homomorphism}\right) \\
&= \sum_{k=0}^{b}(-1)^k \left(\sum_{i\in\mathbb{N}} h_i\otimes h_{a+k+1-i}\right)\left(\sum_{i\in\mathbb{N}} e_i\otimes e_{b-k-i}\right) \\
&= \sum_{k=0}^{b}(-1)^k \left(\sum_{i\in\mathbb{N}} h_i\otimes h_{a+k+1-i}\right)\left(\sum_{j\in\mathbb{N}} e_j\otimes e_{b-k-j}\right) \\
&\qquad \left(\text{here, we renamed the summation index } i \text{ as } j \text{ in the third sum}\right) \\
&= \sum_{k=0}^{b}(-1)^k \sum_{i\in\mathbb{N}}\sum_{j\in\mathbb{N}} \underbrace{\left(h_i\otimes h_{a+k+1-i}\right)\left(e_j\otimes e_{b-k-j}\right)}_{=h_ie_j\otimes h_{a+k+1-i}e_{b-k-j}} \\
&= \sum_{k=0}^{b}(-1)^k \sum_{i\in\mathbb{N}}\sum_{j\in\mathbb{N}} h_ie_j\otimes h_{a+k+1-i}e_{b-k-j} \\
&= \underbrace{\sum_{i\in\mathbb{N}}\sum_{j\in\mathbb{N}}}_{=\sum_{(i,j)\in\mathbb{N}^2}} \underbrace{h_ie_j}_{\substack{=e_jh_i \\ =\mathbf{s}(i,j-1)+\mathbf{s}(i-1,j) \\ \text{(by (13.90.5), applied to } j \text{ and } i \\ \text{instead of } n \text{ and } m\text{)}}} \otimes \underbrace{\left(\sum_{k=0}^{b}(-1)^k h_{a-i+k+1}e_{b-j-k}\right)}_{\substack{=\mathbf{s}(a-i,b-j) \\ \text{(by (13.90.10), applied to } a-i,\ b-j \text{ and } b \\ \text{instead of } a,\ b \text{ and } c \text{ (since } b\geq b-j\text{))}}} \\
&= \sum_{(i,j)\in\mathbb{N}^2}\left(\mathbf{s}\left(i,j-1\right)+\mathbf{s}\left(i-1,j\right)\right)\otimes\mathbf{s}\left(a-i,b-j\right). \qquad (13.90.17)
\end{aligned}$$

Now, applying (13.90.13) to $x = s_{(a+1,1^b)}$, we obtain

$$\begin{aligned}
&\Delta s_{(a+1,1^b)} - 1\otimes s_{(a+1,1^b)} - s_{(a+1,1^b)}\otimes 1 \\
&= \left(\overline{\mathrm{id}}\otimes\overline{\mathrm{id}}\right) \underbrace{\left(\Delta s_{(a+1,1^b)}\right)}_{\substack{=\sum_{(i,j)\in\mathbb{N}^2}(\mathbf{s}(i,j-1)+\mathbf{s}(i-1,j))\otimes\mathbf{s}(a-i,b-j) \\ \text{(by (13.90.17))}}} - \underbrace{\epsilon\left(s_{(a+1,1^b)}\right)}_{\substack{=0 \\ \text{(since } \left(a+1,1^b\right) \text{ is a} \\ \text{nonempty partition)}}} 1\otimes 1 \\
&= \left(\overline{\mathrm{id}}\otimes\overline{\mathrm{id}}\right)\left(\sum_{(i,j)\in\mathbb{N}^2}\left(\mathbf{s}\left(i,j-1\right)+\mathbf{s}\left(i-1,j\right)\right)\otimes\mathbf{s}\left(a-i,b-j\right)\right) - \underbrace{0\cdot 1\otimes 1}_{=0} \\
&= \left(\overline{\mathrm{id}}\otimes\overline{\mathrm{id}}\right)\left(\sum_{(i,j)\in\mathbb{N}^2}\left(\mathbf{s}\left(i,j-1\right)+\mathbf{s}\left(i-1,j\right)\right)\otimes\mathbf{s}\left(a-i,b-j\right)\right) \\
&= \sum_{(i,j)\in\mathbb{N}^2}\overline{\mathrm{id}}\left(\mathbf{s}\left(i,j-1\right)+\mathbf{s}\left(i-1,j\right)\right)\otimes\overline{\mathrm{id}}\left(\mathbf{s}\left(a-i,b-j\right)\right) \\
&= \sum_{(i,j)\in\mathbb{N}^2}\overline{\mathrm{id}}\left(\mathbf{s}\left(i-1,j\right)\right)\otimes\overline{\mathrm{id}}\left(\mathbf{s}\left(a-i,b-j\right)\right) + \sum_{(i,j)\in\mathbb{N}^2}\overline{\mathrm{id}}\left(\mathbf{s}\left(i,j-1\right)\right)\otimes\overline{\mathrm{id}}\left(\mathbf{s}\left(a-i,b-j\right)\right). \qquad (13.90.18)
\end{aligned}$$

We can now observe that

$$\sum_{(i,j)\in\mathbb{N}^2} \overline{\operatorname{id}}\left(\mathbf{s}\left(i-1,j\right)\right)\otimes\overline{\operatorname{id}}\left(\mathbf{s}\left(a-i,b-j\right)\right) = \sum_{\substack{(c,d,e,f)\in\mathbb{N}^4;\\ c+e=a-1;\\ d+f=b}} s_{(c+1,1^d)}\otimes s_{(e+1,1^f)}. \tag{13.90.19}$$

[717] Similarly,

$$\sum_{(i,j)\in\mathbb{N}^2} \overline{\operatorname{id}}\left(\mathbf{s}\left(i,j-1\right)\right)\otimes\overline{\operatorname{id}}\left(\mathbf{s}\left(a-i,b-j\right)\right) = \sum_{\substack{(c,d,e,f)\in\mathbb{N}^4;\\ c+e=a;\\ d+f=b-1}} s_{(c+1,1^d)}\otimes s_{(e+1,1^f)}. \tag{13.90.21}$$

Now, (13.90.18) becomes

$$\begin{aligned}
&\Delta s_{(a+1,1^b)} - 1\otimes s_{(a+1,1^b)} - s_{(a+1,1^b)}\otimes 1\\
&= \underbrace{\sum_{(i,j)\in\mathbb{N}^2} \overline{\operatorname{id}}\left(\mathbf{s}\left(i-1,j\right)\right)\otimes\overline{\operatorname{id}}\left(\mathbf{s}\left(a-i,b-j\right)\right)}_{\substack{=\sum_{\substack{(c,d,e,f)\in\mathbb{N}^4;\\ c+e=a-1;\\ d+f=b}} s_{(c+1,1^d)}\otimes s_{(e+1,1^f)}\\ \text{(by (13.90.19))}}} + \underbrace{\sum_{(i,j)\in\mathbb{N}^2} \overline{\operatorname{id}}\left(\mathbf{s}\left(i,j-1\right)\right)\otimes\overline{\operatorname{id}}\left(\mathbf{s}\left(a-i,b-j\right)\right)}_{\substack{=\sum_{\substack{(c,d,e,f)\in\mathbb{N}^4;\\ c+e=a;\\ d+f=b-1}} s_{(c+1,1^d)}\otimes s_{(e+1,1^f)}\\ \text{(by (13.90.21))}}}\\
&= \sum_{\substack{(c,d,e,f)\in\mathbb{N}^4;\\ c+e=a-1;\\ d+f=b}} s_{(c+1,1^d)}\otimes s_{(e+1,1^f)} + \sum_{\substack{(c,d,e,f)\in\mathbb{N}^4;\\ c+e=a;\\ d+f=b-1}} s_{(c+1,1^d)}\otimes s_{(e+1,1^f)}.
\end{aligned}$$

Adding $1\otimes s_{(a+1,1^b)} + s_{(a+1,1^b)}\otimes 1$ to both sides of this equality, we obtain the claim of Exercise 2.9.14(d).

[717] *Proof of (13.90.19):* Let $\overline{\mathbb{N}} = \{-1,0,1,2,\ldots\} = \{-1\}\cup\mathbb{N}$. We have

$$\sum_{\substack{(c,d,e,f)\in\overline{\mathbb{N}}\times\mathbb{N}\times\mathbb{Z}\times\mathbb{Z};\\ c+e=a-1;\\ d+f=b}} \overline{\operatorname{id}}\left(\mathbf{s}\left(c,d\right)\right)\otimes\overline{\operatorname{id}}\left(\mathbf{s}\left(e,f\right)\right) = \sum_{\substack{(c,d,e,f)\in\mathbb{N}^4;\\ c+e=a-1;\\ d+f=b}} \overline{\operatorname{id}}\left(\mathbf{s}\left(c,d\right)\right)\otimes\overline{\operatorname{id}}\left(\mathbf{s}\left(e,f\right)\right). \tag{13.90.20}$$

(In fact, the sum on the left hand side of (13.90.20) differs from that on the right hand side of (13.90.20) only by the presence of addends of the form $\overline{\operatorname{id}}\left(\mathbf{s}\left(c,d\right)\right)\otimes\overline{\operatorname{id}}\left(\mathbf{s}\left(e,f\right)\right)$ for certain quadruples $(c,d,e,f)\in\overline{\mathbb{N}}\times\mathbb{N}\times\mathbb{Z}\times\mathbb{Z}$ which don't belong to $\mathbb{N}^4$. But all such addends are 0 (as can be easily seen using (13.90.12)), and so the two sums have the same value, and (13.90.20) is proven.)

The map

$$\begin{aligned}
\mathbb{N}^2 &\to \left\{(c,d,e,f)\in\overline{\mathbb{N}}\times\mathbb{N}\times\mathbb{Z}\times\mathbb{Z} \mid c+e=a-1;\ d+f=b\right\},\\
(i,j) &\mapsto (i-1,j,a-i,b-j)
\end{aligned}$$

is a bijection. Hence, we can substitute $(i-1,j,a-i,b-j)$ for (c,d,e,f) in the sum $\sum_{\substack{(c,d,e,f)\in\overline{\mathbb{N}}\times\mathbb{N}\times\mathbb{Z}\times\mathbb{Z};\\ c+e=a-1;\\ d+f=b}} \overline{\operatorname{id}}\left(\mathbf{s}\left(c,d\right)\right)\otimes\overline{\operatorname{id}}\left(\mathbf{s}\left(e,f\right)\right)$, and thus obtain

$$\sum_{\substack{(c,d,e,f)\in\overline{\mathbb{N}}\times\mathbb{N}\times\mathbb{Z}\times\mathbb{Z};\\ c+e=a-1;\\ d+f=b}} \overline{\operatorname{id}}\left(\mathbf{s}\left(c,d\right)\right)\otimes\overline{\operatorname{id}}\left(\mathbf{s}\left(e,f\right)\right) = \sum_{(i,j)\in\mathbb{N}^2} \overline{\operatorname{id}}\left(\mathbf{s}\left(i-1,j\right)\right)\otimes\overline{\operatorname{id}}\left(\mathbf{s}\left(a-i,b-j\right)\right).$$

Comparing this with (13.90.20), we obtain

$$\begin{aligned}
&\sum_{(i,j)\in\mathbb{N}^2} \overline{\operatorname{id}}\left(\mathbf{s}\left(i-1,j\right)\right)\otimes\overline{\operatorname{id}}\left(\mathbf{s}\left(a-i,b-j\right)\right)\\
&= \sum_{\substack{(c,d,e,f)\in\mathbb{N}^4;\\ c+e=a-1;\\ d+f=b}} \underbrace{\overline{\operatorname{id}}\left(\mathbf{s}\left(c,d\right)\right)}_{\substack{=s_{(c+1,1^d)}\\ \text{(by (13.90.11), applied to}\\ c\text{ and }d\text{ instead of }a\text{ and }b)}} \otimes \underbrace{\overline{\operatorname{id}}\left(\mathbf{s}\left(e,f\right)\right)}_{\substack{=s_{(e+1,1^f)}\\ \text{(by (13.90.11), applied to}\\ e\text{ and }f\text{ instead of }a\text{ and }b)}}\\
&= \sum_{\substack{(c,d,e,f)\in\mathbb{N}^4;\\ c+e=a-1;\\ d+f=b}} s_{(c+1,1^d)}\otimes s_{(e+1,1^f)}.
\end{aligned}$$

This proves (13.90.19).

Remark: In our above solution, we have first proved (13.90.5) and then used it to derive (13.90.8). There is also an alternative way to prove these two identities, which proceeds the other way round. The main idea is to obtain (13.90.8) by applying Exercise 2.9.1(b) to $\lambda = (1^b)$ and $m = a + 1$. (This works in the case of $a \in \mathbb{N}$ and $b \in \mathbb{N}$ only. The remaining cases, however, are easy to either check directly or reduce to (2.4.4).) Once this is done, (13.90.5) can be verified by rewriting both $\mathbf{s}(m, n-1)$ and $\mathbf{s}(m-1, n)$ using (13.90.8).

13.91. **Solution to Exercise 2.9.15.** *Solution to Exercise 2.9.15.* Consider the partition $(m^k) = \left(\underbrace{m, m, \ldots, m}_{k \text{ times}}\right)$.

(a) The fact that $\lambda^\vee$ and $\mu^\vee$ are partitions is easy to check. It remains to show that $s_{\lambda/\mu} = s_{\mu^\vee/\lambda^\vee}$.

It is easy to show that if $\mu \subseteq \lambda$ does not hold, then $\lambda^\vee \subseteq \mu^\vee$ does not hold either (because if some positive integer i fails to satisfy $\mu_i \leq \lambda_i$, then this i belongs to $\{1, 2, \ldots, k\}$ and fails to satisfy $m - \lambda_i \leq m - \mu_i$ as well). Hence, if $\mu \subseteq \lambda$ does not hold, then both $s_{\lambda/\mu}$ and $s_{\mu^\vee/\lambda^\vee}$ are 0, and therefore the equality $s_{\lambda/\mu} = s_{\mu^\vee/\lambda^\vee}$ is obvious. Hence, for the rest of the proof of $s_{\lambda/\mu} = s_{\mu^\vee/\lambda^\vee}$, we WLOG assume that $\mu \subseteq \lambda$ does hold. Then, it is easy to see that $\lambda^\vee \subseteq \mu^\vee$ holds, too.

Now, let Z denote the 180° rotation around the center of the Ferrers diagram of (m^k). Let $Y(\rho)$ denote the Ferrers diagram of ρ whenever ρ is a partition or a skew partition. It is straightforward to see that $Z\left(Y\left((m^k)\right) \setminus Y(\mu)\right) = Y(\mu^\vee)$ and $Z\left(Y\left((m^k)\right) \setminus Y(\lambda)\right) = Y(\lambda^\vee)$. Now,

$$Y(\mu^\vee/\lambda^\vee) = \underbrace{Y(\mu^\vee)}_{=Z(Y((m^k))\setminus Y(\mu))} \setminus \underbrace{Y(\lambda^\vee)}_{=Z(Y((m^k))\setminus Y(\lambda))} = Z\left(Y\left((m^k)\right) \setminus Y(\mu)\right) \setminus Z\left(Y\left((m^k)\right) \setminus Y(\lambda)\right)$$

$$= Z\left(\underbrace{\left(Y\left((m^k)\right) \setminus Y(\mu)\right) \setminus \left(Y\left((m^k)\right) \setminus Y(\lambda)\right)}_{\substack{=Y(\lambda)\setminus Y(\mu) \\ \text{(since one can easily see that both } Y(\lambda) \text{ and } Y(\mu) \\ \text{are subsets of } Y((m^k)))}}\right)$$

$$= Z\left(\underbrace{Y(\lambda) \setminus Y(\mu)}_{=Y(\lambda/\mu)}\right) = Z\left(Y(\lambda/\mu)\right).$$

Hence, the skew Ferrers diagram $\mu^\vee/\lambda^\vee$ can be obtained from the skew Ferrers diagram λ/μ by a 180° rotation (namely, by the 180° rotation Z). Thus, Exercise 2.3.4(b) (applied to $\lambda' = \mu^\vee$ and $\mu' = \lambda^\vee$) yields $s_{\lambda/\mu} = s_{\mu^\vee/\lambda^\vee}$. This completes the solution of Exercise 2.9.15(a).

(b) According to Remark 2.5.9, we have $s_{\lambda/\mu} = \sum_\nu c^\lambda_{\mu,\nu} s_\nu$, where the sum ranges over all partitions ν. In other words, we have

$$s_{\lambda/\mu} = \sum_{\nu \in \mathrm{Par}} c^\lambda_{\mu,\nu} s_\nu. \tag{13.91.1}$$

Exercise 2.9.15(b) yields that $\lambda^\vee$ and $\mu^\vee$ are partitions, and that $s_{\lambda/\mu} = s_{\mu^\vee/\lambda^\vee}$. Applying (13.91.1) to $\mu^\vee$ and $\lambda^\vee$ instead of λ and μ, we obtain

$$s_{\mu^\vee/\lambda^\vee} = \sum_{\nu \in \mathrm{Par}} c^{\mu^\vee}_{\lambda^\vee,\nu} s_\nu.$$

Now, (13.91.1) yields

$$\sum_{\nu \in \mathrm{Par}} c^\lambda_{\mu,\nu} s_\nu = s_{\lambda/\mu} = s_{\mu^\vee/\lambda^\vee} = \sum_{\nu \in \mathrm{Par}} c^{\mu^\vee}_{\lambda^\vee,\nu} s_\nu.$$

Comparing coefficients before s_ν in this equality, we conclude that $c^\lambda_{\mu,\nu} = c^{\mu^\vee}_{\lambda^\vee,\nu}$ for every $\nu \in \mathrm{Par}$ (since $(s_\nu)_{\nu \in \mathrm{Par}}$ is a basis of the $\mathbf{k}$-module Λ). This solves Exercise 2.9.15(b).

(c) Exercise 2.9.15(a) (applied to ν instead of μ) yields that $\lambda^\vee$ and $\nu^\vee$ are partitions and that $s_{\lambda/\nu} = s_{\nu^\vee/\lambda^\vee}$.

From (2.5.8), we obtain

$$\begin{aligned} c^{\lambda}_{\mu,\nu} = c^{\lambda}_{\nu,\mu} = c^{\nu^\vee}_{\lambda^\vee,\mu} \qquad & \text{(by Exercise 2.9.15(b) (applied to } \nu \text{ and } \mu \text{ instead of } \mu \text{ and } \nu\text{))} \\ = c^{\nu^\vee}_{\mu,\lambda^\vee} \qquad & \text{(by (2.5.8) (applied to } \nu^\vee\text{, } \lambda^\vee \text{ and } \mu \text{ instead of } \lambda\text{, } \mu \text{ and } \nu\text{))} . \end{aligned}$$

On the other hand, Exercise 2.9.15(b) yields $c^{\lambda}_{\mu,\nu} = c^{\mu^\vee}_{\lambda^\vee,\nu} = c^{\mu^\vee}_{\nu,\lambda^\vee}$ (by (2.5.8) (applied to $\mu^\vee$, $\lambda^\vee$ and ν instead of λ, μ and ν)). Thus, $c^{\mu^\vee}_{\nu,\lambda^\vee} = c^{\mu^\vee}_{\lambda^\vee,\nu} = c^{\lambda}_{\mu,\nu} = c^{\lambda}_{\nu,\mu} = c^{\nu^\vee}_{\lambda^\vee,\mu} = c^{\nu^\vee}_{\mu,\lambda^\vee}$. This solves Exercise 2.9.15(c).

(d) *First solution to Exercise 2.9.15(d):* Notice that $\lambda^\vee = (m - \lambda_k, m - \lambda_{k-1}, \ldots, m - \lambda_1)$. Thus, $\ell(\lambda^\vee) \leq k$.

Let $n = k$. We shall use the notations of Section 2.6; in particular, we set $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ and $\rho = (n-1, n-2, \ldots, 2, 1, 0)$. Clearly, $\mathbf{x} = (x_1, x_2, \ldots, x_n) = (x_1, x_2, \ldots, x_k)$ (since $n = k$). Notice that $\ell(\lambda) \leq k = n$. Hence, we can regard λ as an element of $\mathbb{N}^n$; therefore, $\lambda + \rho$ is a well-defined element of $\mathbb{N}^n$, and the alternant $a_{\lambda+\rho}$ is well-defined. Corollary 2.6.7 yields that $s_\lambda(\mathbf{x}) = \dfrac{a_{\lambda+\rho}}{a_\rho}$, so that

$$a_\rho \cdot s_\lambda(\mathbf{x}) = a_{\lambda+\rho}. \tag{13.91.2}$$

Applying this equality to $\lambda^\vee$ instead of λ, we obtain

$$a_\rho \cdot s_{\lambda^\vee}(\mathbf{x}) = a_{\lambda^\vee+\rho} \tag{13.91.3}$$

(since $\lambda^\vee$ is also a partition satisfying $\ell(\lambda^\vee) \leq k = n$).

Substituting the variables $x_1^{-1}, x_2^{-1}, \ldots, x_n^{-1}$ for $x_1, x_2, \ldots, x_n$ in the equality (13.91.2), we obtain

$$a_\rho\left(x_1^{-1}, x_2^{-1}, \ldots, x_n^{-1}\right) \cdot s_\lambda\left(x_1^{-1}, x_2^{-1}, \ldots, x_n^{-1}\right) = a_{\lambda+\rho}\left(x_1^{-1}, x_2^{-1}, \ldots, x_n^{-1}\right). \tag{13.91.4}$$

Let $w_\circ : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$ be the map which sends every $i \in \{1, 2, \ldots, n\}$ to $n + 1 - i$. Then, $w_\circ$ is a permutation of $\{1, 2, \ldots, n\}$, thus an element of $\mathfrak{S}_n$. For every $i \in \{1, 2, \ldots, n\}$, we have

$$(w \circ w_\circ)(i) = w(n + 1 - i). \tag{13.91.5}$$

The map $\mathfrak{S}_n \to \mathfrak{S}_n,\ w \mapsto w \circ w_\circ$ is a bijection (since $\mathfrak{S}_n$ is a group, and since $w_\circ \in \mathfrak{S}_n$).

But every $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{N}^n$ satisfies

$$\begin{aligned} a_\alpha &= \sum_{w \in \mathfrak{S}_n} \operatorname{sgn}(w) \underbrace{w(\mathbf{x}^\alpha)}_{\substack{= \prod_{i=1}^n x_{w(i)}^{\alpha_i} \\ \text{(by the definition of } w(\mathbf{x}^\alpha))}} \qquad \text{(by the definition of } a_\alpha) \\ &= \sum_{w \in \mathfrak{S}_n} \operatorname{sgn}(w) \prod_{i=1}^n x_{w(i)}^{\alpha_i}. \end{aligned} \tag{13.91.6}$$

But $\lambda = (\lambda_1, \lambda_2, \lambda_3, \ldots)$, thus $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n)$ (since $\ell(\lambda) \leq k = n$). Hence,

$$\begin{aligned} \underbrace{\lambda}_{=(\lambda_1, \lambda_2, \ldots, \lambda_n)} + \underbrace{\rho}_{\substack{=(n-1, n-2, \ldots, 2, 1, 0) \\ =(n-1, n-2, \ldots, n-n)}} &= (\lambda_1, \lambda_2, \ldots, \lambda_n) + (n-1, n-2, \ldots, n-n) \\ &= (\lambda_1 + n - 1, \lambda_2 + n - 2, \ldots, \lambda_n + n - n). \end{aligned}$$

Thus, (13.91.6) (applied to $\lambda + \rho$ and $(\lambda_1 + n - 1, \lambda_2 + n - 2, \ldots, \lambda_n + n - n)$ instead of α and $(\alpha_1, \alpha_2, \ldots, \alpha_n)$) yields

$$a_{\lambda+\rho} = \sum_{w \in \mathfrak{S}_n} \operatorname{sgn}(w) \prod_{i=1}^n x_{w(i)}^{\lambda_i + n - i}. \tag{13.91.7}$$

Substituting x_1^{-1}, x_2^{-1}, …, x_n^{-1} for x_1, x_2, …, x_n on both sides of this equality, we obtain

$$\begin{aligned}
& a_{\lambda+\rho}\left(x_1^{-1},x_2^{-1},\ldots,x_n^{-1}\right) \\
&= \sum_{w\in\mathfrak{S}_n} \operatorname{sgn}(w) \prod_{i=1}^{n} \underbrace{\left(x_{w(i)}^{-1}\right)^{\lambda_i+n-i}}_{=x_{w(i)}^{-(\lambda_i+n-i)}} = \sum_{w\in\mathfrak{S}_n} \operatorname{sgn}(w) \underbrace{\prod_{i=1}^{n} x_{w(i)}^{-(\lambda_i+n-i)}}_{\substack{=\prod_{i=1}^{n} x_{w(n+1-i)}^{-(\lambda_{n+1-i}+n-(n+1-i))} \\ \text{(here, we have substituted} \\ n+1-i \text{ for } i \text{ in the product)}}} \\
&= \sum_{w\in\mathfrak{S}_n} \underbrace{\operatorname{sgn}(w)}_{=\frac{1}{\operatorname{sgn}(w_\circ)}\operatorname{sgn}(w)\operatorname{sgn}(w_\circ)} \prod_{i=1}^{n} \underbrace{x_{w(n+1-i)}^{-(\lambda_{n+1-i}+n-(n+1-i))}}_{\substack{=x_{w(n+1-i)}^{-(\lambda_{n+1-i}+i-1)} \\ (\text{since } n-(n+1-i)=i-1)}} \\
&= \sum_{w\in\mathfrak{S}_n} \frac{1}{\operatorname{sgn}(w_\circ)} \underbrace{\operatorname{sgn}(w)\operatorname{sgn}(w_\circ)}_{=\operatorname{sgn}(w\circ w_\circ)} \prod_{i=1}^{n} \underbrace{x_{w(n+1-i)}^{-(\lambda_{n+1-i}+i-1)}}_{\substack{=x_{(w\circ w_\circ)(i)}^{-(\lambda_{n+1-i}+i-1)} \\ (\text{since } w(n+1-i)=(w\circ w_\circ)(i) \\ (\text{by } (13.91.5)))}} \\
&= \sum_{w\in\mathfrak{S}_n} \frac{1}{\operatorname{sgn}(w_\circ)} \operatorname{sgn}(w\circ w_\circ) \prod_{i=1}^{n} x_{(w\circ w_\circ)(i)}^{-(\lambda_{n+1-i}+i-1)} = \sum_{w\in\mathfrak{S}_n} \frac{1}{\operatorname{sgn}(w_\circ)} \operatorname{sgn}(w) \prod_{i=1}^{n} x_{w(i)}^{-(\lambda_{n+1-i}+i-1)} \\
&\qquad \left(\begin{array}{c} \text{here, we substituted } w \text{ for } w\circ w_\circ \text{ in the sum} \\ \text{(since the map } \mathfrak{S}_n\to\mathfrak{S}_n,\ w\mapsto w\circ w_\circ \text{ is a bijection)} \end{array}\right) \\
&= \frac{1}{\operatorname{sgn}(w_\circ)} \sum_{w\in\mathfrak{S}_n} \operatorname{sgn}(w) \prod_{i=1}^{n} x_{w(i)}^{-(\lambda_{n+1-i}+i-1)}.
\end{aligned}$$

Multiplying both sides of this equality by $\operatorname{sgn}(w_\circ)$, we obtain

$$\operatorname{sgn}(w_\circ)\cdot a_{\lambda+\rho}\left(x_1^{-1},x_2^{-1},\ldots,x_n^{-1}\right) = \sum_{w\in\mathfrak{S}_n} \operatorname{sgn}(w) \prod_{i=1}^{n} x_{w(i)}^{-(\lambda_{n+1-i}+i-1)}. \tag{13.91.8}$$

But $\lambda^\vee = (m-\lambda_k, m-\lambda_{k-1}, \ldots, m-\lambda_1) = (m-\lambda_n, m-\lambda_{n-1}, \ldots, m-\lambda_1)$ (since $k=n$), so that

$$\begin{aligned}
& \underbrace{\lambda^\vee}_{\substack{=(m-\lambda_n,m-\lambda_{n-1},\ldots,m-\lambda_1) \\ =\left(m-\lambda_{(n+1)-1},m-\lambda_{(n+1)-2},\ldots,m-\lambda_{(n+1)-n}\right)}} + \underbrace{\rho}_{\substack{=(n-1,n-2,\ldots,2,1,0) \\ =(n-1,n-2,\ldots,n-n)}} \\
&= \left(m-\lambda_{(n+1)-1}, m-\lambda_{(n+1)-2}, \ldots, m-\lambda_{(n+1)-n}\right) + (n-1, n-2, \ldots, n-n) \\
&= \left(\left(m-\lambda_{(n+1)-1}\right)+(n-1), \left(m-\lambda_{(n+1)-2}\right)+(n-2), \ldots, \left(m-\lambda_{(n+1)-n}\right)+(n-n)\right).
\end{aligned}$$

Hence, (13.91.6) (applied to $\lambda^\vee+\rho$ and $\left(\left(m-\lambda_{(n+1)-1}\right)+(n-1), \left(m-\lambda_{(n+1)-2}\right)+(n-2), \ldots, \left(m-\lambda_{(n+1)-n}\right)+(n-n)\right)$ instead of α and

$(\alpha_1, \alpha_2, \ldots, \alpha_n))$ yields

$$\begin{aligned}
a_{\lambda^\vee+\rho} &= \sum_{w \in \mathfrak{S}_n} \operatorname{sgn}(w) \prod_{i=1}^{n} \underbrace{x_{w(i)}^{\left(m-\lambda_{(n+1)-i}\right)+(n-i)}}_{\substack{= x_{w(i)}^{-\left(\lambda_{n+1-i}+i-1\right)+(n+m-1)} \\ = x_{w(i)}^{-\left(\lambda_{n+1-i}+i-1\right)} x_{w(i)}^{n+m-1}}} \\
&= \sum_{w \in \mathfrak{S}_n} \operatorname{sgn}(w) \underbrace{\prod_{i=1}^{n} \left(x_{w(i)}^{-\left(\lambda_{n+1-i}+i-1\right)} x_{w(i)}^{n+m-1} \right)}_{=\left(\prod_{i=1}^{n} x_{w(i)}^{-\left(\lambda_{n+1-i}+i-1\right)}\right)\left(\prod_{i=1}^{n} x_{w(i)}^{n+m-1}\right)} \\
&= \sum_{w \in \mathfrak{S}_n} \operatorname{sgn}(w) \left(\prod_{i=1}^{n} x_{w(i)}^{-\left(\lambda_{n+1-i}+i-1\right)} \right) \underbrace{\left(\prod_{i=1}^{n} x_{w(i)}^{n+m-1} \right)}_{\substack{=\prod_{i=1}^{n} x_i^{n+m-1} \\ \text{(here, we have substituted } i \text{ for } w(i) \text{ in the product} \\ \text{(since } w \text{ is a permutation of } \{1,2,\ldots,n\}))}} \\
&= \sum_{w \in \mathfrak{S}_n} \operatorname{sgn}(w) \left(\prod_{i=1}^{n} x_{w(i)}^{-\left(\lambda_{n+1-i}+i-1\right)} \right) \left(\prod_{i=1}^{n} x_i^{n+m-1} \right) \\
&= \underbrace{\left(\prod_{i=1}^{n} x_i^{n+m-1} \right)}_{=\left(\prod_{i=1}^{n} x_i\right)^{n+m-1}} \underbrace{\sum_{w \in \mathfrak{S}_n} \operatorname{sgn}(w) \prod_{i=1}^{n} x_{w(i)}^{-\left(\lambda_{n+1-i}+i-1\right)}}_{\substack{=\operatorname{sgn}(w_\circ) \cdot a_{\lambda+\rho}\left(x_1^{-1}, x_2^{-1}, \ldots, x_n^{-1}\right) \\ \text{(by (13.91.8))}}} \\
&= \left(\prod_{i=1}^{n} x_i \right)^{n+m-1} \cdot \operatorname{sgn}(w_\circ) \cdot a_{\lambda+\rho}\left(x_1^{-1}, x_2^{-1}, \ldots, x_n^{-1}\right). \qquad (13.91.9)
\end{aligned}$$

Finally, $\rho = (n-1, n-2, \ldots, 2, 1, 0) = (n-1, n-2, \ldots, n-n)$. Thus, (13.91.6) (applied to ρ and $(n-1, n-2, \ldots, n-n)$ instead of α and $(\alpha_1, \alpha_2, \ldots, \alpha_n))$ yields

$$a_\rho = \sum_{w \in \mathfrak{S}_n} \operatorname{sgn}(w) \prod_{i=1}^{n} x_{w(i)}^{n-i}. \qquad (13.91.10)$$

Substituting x_1^{-1}, x_2^{-1}, …, x_n^{-1} for x_1, x_2, …, x_n on both sides of this equality, we obtain

$$\begin{aligned}
&a_\rho\left(x_1^{-1}, x_2^{-1}, \ldots, x_n^{-1}\right) \\
&= \sum_{w\in\mathfrak{S}_n} \operatorname{sgn}(w) \prod_{i=1}^{n} \underbrace{\left(x_{w(i)}^{-1}\right)^{n-i}}_{=x_{w(i)}^{-(n-i)}=x_{w(i)}^{i-n}} = \sum_{w\in\mathfrak{S}_n} \operatorname{sgn}(w) \underbrace{\prod_{i=1}^{n} x_{w(i)}^{i-n}}_{\substack{=\prod_{i=1}^{n} x_{w(n+1-i)}^{(n+1-i)-n} \\ \text{(here, we have substituted} \\ n+1-i \text{ for } i \text{ in the product)}}} \\
&= \sum_{w\in\mathfrak{S}_n} \underbrace{\operatorname{sgn}(w)}_{=\frac{1}{\operatorname{sgn}(w_\circ)} \operatorname{sgn}(w)\operatorname{sgn}(w_\circ)} \prod_{i=1}^{n} \underbrace{x_{w(n+1-i)}^{(n+1-i)-n}}_{=x_{w(n+1-i)}^{-i+1}} \\
&= \sum_{w\in\mathfrak{S}_n} \frac{1}{\operatorname{sgn}(w_\circ)} \underbrace{\operatorname{sgn}(w)\operatorname{sgn}(w_\circ)}_{=\operatorname{sgn}(w\circ w_\circ)} \prod_{i=1}^{n} \underbrace{x_{w(n+1-i)}^{-i+1}}_{\substack{=x_{(w\circ w_\circ)(i)}^{-i+1} \\ \text{(since } w(n+1-i)=(w\circ w_\circ)(i) \\ \text{(by (13.91.5)))}}} \\
&= \sum_{w\in\mathfrak{S}_n} \frac{1}{\operatorname{sgn}(w_\circ)} \operatorname{sgn}(w\circ w_\circ) \prod_{i=1}^{n} x_{(w\circ w_\circ)(i)}^{-i+1} = \sum_{w\in\mathfrak{S}_n} \frac{1}{\operatorname{sgn}(w_\circ)} \operatorname{sgn}(w) \prod_{i=1}^{n} x_{w(i)}^{-i+1} \\
&\qquad \left(\begin{array}{c} \text{here, we substituted } w \text{ for } w\circ w_\circ \text{ in the sum} \\ \text{(since the map } \mathfrak{S}_n \to \mathfrak{S}_n,\ w\mapsto w\circ w_\circ \text{ is a bijection)} \end{array}\right) \\
&= \frac{1}{\operatorname{sgn}(w_\circ)} \sum_{w\in\mathfrak{S}_n} \operatorname{sgn}(w) \prod_{i=1}^{n} \underbrace{x_{w(i)}^{-i+1}}_{\substack{=x_{w(i)}^{(n-i)+(1-n)} \\ =x_{w(i)}^{n-i} x_{w(i)}^{1-n}}} = \frac{1}{\operatorname{sgn}(w_\circ)} \sum_{w\in\mathfrak{S}_n} \operatorname{sgn}(w) \underbrace{\prod_{i=1}^{n} \left(x_{w(i)}^{n-i} x_{w(i)}^{1-n}\right)}_{=\left(\prod_{i=1}^{n} x_{w(i)}^{n-i}\right)\left(\prod_{i=1}^{n} x_{w(i)}^{1-n}\right)} \\
&= \frac{1}{\operatorname{sgn}(w_\circ)} \sum_{w\in\mathfrak{S}_n} \operatorname{sgn}(w) \left(\prod_{i=1}^{n} x_{w(i)}^{n-i}\right) \underbrace{\left(\prod_{i=1}^{n} x_{w(i)}^{1-n}\right)}_{\substack{=\prod_{i=1}^{n} x_i^{1-n} \\ \text{(here, we have substituted } i \text{ for } w(i) \text{ in the product} \\ \text{(since } w \text{ is a permutation of } \{1,2,\ldots,n\}))}} \\
&= \frac{1}{\operatorname{sgn}(w_\circ)} \sum_{w\in\mathfrak{S}_n} \operatorname{sgn}(w) \left(\prod_{i=1}^{n} x_{w(i)}^{n-i}\right) \left(\prod_{i=1}^{n} x_i^{1-n}\right) = \frac{1}{\operatorname{sgn}(w_\circ)} \underbrace{\left(\prod_{i=1}^{n} x_i^{1-n}\right)}_{=\left(\prod_{i=1}^{n} x_i\right)^{1-n}} \underbrace{\sum_{w\in\mathfrak{S}_n} \operatorname{sgn}(w) \prod_{i=1}^{n} x_{w(i)}^{n-i}}_{\substack{=a_\rho \\ \text{(by (13.91.10))}}} \\
&= \frac{1}{\operatorname{sgn}(w_\circ)} \left(\prod_{i=1}^{n} x_i\right)^{1-n} a_\rho.
\end{aligned}$$

Multiplying both sides of this equality by $\operatorname{sgn}(w_\circ)$, we obtain

$$\operatorname{sgn}(w_\circ) \cdot a_\rho\left(x_1^{-1}, x_2^{-1}, \ldots, x_n^{-1}\right) = \left(\prod_{i=1}^{n} x_i\right)^{1-n} a_\rho. \tag{13.91.11}$$

Now, (13.91.3) becomes

$$\begin{aligned}
a_{\rho} \cdot s_{\lambda^{\vee}}\left(\mathbf{x}\right) &= a_{\lambda^{\vee}+\rho} \\
&= \left(\prod_{i=1}^{n} x_i\right)^{n+m-1} \cdot \operatorname{sgn}\left(w_{\circ}\right) \cdot \underbrace{a_{\lambda+\rho}\left(x_1^{-1}, x_2^{-1}, \ldots, x_n^{-1}\right)}_{\substack{=a_{\rho}\left(x_1^{-1}, x_2^{-1}, \ldots, x_n^{-1}\right) \cdot s_{\lambda}\left(x_1^{-1}, x_2^{-1}, \ldots, x_n^{-1}\right) \\ \text{(by (13.91.4))}}} \\
&\qquad \text{(by (13.91.9))} \\
&= \left(\prod_{i=1}^{n} x_i\right)^{n+m-1} \cdot \underbrace{\operatorname{sgn}\left(w_{\circ}\right) \cdot a_{\rho}\left(x_1^{-1}, x_2^{-1}, \ldots, x_n^{-1}\right)}_{\substack{=\left(\prod_{i=1}^{n} x_i\right)^{1-n} a_{\rho} \\ \text{(by (13.91.11))}}} \cdot s_{\lambda}\left(x_1^{-1}, x_2^{-1}, \ldots, x_n^{-1}\right) \\
&= \underbrace{\left(\prod_{i=1}^{n} x_i\right)^{n+m-1} \cdot \left(\prod_{i=1}^{n} x_i\right)^{1-n}}_{=\left(\prod_{i=1}^{n} x_i\right)^{(n+m-1)+(1-n)} = \left(\prod_{i=1}^{n} x_i\right)^{m}} a_{\rho} \cdot s_{\lambda}\left(x_1^{-1}, x_2^{-1}, \ldots, x_n^{-1}\right) \\
&= \left(\underbrace{\prod_{i=1}^{n} x_i}_{=x_1 x_2 \cdots x_n}\right)^{m} a_{\rho} \cdot s_{\lambda}\left(x_1^{-1}, x_2^{-1}, \ldots, x_n^{-1}\right) \\
&= \left(x_1 x_2 \cdots x_n\right)^{m} a_{\rho} \cdot s_{\lambda}\left(x_1^{-1}, x_2^{-1}, \ldots, x_n^{-1}\right) \\
&= a_{\rho} \cdot \left(x_1 x_2 \cdots x_n\right)^{m} \cdot s_{\lambda}\left(x_1^{-1}, x_2^{-1}, \ldots, x_n^{-1}\right) . \qquad (13.91.12)
\end{aligned}$$

But[718]

$$a_{\rho} \text{ is a non-zero-divisor in the ring } \mathbf{k}\left[x_1, x_2, \ldots, x_n, x_1^{-1}, x_2^{-1}, \ldots, x_n^{-1}\right] . \tag{13.91.13}$$

[718] In the following, a *non-zero-divisor* in a commutative ring B means an element $b \in B$ such that every element $c \in B$ satisfying $bc = 0$ must satisfy $c = 0$.

[719] Hence, we can cancel the factor a_ρ from the equality (13.91.12). As a result, we obtain $s_{\lambda^\vee}(\mathbf{x}) = (x_1 x_2 \cdots x_n)^m \cdot s_\lambda\left(x_1^{-1}, x_2^{-1}, \ldots, x_n^{-1}\right)$. Compared with $s_{\lambda^\vee}(\mathbf{x}) = s_{\lambda^\vee}(x_1, x_2, \ldots, x_k)$ (since $\mathbf{x} = (x_1, x_2, \ldots, x_k)$), this yields

$$s_{\lambda^\vee}(x_1, x_2, \ldots, x_k) = (x_1 x_2 \cdots x_n)^m \cdot s_\lambda\left(x_1^{-1}, x_2^{-1}, \ldots, x_n^{-1}\right) = (x_1 x_2 \cdots x_k)^m \cdot s_\lambda\left(x_1^{-1}, x_2^{-1}, \ldots, x_k^{-1}\right)$$

(since $n = k$). This solves Exercise 2.9.15(d).

Second solution to Exercise 2.9.15(d) (sketched): Let us give a more combinatorial solution now. In the following, if α is a partition, then an (α, k)*-CST* will mean a column-strict tableau T of shape α such that all entries of T belong to $\{1, 2, \ldots, k\}$.

We have

$$\underbrace{s_{\lambda^\vee}}_{= s_{\lambda^\vee/\varnothing}}(x_1, x_2, \ldots, x_k) = s_{\lambda^\vee/\varnothing}(x_1, x_2, \ldots, x_k) = \sum_{\substack{T \text{ is a column-strict} \\ \text{tableau of shape } \lambda^\vee/\varnothing; \\ \text{all entries of } T \text{ belong} \\ \text{to } \{1,2,\ldots,k\}}} \mathbf{x}^{\operatorname{cont}(T)}$$

(by Exercise 2.3.8(a), applied to k, $\lambda^\vee$ and $\varnothing$ instead of n, λ and μ). This rewrites as

$$s_{\lambda^\vee}(x_1, x_2, \ldots, x_k) = \sum_{T \text{ is a } (\lambda^\vee, k)\text{-CST}} \mathbf{x}^{\operatorname{cont}(T)} \qquad (13.91.14)$$

(because the column-strict tableaux T of shape $\lambda^\vee/\varnothing$ such that all entries of T belong to $\{1, 2, \ldots, k\}$ are precisely the $(\lambda^\vee, k)$-CSTs).

On the other hand,

$$\underbrace{s_{\lambda}}_{= s_{\lambda/\varnothing}}(x_1, x_2, \ldots, x_k) = s_{\lambda/\varnothing}(x_1, x_2, \ldots, x_k) = \sum_{\substack{T \text{ is a column-strict} \\ \text{tableau of shape } \lambda/\varnothing; \\ \text{all entries of } T \text{ belong} \\ \text{to } \{1,2,\ldots,k\}}} \mathbf{x}^{\operatorname{cont}(T)}$$

[719] *Proof.* We know (from a footnote in Corollary 2.6.7) that a_ρ is not a zero-divisor in the ring $\mathbf{k}[x_1, x_2, \ldots, x_n]$. In other words, a_ρ is a non-zero-divisor in the ring $\mathbf{k}[x_1, x_2, \ldots, x_n]$.

Now, let b be any element of the ring $\mathbf{k}[x_1, x_2, \ldots, x_n]$ such that b is a non-zero-divisor in the ring $\mathbf{k}[x_1, x_2, \ldots, x_n]$. We shall show that b is a non-zero-divisor in the ring $\mathbf{k}\left[x_1, x_2, \ldots, x_n, x_1^{-1}, x_2^{-1}, \ldots, x_n^{-1}\right]$ as well.

Indeed, let $c \in \mathbf{k}\left[x_1, x_2, \ldots, x_n, x_1^{-1}, x_2^{-1}, \ldots, x_n^{-1}\right]$ be such that $bc = 0$. It is known that every element of $\mathbf{k}\left[x_1, x_2, \ldots, x_n, x_1^{-1}, x_2^{-1}, \ldots, x_n^{-1}\right]$ has the form $\dfrac{p}{x_1^{g_1} x_2^{g_2} \cdots x_n^{g_n}}$ for some $p \in \mathbf{k}[x_1, x_2, \ldots, x_n]$ and some $(g_1, g_2, \ldots, g_n) \in \mathbb{N}^n$. So let us write c in this form, and consider the corresponding p and $(g_1, g_2, \ldots, g_n)$. We thus have $c = \dfrac{p}{x_1^{g_1} x_2^{g_2} \cdots x_n^{g_n}}$. Now, $bc = 0$. This rewrites as $b \cdot \dfrac{p}{x_1^{g_1} x_2^{g_2} \cdots x_n^{g_n}} = 0$ (since $c = \dfrac{p}{x_1^{g_1} x_2^{g_2} \cdots x_n^{g_n}}$). Multiplying both sides of this equality by $x_1^{g_1} x_2^{g_2} \cdots x_n^{g_n}$, we obtain $bp = 0$. This equality holds in the ring $\mathbf{k}\left[x_1, x_2, \ldots, x_n, x_1^{-1}, x_2^{-1}, \ldots, x_n^{-1}\right]$, and thus also in the ring $\mathbf{k}[x_1, x_2, \ldots, x_n]$ (since $\mathbf{k}[x_1, x_2, \ldots, x_n]$ is a subring of $\mathbf{k}\left[x_1, x_2, \ldots, x_n, x_1^{-1}, x_2^{-1}, \ldots, x_n^{-1}\right]$). Thus, $p = 0$ (since b is a non-zero-divisor in the ring $\mathbf{k}[x_1, x_2, \ldots, x_n]$, and since $p \in \mathbf{k}[x_1, x_2, \ldots, x_n]$). Hence,

$$c = \frac{p}{x_1^{g_1} x_2^{g_2} \cdots x_n^{g_n}} = \frac{0}{x_1^{g_1} x_2^{g_2} \cdots x_n^{g_n}} \qquad (\text{since } p = 0)$$
$$= 0.$$

Now, let us forget that we fixed c. We thus have proven that every element $c \in \mathbf{k}\left[x_1, x_2, \ldots, x_n, x_1^{-1}, x_2^{-1}, \ldots, x_n^{-1}\right]$ satisfying $bc = 0$ must satisfy $c = 0$. In other words, b is a non-zero-divisor in the ring $\mathbf{k}\left[x_1, x_2, \ldots, x_n, x_1^{-1}, x_2^{-1}, \ldots, x_n^{-1}\right]$.

Now, let us forget that we fixed b. We thus have proven that if b is any element of the ring $\mathbf{k}[x_1, x_2, \ldots, x_n]$ such that b is a non-zero-divisor in the ring $\mathbf{k}[x_1, x_2, \ldots, x_n]$, then b is a non-zero-divisor in the ring $\mathbf{k}\left[x_1, x_2, \ldots, x_n, x_1^{-1}, x_2^{-1}, \ldots, x_n^{-1}\right]$ as well. Applying this to $b = a_\rho$, we conclude that a_ρ is a non-zero-divisor in the ring $\mathbf{k}\left[x_1, x_2, \ldots, x_n, x_1^{-1}, x_2^{-1}, \ldots, x_n^{-1}\right]$ (since we know that a_ρ is a non-zero-divisor in the ring $\mathbf{k}[x_1, x_2, \ldots, x_n]$). This proves (13.91.13).

(by Exercise 2.3.8(a), applied to k and $\varnothing$ instead of n and μ). Substituting x_1^{-1}, x_2^{-1}, ..., x_k^{-1} for x_1, x_2, ..., x_k on both sides of this equality, we obtain

$$s_\lambda\left(x_1^{-1}, x_2^{-1}, \ldots, x_k^{-1}\right) = \sum_{\substack{T \text{ is a column-strict} \\ \text{tableau of shape } \lambda/\varnothing; \\ \text{all entries of } T \text{ belong} \\ \text{to } \{1,2,\ldots,k\}}} \left(\mathbf{x}^{-1}\right)^{\operatorname{cont}(T)}, \tag{13.91.15}$$

where $\left(\mathbf{x}^{-1}\right)^{\operatorname{cont}(T)}$ is defined as $\prod_{i\geq 1}\left(x_i^{-1}\right)^{\left|T^{-1}(i)\right|}$ for any column-strict tableau T. The equality (13.91.15) rewrites as

$$s_\lambda\left(x_1^{-1}, x_2^{-1}, \ldots, x_k^{-1}\right) = \sum_{T \text{ is a } (\lambda,k)\text{-CST}} \left(\mathbf{x}^{-1}\right)^{\operatorname{cont}(T)}$$

(since the column-strict tableaux T of shape $\lambda/\varnothing$ such that all entries of T belong to $\{1, 2, \ldots, k\}$ are precisely the (λ, k)-CSTs). Hence,

$$\begin{aligned} (x_1x_2\cdots x_k)^m \cdot s_\lambda\left(x_1^{-1}, x_2^{-1}, \ldots, x_k^{-1}\right) &= (x_1x_2\cdots x_k)^m \cdot \sum_{T \text{ is a } (\lambda,k)\text{-CST}} \left(\mathbf{x}^{-1}\right)^{\operatorname{cont}(T)} \\ &= \sum_{T \text{ is a } (\lambda,k)\text{-CST}} (x_1x_2\cdots x_k)^m \cdot \left(\mathbf{x}^{-1}\right)^{\operatorname{cont}(T)}. \end{aligned} \tag{13.91.16}$$

We need to prove that

$$s_{\lambda^\vee}(x_1, x_2, \ldots, x_k) = (x_1x_2\cdots x_k)^m \cdot s_\lambda\left(x_1^{-1}, x_2^{-1}, \ldots, x_k^{-1}\right).$$

Due to (13.91.14) and (13.91.16), this rewrites as

$$\sum_{T \text{ is a } (\lambda^\vee,k)\text{-CST}} \mathbf{x}^{\operatorname{cont}(T)} = \sum_{T \text{ is a } (\lambda,k)\text{-CST}} (x_1x_2\cdots x_k)^m \cdot \left(\mathbf{x}^{-1}\right)^{\operatorname{cont}(T)}. \tag{13.91.17}$$

So it remains to prove (13.91.17).

In order to prove (13.91.17), it is clearly sufficient to construct a bijection

$$\Omega : (\text{the set of all } (\lambda, k)\text{-CSTs}) \to (\text{the set of all } (\lambda^\vee, k)\text{-CSTs})$$

with the property that every (λ, k)-CST T satisfies

$$\mathbf{x}^{\operatorname{cont}(\Omega(T))} = (x_1x_2\cdots x_k)^m \cdot \left(\mathbf{x}^{-1}\right)^{\operatorname{cont}(T)}. \tag{13.91.18}$$

Let us construct such a Ω now. We begin by doing some elementary combinatorics.

We define a relation $\leq_\#$ on the set of all subsets of $\{1, 2, \ldots, k\}$ as follows:

Definition 13.91.1. Let k be a nonnegative integer. Let $[k] = \{1, 2, \ldots, k\}$. Let I and J be two subsets of $[k]$. We say that $I \leq_\# J$ if the following two properties hold:

– We have $|I| \geq |J|$.

– For every $r \in \{1, 2, \ldots, |J|\}$, the r-th smallest element of I is $\leq$ to the r-th smallest element of J.

We notice that this relation $\leq_\#$ is the less-or-equal relation of a partial order (as follows easily from the definition); but we will not have any use for this fact. Instead, we need a symmetry property:

Proposition 13.91.2. *Let k be a nonnegative integer. Let $[k] = \{1, 2, \ldots, k\}$. Let I and J be two subsets of $[k]$.*

(a) *We have $I \leq_\# J$ if and only if $[k] \setminus J \leq_\# [k] \setminus I$.*

(b) *For every $\ell \in [k]$ and $S \subset [k]$, let $\alpha_S(\ell)$ denote the number $|\{s \in S \mid s \leq \ell\}|$. Then,*

$$I \leq_\# J \text{ holds if and only if every } \ell \in [k] \text{ satisfies } \alpha_I(\ell) \geq \alpha_J(\ell). \tag{13.91.19}$$

Proof of Proposition 13.91.2. (b) *Proof of (13.91.19):* $\Longrightarrow$: Assume that $I \leq_\# J$. In other words, the following two properties hold:

Property α: We have $|I| \geq |J|$.

Property β: For every $r \in \{1, 2, \ldots, |J|\}$, the r-th smallest element of I is $\leq$ to the r-th smallest element of J.

Now, let $\ell \in [k]$. Then, we need to show that $\alpha_I(\ell) \geq \alpha_J(\ell)$. Since this is obvious if $\alpha_J(\ell) = 0$ (because $\alpha_I(\ell) \geq 0$), we can WLOG assume that $\alpha_J(\ell) \neq 0$. Assume this. Thus, $\alpha_J(\ell) \geq 1$. Also, $\alpha_J(\ell) = \left|\underbrace{\{s \in J \mid s \leq \ell\}}_{\subset J}\right| \leq |J| \leq |I|$ (since $|I| \geq |J|$). Hence, both the $\alpha_J(\ell)$-th smallest element of J and the $\alpha_J(\ell)$-th smallest element of I are well-defined.

Since $\alpha_J(\ell) = |\{s \in J \mid s \leq \ell\}|$, we know that the elements of J which are $\leq \ell$ are precisely the $\alpha_J(\ell)$ smallest elements of J. Thus,

$$(\text{the } \alpha_J(\ell)\text{-th smallest element of } J) = (\text{the largest element of } J \text{ which is } \leq \ell).$$

But by Property β (applied to $r = \alpha_J(\ell)$), we have

$$\begin{aligned}(\text{the } \alpha_J(\ell)\text{-th smallest element of } I) &\leq (\text{the } \alpha_J(\ell)\text{-th smallest element of } J) \\ &= (\text{the largest element of } J \text{ which is } \leq \ell) \leq \ell.\end{aligned}$$

Hence, there are at least $\alpha_J(\ell)$ elements of I which are $\leq \ell$ (namely, the $\alpha_J(\ell)$ smallest ones). In other words, $|\{s \in I \mid s \leq \ell\}| \geq \alpha_J(\ell)$. Now, $\alpha_I(\ell) = |\{s \in I \mid s \leq \ell\}| \geq \alpha_J(\ell)$. We thus have proven the $\Longrightarrow$ direction of (13.91.19).

$\Longleftarrow$: Assume that every $\ell \in [k]$ satisfies $\alpha_I(\ell) \geq \alpha_J(\ell)$. We need to prove that $I \leq_{\#} J$. In other words, we need to prove that the following two properties hold:

Property α: We have $|I| \geq |J|$.

Property β: For every $r \in \{1, 2, \ldots, |J|\}$, the r-th smallest element of I is $\leq$ to the r-th smallest element of J.

First of all, $\{s \in I \mid s \leq k\} = I$ (since every $s \in I$ satisfies $s \leq k$), and the definition of $\alpha_I(k)$ yields $\alpha_I(k) = \left|\underbrace{\{s \in I \mid s \leq k\}}_{=I}\right| = |I|$. Similarly, $\alpha_J(k) = |J|$. Applying $\alpha_I(\ell) \geq \alpha_J(\ell)$ to $\ell = k$, we obtain $\alpha_I(k) \geq \alpha_J(k)$, so that $|I| = \alpha_I(k) \geq \alpha_J(k) = |J|$, and thus Property α is proven.

Now, let $r \in \{1, 2, \ldots, |J|\}$. The r-th smallest element of I and the r-th smallest element of J are then well-defined (because of $r \leq |J| \leq |I|$). Let ℓ be the r-th smallest element of J. Then, $\{s \in J \mid s \leq \ell\}$ is the set consisting of the r smallest elements of J, so that $|\{s \in J \mid s \leq \ell\}| = r$. Now, $\alpha_J(\ell) = |\{s \in J \mid s \leq \ell\}| = r$.

But $\alpha_I(\ell) = |\{s \in I \mid s \leq \ell\}|$, so that

$$|\{s \in I \mid s \leq \ell\}| = \alpha_I(\ell) \geq \alpha_J(\ell) = r.$$

In other words, there exist at least r elements of I which are $\leq \ell$. Hence, the r-th smallest element of I must be $\leq \ell$. Since ℓ is the r-th smallest element of J, this rewrites as follows: The r-th smallest element of I is $\leq$ to the r-th smallest element of J. Thus, Property β holds. Now we know that both Properties α and β hold. Hence, $I \leq_{\#} J$ holds (which, as we know, is equivalent to the conjunction of said properties). This proves the $\Longleftarrow$ direction of (13.91.19). Thus, (13.91.19) is proven. In other words, Proposition 13.91.2(b) is proven.

(a) For every $\ell \in [k]$ and $S \subset [k]$, let $\alpha_S(\ell)$ denote the number $|\{s \in S \mid s \leq \ell\}|$. Thus, every $\ell \in [k]$ satisfies

$$\begin{aligned}\alpha_I(\ell) + \alpha_{[k]\setminus I}(\ell) &= |\{s \in I \mid s \leq \ell\}| + |\{s \in [k] \setminus I \mid s \leq \ell\}| \\ &= \left|\left\{s \in \underbrace{I \cup ([k] \setminus I)}_{=[k]} \mid s \leq \ell\right\}\right| \qquad (\text{since } I \text{ and } [k] \setminus I \text{ are disjoint}) \\ &= |\{s \in [k] \mid s \leq \ell\}| = \ell,\end{aligned}$$

so that $\alpha_{[k]\setminus I}(\ell) = \ell - \alpha_I(\ell)$. Similarly, every $\ell \in [k]$ satisfies $\alpha_{[k]\setminus J}(\ell) = \ell - \alpha_J(\ell)$.

Applying (13.91.19) to $[k] \setminus J$ and $[k] \setminus I$ in lieu of I and J, we obtain that

$$[k] \setminus J \leq_{\#} [k] \setminus I \text{ holds if and only if every } \ell \in [k] \text{ satisfies } \alpha_{[k]\setminus J}(\ell) \geq \alpha_{[k]\setminus I}(\ell). \tag{13.91.20}$$

Now, we have the following equivalence of assertions:

$$\begin{aligned}
&\left([k] \setminus J \leq_{\#} [k] \setminus I\right) \\
\iff &\left(\text{every } \ell \in [k] \text{ satisfies } \underbrace{\alpha_{[k]\setminus J}(\ell)}_{=\ell - \alpha_J(\ell)} \geq \underbrace{\alpha_{[k]\setminus I}(\ell)}_{=\ell-\alpha_I(\ell)}\right) \qquad (\text{by (13.91.20)}) \\
\iff &\left(\text{every } \ell \in [k] \text{ satisfies } \ell - \alpha_J(\ell) \geq \ell - \alpha_I(\ell)\right) \\
\iff &\left(\text{every } \ell \in [k] \text{ satisfies } \alpha_I(\ell) \geq \alpha_J(\ell)\right) \\
\iff &\left(I \leq_{\#} J\right) \qquad (\text{by (13.91.19)}).
\end{aligned}$$

This proves Proposition 13.91.2(a).

Returning to the solution of Exercise 2.9.15(d), we now define some more notations.

If T is a column-strict tableau and i is an integer, then the *i-th set column* of T will mean the set of the entries in the i-th column of T. Notice that the cardinality of the i-th set column of T is the length of the i-th column of the shape of T (since every column of a column-strict tableau has all its entries distinct), and that the i-th column of T can be uniquely reconstructed from the i-th set column of T (because the order of the entries in a column can only be increasing).

For every subset S of $[k]$, we define $\mathbf{x}_S$ to be the monomial $\prod_{s\in S} x_s$ in $\mathbf{k}[x_1, x_2, \ldots, x_k]$. If $(S_1, S_2, \ldots, S_m)$ is an m-tuple of subsets of $[k]$, then we set $\mathbf{x}_{(S_1,S_2,\ldots,S_m)} = \prod_{i=1}^m \mathbf{x}_{S_i}$.

Let λ^t denote the conjugate of the partition λ. Then, $(\lambda^t)_i$ is the length of the i-th column of the Ferrers diagram of λ for every $i \in \{1, 2, \ldots, m\}$.

Let $A(\lambda)$ denote the set of all m-tuples $(I_1, I_2, \ldots, I_m)$ of subsets of $[k]$ satisfying $I_1 \leq_{\#} I_2 \leq_{\#} \cdots \leq_{\#} I_m$ and $\left(|I_i| = (\lambda^t)_i \text{ for every } i \in \{1, 2, \ldots, m\}\right)$. Note that we denote it by $A(\lambda)$ to stress its dependency on λ.

Let $(\lambda^\vee)^t$ denote the conjugate of the partition $\lambda^\vee$. It is easy to see that

$$\left((\lambda^\vee)^t\right)_i = k - \left(\lambda^t\right)_{m+1-i} \qquad \text{for every } i \in \{1, 2, \ldots, m\}. \tag{13.91.21}$$

720

[720] *Proof of (13.91.21):* Fix $i \in \{1, 2, \ldots, m\}$. We recall that $\lambda^\vee = (m - \lambda_k, m - \lambda_{k-1}, \ldots, m - \lambda_1)$. Hence,

$$\left(\lambda^\vee\right)_j = m - \lambda_{k+1-j} \qquad \text{for every } j \in \{1, 2, \ldots, k\}. \tag{13.91.22}$$

Also, $\ell(\lambda^\vee) \leq k$ (since $\lambda^\vee = (m - \lambda_k, m - \lambda_{k-1}, \ldots, m - \lambda_1)$), so that every positive integer $j > k$ satisfies $(\lambda^\vee)_j = 0$. Thus, every positive integer j satisfying $(\lambda^\vee)_j \geq i$ must belong to $\{1, 2, \ldots, k\}$ (since otherwise, this j would satisfy $j > k$, and thus $(\lambda^\vee)_j = 0$, which would contradict $(\lambda^\vee)_j \geq i > 0$). Hence,

$$\begin{aligned}
\left\{j \mid (\lambda^\vee)_j \geq i\right\} &= \left\{ j \in \{1, 2, \ldots, k\} \mid \underbrace{(\lambda^\vee)_j}_{\substack{=m-\lambda_{k+1-j} \\ (\text{by (13.91.22)})}} \geq i \right\} = \left\{ j \in \{1, 2, \ldots, k\} \mid \underbrace{m - \lambda_{k+1-j} \geq i}_{\substack{\text{this is equivalent to } \lambda_{k+1-j} \leq m-i, \text{ and thus} \\ \text{also equivalent to } \lambda_{k+1-j} < m-i+1 \\ (\text{since } \lambda_{k+1-j} \text{ and } m-i \text{ are integers})}} \right\} \\
&= \left\{ j \in \{1, 2, \ldots, k\} \mid \lambda_{k+1-j} < m - i + 1 \right\} \\
&= \{1, 2, \ldots, k\} \setminus \left\{ j \in \{1, 2, \ldots, k\} \mid \lambda_{k+1-j} \geq m - i + 1 \right\}.
\end{aligned}$$

But applying (2.2.7) to $\lambda^\vee$ instead of λ, we obtain

$$\begin{aligned}
\left((\lambda^\vee)^t\right)_i &= \left| \underbrace{\left\{ j \mid (\lambda^\vee)_j \geq i \right\}}_{=\{1,2,\ldots,k\} \setminus \left\{j \in \{1,2,\ldots,k\} \mid \lambda_{k+1-j} \geq m-i+1\right\}} \right| = \left| \{1, 2, \ldots, k\} \setminus \left\{ j \in \{1, 2, \ldots, k\} \mid \lambda_{k+1-j} \geq m - i + 1 \right\} \right| \\
&= \underbrace{\left|\{1, 2, \ldots, k\}\right|}_{=k} - \underbrace{\left| \left\{ j \in \{1, 2, \ldots, k\} \mid \lambda_{k+1-j} \geq m - i + 1 \right\} \right|}_{\substack{=\left|\left\{j \in \{1,2,\ldots,k\} \mid \lambda_j \geq m-i+1\right\}\right| \\ (\text{here, we have substituted } j \text{ for } k+1-j)}} \\
&= k - \left| \left\{ j \in \{1, 2, \ldots, k\} \mid \lambda_j \geq m - i + 1 \right\} \right|. \qquad (13.91.23)
\end{aligned}$$

But $\ell(\lambda) \leq k$, so that every positive integer $j > k$ satisfies $\lambda_j = 0$. Thus, every positive integer j satisfying $\lambda_j \geq m-i+1$ must belong to $\{1, 2, \ldots, k\}$ (since otherwise, this j would satisfy $j > k$, and thus $\lambda_j = 0$, which would contradict $\lambda_j \geq m-i+1 > 0$).

It is easy to see (from the definition of $\lambda^\vee$) that $\lambda^\vee$ is a partition satisfying $\ell(\lambda^\vee) \leq k$, and all parts of $\lambda^\vee$ are $\leq m$. Hence, we can define a set $A(\lambda^\vee)$ in the same way as we defined the set $A(\lambda)$ (but with every λ replaced by $\lambda^\vee$). Explicitly, $A(\lambda^\vee)$ is the set of all m-tuples $(I_1, I_2, \ldots, I_m)$ of subsets of $[k]$ satisfying $I_1 \leq_\# I_2 \leq_\# \cdots \leq_\# I_m$ and $\left(|I_i| = \left((\lambda^\vee)^t\right)_i \text{ for every } i \in \{1, 2, \ldots, m\}\right)$.

Define a map

$$\varphi(\lambda) : (\text{the set of all } (\lambda, k)\text{-CSTs}) \to A(\lambda)$$

by sending every (λ, k)-CST T to the m-tuple whose i-th entry is the i-th set column of T. This map is well-defined[721], injective[722] and surjective[723]. Hence, $\varphi(\lambda)$ is a bijection. This bijection $\varphi(\lambda)$ furthermore satisfies

$$\mathbf{x}_{(\varphi(\lambda))(T)} = \mathbf{x}^{\operatorname{cont}(T)} \qquad \text{for every } (\lambda, k)\text{-CST } T \tag{13.91.25}$$

[724]. Substituting x_1^{-1}, x_2^{-1}, $\ldots$, x_k^{-1} for x_1, x_2, $\ldots$, x_k on both sides of this equality, we obtain

$$\left(\mathbf{x}_{(\varphi(\lambda))(T)}\right)^{-1} = \left(\mathbf{x}^{-1}\right)^{\operatorname{cont}(T)} \qquad \text{for every } (\lambda, k)\text{-CST } T \tag{13.91.26}$$

(in fact, $\mathbf{x}_{(\varphi(\lambda))(T)}$ is a monomial, so that substituting x_1^{-1}, x_2^{-1}, $\ldots$, x_k^{-1} for x_1, x_2, $\ldots$, x_k transforms it into its inverse $\left(\mathbf{x}_{(\varphi(\lambda))(T)}\right)^{-1}$).

Recall that $\lambda^\vee$ is a partition satisfying $\ell(\lambda^\vee) \leq k$, and all parts of $\lambda^\vee$ are $\leq m$. Hence, we can define a bijection

$$\varphi(\lambda^\vee) : (\text{the set of all } (\lambda^\vee, k)\text{-CSTs}) \to A(\lambda^\vee)$$

in the same way as we defined the bijection $\varphi(\lambda) : (\text{the set of all } (\lambda, k)\text{-CSTs}) \to A(\lambda)$ (but with $\lambda^\vee$ instead of λ). This bijection satisfies

$$\mathbf{x}_{(\varphi(\lambda^\vee))(T)} = \mathbf{x}^{\operatorname{cont}(T)} \qquad \text{for every } (\lambda^\vee, k)\text{-CST } T \tag{13.91.27}$$

Hence,

$$\{j \mid \lambda_j \geq m - i + 1\} = \{j \in \{1, 2, \ldots, k\} \mid \lambda_j \geq m - i + 1\}.$$

Now, (2.2.7) (applied to $m - i + 1$ instead of i) yields

$$\left(\lambda^t\right)_{m-i+1} = \left| \underbrace{\{j \mid \lambda_j \geq m - i + 1\}}_{= \{j \in \{1,2,\ldots,k\} \mid \lambda_j \geq m-i+1\}} \right| = \left|\{j \in \{1, 2, \ldots, k\} \mid \lambda_j \geq m - i + 1\}\right|. \tag{13.91.24}$$

Thus, (13.91.23) becomes

$$\left((\lambda^\vee)^t\right)_i = k - \underbrace{\left|\{j \in \{1, 2, \ldots, k\} \mid \lambda_j \geq m - i + 1\}\right|}_{\substack{= (\lambda^t)_{m-i+1} \\ \text{(by (13.91.24))}}} = k - \left(\lambda^t\right)_{m-i+1} = k - \left(\lambda^t\right)_{m+1-i}.$$

This proves (13.91.21).

[721] Indeed, if T is an (λ, k)-CST, then the m-tuple whose i-th entry is the i-th set column of T belongs to $A(\lambda)$ (because if we denote this m-tuple by $(I_1, I_2, \ldots, I_m)$, then

$$\begin{aligned} |I_i| &= |(\text{the } i\text{-th set column of } T)| = (\text{the length of the } i\text{-th column of } T) \\ &= (\text{the length of the } i\text{-th column of the Ferrers diagram of } \lambda) \\ &\qquad (\text{since } T \text{ is a } (\lambda, k)\text{-CST, thus a column-strict tableau of shape } \lambda) \\ &= \left(\lambda^t\right)_i, \end{aligned}$$

and the fact that the entries of T increase weakly along rows (because T is a column-strict tableau) translates precisely into the inequality chain $I_1 \leq_\# I_2 \leq_\# \cdots \leq_\# I_m$).

[722] This is because the i-th column of a (λ, k)-CST T can be uniquely reconstructed from the i-th set column of T (indeed, the entries of the i-th column of T must be strictly increasing down the column, and therefore the knowledge of the set of these entries is sufficient to recover the i-th column). Here, we are using the fact that T has at most m columns (since every part of λ is $\leq m$).

[723] Indeed, given any $(I_1, I_2, \ldots, I_m) \in A(\lambda)$. Then, $(I_1, I_2, \ldots, I_m)$ is an m-tuple of subsets of $[k]$ satisfying $I_1 \leq_\# I_2 \leq_\# \cdots \leq_\# I_m$ and $\left(|I_i| = \left(\lambda^t\right)_i \text{ for every } i \in \{1, 2, \ldots, m\}\right)$. Now, we can fill in each column of the Ferrers diagram of λ with the entries of the corresponding set I_i in increasing order, and then the resulting filling is a column-strict tableau (indeed, its entries increase weakly along its rows (due to $I_1 \leq_\# I_2 \leq_\# \cdots \leq_\# I_m$)), and more precisely a (λ, k)-CST. Our m-tuple $(I_1, I_2, \ldots, I_m)$ is the image of this (λ, k)-CST under $\varphi(\lambda)$; therefore, $(I_1, I_2, \ldots, I_m)$ lies in the image of $\varphi(\lambda)$. Hence, we have shown that every $(I_1, I_2, \ldots, I_m) \in A(\lambda)$ lies in the image of $\varphi(\lambda)$. The map $\varphi(\lambda)$ is thus surjective, qed.

[724] Indeed, it is easy to check that both $\mathbf{x}_{(\varphi(\lambda))(T)}$ and $\mathbf{x}^{\operatorname{cont}(T)}$ are equal to the product $\prod_{c \text{ is a cell of } T} x_{(\text{entry of } T \text{ in } c)}$.

(this can be shown in analogy to (13.91.25)).

Finally, we define a map $\psi(\lambda) : A(\lambda) \to A(\lambda^\vee)$ as follows: For every $(I_1, I_2, \ldots, I_m) \in A(\lambda)$, let

$$(\psi(\lambda))(I_1, I_2, \ldots, I_m) = ([k] \setminus I_m, [k] \setminus I_{m-1}, \ldots, [k] \setminus I_1).$$

This map $\psi(\lambda)$ is well-defined[725] and bijective[726]. This bijective map $\psi(\lambda)$ has the property that

$$\mathbf{x}_{(\psi(\lambda))(S)} = (x_1 x_2 \cdots x_k)^m \cdot (\mathbf{x}_S)^{-1} \qquad \text{for every } S \in A(\lambda) \tag{13.91.28}$$

[727].

Using the three bijective maps

$$\varphi(\lambda) : (\text{the set of all } (\lambda, k)\text{-CSTs}) \to A(\lambda),$$
$$\psi(\lambda) : A(\lambda) \to A(\lambda^\vee),$$
$$\varphi(\lambda^\vee) : (\text{the set of all } (\lambda^\vee, k)\text{-CSTs}) \to A(\lambda^\vee),$$

we can define a map

$$\Omega : (\text{the set of all } (\lambda, k)\text{-CSTs}) \to (\text{the set of all } (\lambda^\vee, k)\text{-CSTs})$$

by $\Omega = (\varphi(\lambda^\vee))^{-1} \circ (\psi(\lambda)) \circ (\varphi(\lambda))$. Clearly, this Ω is a bijection. If we can show that every (λ, k)-CST T satisfies (13.91.18), then (13.91.17) will be proven, and thus Exercise 2.9.15(d) will be solved again. Hence, all that remains to prove now is that every (λ, k)-CST T satisfies (13.91.18).

[725] *Proof.* Let $(I_1, I_2, \ldots, I_m) \in A(\lambda)$. Then, $(I_1, I_2, \ldots, I_m)$ is an m-tuple of subsets of $[k]$ satisfying $I_1 \leq_\# I_2 \leq_\# \cdots \leq_\# I_m$ and $\left(|I_i| = (\lambda^t)_i \text{ for every } i \in \{1, 2, \ldots, m\}\right)$.

In order to prove the well-definedness of $\psi(\lambda)$, we need to show that $([k] \setminus I_m, [k] \setminus I_{m-1}, \ldots, [k] \setminus I_1) \in A(\lambda^\vee)$. In other words, we need to prove that $([k] \setminus I_m, [k] \setminus I_{m-1}, \ldots, [k] \setminus I_1)$ is an m-tuple of subsets of $[k]$ satisfying $[k] \setminus I_m \leq_\# [k] \setminus I_{m-1} \leq_\# \cdots \leq_\# [k] \setminus I_1$ and $\left(|[k] \setminus I_{m+1-i}| = \left((\lambda^\vee)^t\right)_i \text{ for every } i \in \{1, 2, \ldots, m\}\right)$.

First of all, it is clear that $([k] \setminus I_m, [k] \setminus I_{m-1}, \ldots, [k] \setminus I_1)$ is an m-tuple of subsets of $[k]$. Furthermore, $[k] \setminus I_m \leq_\# [k] \setminus I_{m-1} \leq_\# \cdots \leq_\# [k] \setminus I_1$ follows from $I_1 \leq_\# I_2 \leq_\# \cdots \leq_\# I_m$ (according to Proposition 13.91.2(a)). Thus, it remains to prove that $|[k] \setminus I_{m+1-i}| = \left((\lambda^\vee)^t\right)_i$ for every $i \in \{1, 2, \ldots, m\}$.

So fix some $i \in \{1, 2, \ldots, m\}$. Then, $|I_{m+1-i}| = (\lambda^t)_{m+1-i}$ (this follows from the $|I_i| = (\lambda^t)_i$ formula, but applied to $m + 1 - i$ instead of i). Since $I_{m+1-i} \subset [k]$, we have

$$|[k] \setminus I_{m+1-i}| = \underbrace{|[k]|}_{=k} - \underbrace{|I_{m+1-i}|}_{=(\lambda^t)_{m+1-i}} = k - (\lambda^t)_{m+1-i} = \left((\lambda^\vee)^t\right)_i \qquad (\text{by } (13.91.21)),$$

qed.

[726] Indeed, we can define a map $\psi(\lambda^\vee)$ in the same way as we have defined $\psi(\lambda)$ (but with $\lambda^\vee$ instead of λ). It is then easy to see that the maps $\psi(\lambda)$ and $\psi(\lambda^\vee)$ are mutually inverse, so that $\psi(\lambda)$ is bijective, qed.

[727] *Proof of (13.91.28):* Let $S \in A(\lambda)$. Then, S is an m-tuple of subsets of $[k]$. Write S as $S = (I_1, I_2, \ldots, I_m)$. The definition of $\mathbf{x}_S$ then yields

$$\mathbf{x}_S = \prod_{i=1}^m \mathbf{x}_{I_i}.$$

Hence,

$$(x_1 x_2 \cdots x_k)^m \cdot (\mathbf{x}_S)^{-1} = (x_1 x_2 \cdots x_k)^m \cdot \left(\prod_{i=1}^m \mathbf{x}_{I_i}\right)^{-1} = \prod_{i=1}^m \frac{x_1 x_2 \cdots x_k}{\mathbf{x}_{I_i}}. \tag{13.91.29}$$

On the other hand, the definition of $\psi(\lambda)$ yields $(\psi(\lambda))(S) = ([k] \setminus I_m, [k] \setminus I_{m-1}, \ldots, [k] \setminus I_1)$, and thus (by the definition of $\mathbf{x}_{(\psi(\lambda))(S)}$) we have

$$\mathbf{x}_{(\psi(\lambda))(S)} = \prod_{i=1}^m \mathbf{x}_{[k] \setminus I_{m+1-i}} = \prod_{i=1}^m \mathbf{x}_{[k] \setminus I_i} \qquad (\text{here, we substituted } m + 1 - i \text{ for } i \text{ in the product}). \tag{13.91.30}$$

But we need to prove (13.91.28). In view of (13.91.29) and (13.91.30), this boils down to proving that $\prod_{i=1}^m \mathbf{x}_{[k] \setminus I_i} = \prod_{i=1}^m \frac{x_1 x_2 \cdots x_k}{\mathbf{x}_{I_i}}$. But this will immediately follow if we can prove the identity $\mathbf{x}_{[k] \setminus I_i} = \frac{x_1 x_2 \cdots x_k}{\mathbf{x}_{I_i}}$ for every $i \in \{1, 2, \ldots, m\}$. But the latter identity follows from the (obvious) fact that $\mathbf{x}_{[k] \setminus S} = \frac{x_1 x_2 \cdots x_k}{\mathbf{x}_S}$ for every $S \subset [k]$. Thus, (13.91.28) is proven.

Let T be a (λ, k)-CST. We have $\Omega = \left(\varphi\left(\lambda^\vee\right)\right)^{-1} \circ \left(\psi\left(\lambda\right)\right) \circ \left(\varphi\left(\lambda\right)\right)$, thus $\left(\varphi\left(\lambda^\vee\right)\right) \circ \Omega = \left(\psi\left(\lambda\right)\right) \circ \left(\varphi\left(\lambda\right)\right)$. Applying (13.91.27) to $\Omega\left(T\right)$ instead of T, we obtain $\mathbf{x}_{\left(\varphi\left(\lambda^\vee\right)\right)\left(\Omega\left(T\right)\right)} = \mathbf{x}^{\operatorname{cont}\left(\Omega\left(T\right)\right)}$, whence

$$\begin{aligned}
\mathbf{x}^{\operatorname{cont}\left(\Omega\left(T\right)\right)} &= \mathbf{x}_{\left(\varphi\left(\lambda^\vee\right)\right)\left(\Omega\left(T\right)\right)} = \mathbf{x}_{\left(\psi\left(\lambda\right)\right)\left(\left(\varphi\left(\lambda\right)\right)\left(T\right)\right)} \\
&\qquad \left(\text{since } \left(\varphi\left(\lambda^\vee\right)\right)\left(\Omega\left(T\right)\right) = \underbrace{\left(\left(\varphi\left(\lambda^\vee\right)\right) \circ \Omega\right)}_{=\left(\psi\left(\lambda\right)\right)\circ\left(\varphi\left(\lambda\right)\right)}\left(T\right) = \left(\left(\psi\left(\lambda\right)\right) \circ \left(\varphi\left(\lambda\right)\right)\right)\left(T\right) = \left(\psi\left(\lambda\right)\right)\left(\left(\varphi\left(\lambda\right)\right)\left(T\right)\right)\right) \\
&= \left(x_1 x_2 \cdots x_k\right)^m \cdot \underbrace{\left(\mathbf{x}_{\left(\varphi\left(\lambda\right)\right)\left(T\right)}\right)^{-1}}_{\substack{=\left(\mathbf{x}^{-1}\right)^{\operatorname{cont}\left(T\right)} \\ \text{(by (13.91.26))}}} \qquad \left(\text{by (13.91.28), applied to } S = \left(\varphi\left(\lambda\right)\right)\left(T\right)\right) \\
&= \left(x_1 x_2 \cdots x_k\right)^m \cdot \left(\mathbf{x}^{-1}\right)^{\operatorname{cont}\left(T\right)}.
\end{aligned}$$

Thus, (13.91.18) holds. We thus have proven that every (λ, k)-CST T satisfies (13.91.18). Exercise 2.9.15(d) is thus solved.

(e) *First solution to Exercise 2.9.15(e):* Obviously, the k-tuple $\left(r+\lambda_1, r+\lambda_2, \ldots, r+\lambda_k\right)$ is a partition. Let us denote this partition by $\lambda^{[r]}$. Then, $\lambda^{[r]} = \left(r+\lambda_1, r+\lambda_2, \ldots, r+\lambda_k\right)$ and $\ell\left(\lambda^{[r]}\right) \leq k$.

Let $n = k$. We shall use the notations of Section 2.6; in particular, we set $\mathbf{x} = \left(x_1, x_2, \ldots, x_n\right)$ and $\rho = \left(n-1, n-2, \ldots, 2, 1, 0\right)$. Clearly, $\mathbf{x} = \left(x_1, x_2, \ldots, x_n\right) = \left(x_1, x_2, \ldots, x_k\right)$ (since $n = k$). Notice that $\ell\left(\lambda\right) \leq k = n$. Hence, we can regard λ as an element of $\mathbb{N}^n$; therefore, $\lambda + \rho$ is a well-defined element of $\mathbb{N}^n$, and the alternant $a_{\lambda+\rho}$ is well-defined. Corollary 2.6.7 yields that $s_\lambda\left(\mathbf{x}\right) = \dfrac{a_{\lambda+\rho}}{a_\rho}$, so that

$$a_\rho \cdot s_\lambda\left(\mathbf{x}\right) = a_{\lambda+\rho}. \tag{13.91.31}$$

Applying this equality to $\lambda^{[r]}$ instead of λ, we obtain

$$a_\rho \cdot s_{\lambda^{[r]}}\left(\mathbf{x}\right) = a_{\lambda^{[r]}+\rho} \tag{13.91.32}$$

(since $\lambda^{[r]}$ is also a partition satisfying $\ell\left(\lambda^{[r]}\right) \leq k = n$).

We have (13.91.7), and every $\alpha = \left(\alpha_1, \alpha_2, \ldots, \alpha_n\right) \in \mathbb{N}^n$ satisfies (13.91.6). (This can be proven just as in the First solution to Exercise 2.9.15(d).)

We have $\lambda^{[r]} = \left(r+\lambda_1, r+\lambda_2, \ldots, r+\lambda_k\right) = \left(r+\lambda_1, r+\lambda_2, \ldots, r+\lambda_n\right)$ (since $k = n$), so that

$$\begin{aligned}
\underbrace{\lambda^{[r]}}_{=\left(r+\lambda_1, r+\lambda_2, \ldots, r+\lambda_n\right)} + \underbrace{\rho}_{\substack{=\left(n-1, n-2, \ldots, 2, 1, 0\right) \\ =\left(n-1, n-2, \ldots, n-n\right)}} &= \left(r+\lambda_1, r+\lambda_2, \ldots, r+\lambda_n\right) + \left(n-1, n-2, \ldots, n-n\right) \\
&= \left(\left(r+\lambda_1\right) + \left(n-1\right), \left(r+\lambda_2\right) + \left(n-2\right), \ldots, \left(r+\lambda_n\right) + \left(n-n\right)\right).
\end{aligned}$$

Hence, (13.91.6) (applied to $\lambda^{[r]} + \rho$ and $((r+\lambda_1)+(n-1), (r+\lambda_2)+(n-2), \ldots, (r+\lambda_n)+(n-n))$ instead of α and $(\alpha_1, \alpha_2, \ldots, \alpha_n)$) yields

$$\begin{aligned}
& a_{\lambda^{[r]}+\rho} \\
&= \sum_{w\in\mathfrak{S}_n} \operatorname{sgn}(w) \prod_{i=1}^{n} \underbrace{x_{w(i)}^{(r+\lambda_i)+(n-i)}}_{=x_{w(i)}^{r+(\lambda_i+n-i)}=x_{w(i)}^{r}x_{w(i)}^{\lambda_i+n-i}} \\
&= \sum_{w\in\mathfrak{S}_n} \operatorname{sgn}(w) \underbrace{\prod_{i=1}^{n}\left(x_{w(i)}^{r}x_{w(i)}^{\lambda_i+n-i}\right)}_{=\left(\prod_{i=1}^{n} x_{w(i)}^{r}\right)\left(\prod_{i=1}^{n} x_{w(i)}^{\lambda_i+n-i}\right)} = \sum_{w\in\mathfrak{S}_n} \operatorname{sgn}(w) \underbrace{\left(\prod_{i=1}^{n} x_{w(i)}^{r}\right)}_{\substack{=\prod_{i=1}^{n} x_i^r \\ \text{(here, we have substituted } i \text{ for} \\ w(i) \text{ in the product (since } w \text{ is} \\ \text{a permutation of } \{1,2,\ldots,n\}))}} \left(\prod_{i=1}^{n} x_{w(i)}^{\lambda_i+n-i}\right) \\
&= \sum_{w\in\mathfrak{S}_n} \operatorname{sgn}(w) \left(\prod_{i=1}^{n} x_i^r\right)\left(\prod_{i=1}^{n} x_{w(i)}^{\lambda_i+n-i}\right) = \underbrace{\left(\prod_{i=1}^{n} x_i^r\right)}_{=x_1^r x_2^r \cdots x_n^r=(x_1x_2\cdots x_n)^r} \cdot \underbrace{\sum_{w\in\mathfrak{S}_n} \operatorname{sgn}(w) \prod_{i=1}^{n} x_{w(i)}^{\lambda_i+n-i}}_{\substack{=a_{\lambda+\rho} \\ \text{(by (13.91.7))}}} \\
&= (x_1x_2\cdots x_n)^r \cdot \underbrace{a_{\lambda+\rho}}_{\substack{=a_\rho \cdot s_\lambda(\mathbf{x}) \\ \text{(by (13.91.31))}}} = (x_1x_2\cdots x_n)^r \cdot a_\rho \cdot s_\lambda(\mathbf{x}) = a_\rho \cdot (x_1x_2\cdots x_n)^r \cdot s_\lambda(\mathbf{x}).
\end{aligned}$$

Hence, $a_\rho \cdot (x_1x_2\cdots x_n)^r \cdot s_\lambda(\mathbf{x}) = a_{\lambda^{[r]}+\rho} = a_\rho \cdot s_{\lambda^{[r]}}(\mathbf{x})$ (by (13.91.31)).

But we know (from a footnote in Corollary 2.6.7) that a_ρ is not a zero-divisor in the ring $\mathbf{k}[x_1, x_2, \ldots, x_n]$. In other words, a_ρ is a non-zero-divisor in the ring $\mathbf{k}[x_1, x_2, \ldots, x_n]$. Hence, we can cancel the factor a_ρ from the equality $a_\rho \cdot (x_1x_2\cdots x_n)^r \cdot s_\lambda(\mathbf{x}) = a_\rho \cdot s_{\lambda^{[r]}}(\mathbf{x})$. As a result, we obtain $(x_1x_2\cdots x_n)^r \cdot s_\lambda(\mathbf{x}) = s_{\lambda^{[r]}}(\mathbf{x}) = s_{\lambda^{[r]}}(x_1, x_2, \ldots, x_k)$ (since $\mathbf{x} = (x_1, x_2, \ldots, x_k)$), so that

$$s_{\lambda^{[r]}}(x_1, x_2, \ldots, x_k) = \Bigg(\underbrace{x_1x_2\cdots x_n}_{\substack{=x_1x_2\cdots x_k \\ \text{(since } n=k)}}\Bigg)^r \cdot \underbrace{s_\lambda(\mathbf{x})}_{\substack{=s_\lambda(x_1,x_2,\ldots,x_k) \\ \text{(since } \mathbf{x}=(x_1,x_2,\ldots,x_k))}} = (x_1x_2\cdots x_k)^r \cdot s_\lambda(x_1, x_2, \ldots, x_k).$$

Since $\lambda^{[r]} = s_{(r+\lambda_1, r+\lambda_2, \ldots, r+\lambda_k)}$, this rewrites as

$$s_{(r+\lambda_1, r+\lambda_2, \ldots, r+\lambda_k)}(x_1, x_2, \ldots, x_k) = (x_1x_2\cdots x_k)^r \cdot s_\lambda(x_1, x_2, \ldots, x_k).$$

This solves Exercise 2.9.15(e).

Second solution to Exercise 2.9.15(e) (sketched): Obviously, the k-tuple $(r+\lambda_1, r+\lambda_2, \ldots, r+\lambda_k)$ is a partition. Let us denote this partition by $\lambda^{[r]}$. Then, $\lambda^{[r]} = (r+\lambda_1, r+\lambda_2, \ldots, r+\lambda_k)$ and $\ell\left(\lambda^{[r]}\right) \leq k$.

We have $(r+\lambda_1, r+\lambda_2, \ldots, r+\lambda_k) = \lambda^{[r]}$, thus $s_{(r+\lambda_1, r+\lambda_2, \ldots, r+\lambda_k)} = s_{\lambda^{[r]}} = s_{\lambda^{[r]}/\varnothing}$, hence

$$\begin{aligned}
& \underbrace{s_{(r+\lambda_1, r+\lambda_2, \ldots, r+\lambda_k)}}_{=s_{\lambda^{[r]}/\varnothing}}(x_1, x_2, \ldots, x_k) \\
&= s_{\lambda^{[r]}/\varnothing}(x_1, x_2, \ldots, x_k) = \sum_{\substack{T \text{ is a column-strict} \\ \text{tableau of shape } \lambda^{[r]}/\varnothing; \\ \text{all entries of } T \text{ belong} \\ \text{to } \{1,2,\ldots,k\}}} \mathbf{x}^{\operatorname{cont}(T)} \\
& \qquad \left(\text{by Exercise 2.3.8(a), applied to } k,\ \lambda^{[r]} \text{ and } \varnothing \text{ instead of } n,\ \lambda \text{ and } \mu\right) \\
&= \sum_{\substack{T \text{ is a column-strict} \\ \text{tableau of shape } \lambda^{[r]}; \\ \text{all entries of } T \text{ belong} \\ \text{to } \{1,2,\ldots,k\}}} \mathbf{x}^{\operatorname{cont}(T)} \qquad (13.91.33)
\end{aligned}$$

(since column-strict tableaux of shape $\lambda^{[r]}/\varnothing$ are the same as column-strict tableaux of shape $\lambda^{[r]}$).

But an argument analogous to the one we just used to prove (13.91.33) (but with λ in place of $\lambda^{[r]}$) shows that

$$s_\lambda\left(x_1, x_2, \ldots, x_k\right) = \sum_{\substack{T \text{ is a column-strict} \\ \text{tableau of shape } \lambda; \\ \text{all entries of } T \text{ belong} \\ \text{to } \{1,2,\ldots,k\}}} \mathbf{x}^{\operatorname{cont}(T)}.$$

Multiplied by $\left(x_1 x_2 \cdots x_k\right)^r$, this equality becomes

$$\begin{aligned} \left(x_1 x_2 \cdots x_k\right)^r \cdot s_\lambda\left(x_1, x_2, \ldots, x_k\right) &= \left(x_1 x_2 \cdots x_k\right)^r \cdot \sum_{\substack{T \text{ is a column-strict} \\ \text{tableau of shape } \lambda; \\ \text{all entries of } T \text{ belong} \\ \text{to } \{1,2,\ldots,k\}}} \mathbf{x}^{\operatorname{cont}(T)} \\ &= \sum_{\substack{T \text{ is a column-strict} \\ \text{tableau of shape } \lambda; \\ \text{all entries of } T \text{ belong} \\ \text{to } \{1,2,\ldots,k\}}} \left(x_1 x_2 \cdots x_k\right)^r \cdot \mathbf{x}^{\operatorname{cont}(T)}. \qquad (13.91.34) \end{aligned}$$

We need to prove that

$$s_{(r+\lambda_1, r+\lambda_2, \ldots, r+\lambda_k)}\left(x_1, x_2, \ldots, x_k\right) = \left(x_1 x_2 \cdots x_k\right)^r \cdot s_\lambda\left(x_1, x_2, \ldots, x_k\right).$$

Due to (13.91.33) and (13.91.34), this rewrites as

$$\sum_{\substack{T \text{ is a column-strict} \\ \text{tableau of shape } \lambda^{[r]}; \\ \text{all entries of } T \text{ belong} \\ \text{to } \{1,2,\ldots,k\}}} \mathbf{x}^{\operatorname{cont}(T)} = \sum_{\substack{T \text{ is a column-strict} \\ \text{tableau of shape } \lambda; \\ \text{all entries of } T \text{ belong} \\ \text{to } \{1,2,\ldots,k\}}} \left(x_1 x_2 \cdots x_k\right)^r \cdot \mathbf{x}^{\operatorname{cont}(T)}. \qquad (13.91.35)$$

So it remains to prove (13.91.35).

In order to prove (13.91.35), it is clearly sufficient to construct a bijection

$$\begin{aligned} \Omega : &\left(\text{the set of all column-strict tableaux } T \text{ of shape } \lambda^{[r]} \text{ such that all entries of } T \text{ belong to } \{1, 2, \ldots, k\}\right) \\ &\to \left(\text{the set of all column-strict tableaux } T \text{ of shape } \lambda \text{ such that all entries of } T \text{ belong to } \{1, 2, \ldots, k\}\right) \end{aligned}$$

with the property that every column-strict tableau T of shape $\lambda^{[r]}$ such that all entries of T belong to $\{1, 2, \ldots, k\}$ satisfies

$$\left(x_1 x_2 \cdots x_k\right)^r \cdot \mathbf{x}^{\operatorname{cont}(\Omega(T))} = \mathbf{x}^{\operatorname{cont}(T)}. \qquad (13.91.36)$$

Constructing such an Ω is easy: The map Ω just maps every column-strict tableau T of shape $\lambda^{[r]}$ to the tableau of shape λ obtained by removing the first r entries of each row of T, and moving all other entries to the left by r cells. To prove that Ω is bijective, we need to construct a map inverse to Ω; this latter map sends every column-strict tableau T of shape λ to the tableau of shape $\lambda^{[r]}$ obtained by moving all entries of T to the right by r cells, and filling the now-vacant first r columns as follows:

$$\begin{array}{cccc} 1 & 1 & \cdots & 1 \\ 2 & 2 & \cdots & 2 \\ \vdots & \vdots & \ddots & \vdots \\ k & k & \cdots & k \end{array}.$$

Proving that this map is well-defined and really inverse to Ω is left to the reader[728]. Anyway, we now know that Ω is bijective, and it is easy to check that (13.91.36) holds. This completes the second solution of Exercise 2.9.15(e).

[728] The main ingredient of this proof is the observation that if T is a column-strict tableau of shape $\lambda^{[r]}$ such that all entries of T belong to $\{1, 2, \ldots, k\}$, then the first r columns of T must look like this:

$$\begin{array}{cccc} 1 & 1 & \cdots & 1 \\ 2 & 2 & \cdots & 2 \\ \vdots & \vdots & \ddots & \vdots \\ k & k & \cdots & k \end{array}.$$

13.92. **Solution to Exercise 2.9.16.** *Solution to Exercise 2.9.16.* (a) We can apply Exercise 2.9.15(a) to μ, μ, $\mu^{\vee\{m\}}$ and $\mu^{\vee\{m\}}$ instead of λ, μ, $\lambda^{\vee}$ and $\mu^{\vee}$. As a consequence, we conclude that $\mu^{\vee\{m\}}$ and $\mu^{\vee\{m\}}$ are partitions, and that $s_{\mu/\mu} = s_{\mu^{\vee\{m\}}/\mu^{\vee\{m\}}}$. Thus, $\mu^{\vee\{m\}}$ is a partition. The same argument (but with m, μ and $\mu^{\vee\{m\}}$ replaced by n, ν and $\nu^{\vee\{n\}}$) yields that $\nu^{\vee\{n\}}$ is a partition. This solves Exercise 2.9.16(a).

(b) Assume that not all parts of λ are $\leq m+n$. Then, some part of λ is $> m+n$; hence, the greatest part of λ is $> m+n$. Thus, $\lambda_1 > m+n$ (since λ_1 is the greatest part of λ). As a consequence, $\lambda_1 - \mu_1 > \nu_1$ [729].

Now, we shall show that

$$\left(\begin{array}{c} \text{there exists no column-strict tableau } T \text{ of shape } \lambda/\mu \text{ with } \operatorname{cont}(T) = \nu \\ \text{having the property that each } \operatorname{cont}(T\mid_{\text{cols}\geq j}) \text{ is a partition} \end{array} \right). \tag{13.92.1}$$

Proof of (13.92.1): Assume the contrary. Thus, there exists a column-strict tableau T of shape λ/μ with $\operatorname{cont}(T) = \nu$ having the property that each $\operatorname{cont}(T\mid_{\text{cols}\geq j})$ is a partition. Consider this T.

The column-strict tableau T has shape λ/μ, and thus its 1-st row has $\lambda_1 - \mu_1$ entries. In other words,

$$\begin{aligned} &(\text{the number of entries in the 1-st row of } T) \\ &= \lambda_1 - \mu_1 > \nu_1 = (\operatorname{cont}(T))_1 \qquad (\text{since } \nu = \operatorname{cont}(T)) \\ &= \left|T^{-1}(1)\right| \qquad (\text{by the definition of } \operatorname{cont}(T)) \\ &= (\text{the number of entries of } T \text{ equal to } 1). \end{aligned}$$

Hence, the tableau T has more entries in its 1-st row than it has entries equal to 1. Consequently, not every entry in the 1-st row is equal to 1. Thus, there exists an entry in the 1-st row of T which is > 1. Let k be this entry, and let $\mathbf{c}$ be the cell it occupies. Then, $\mathbf{c}$ is a cell in the 1-st row of T, and thus can be written in the form $(1, j)$ for some positive integer j. Consider this j. The cell $\mathbf{c}$ lies in column j, and thus is a cell of the skew tableau $T\mid_{\text{cols}\geq j}$; its entry is $(T\mid_{\text{cols}\geq j})(\mathbf{c}) = T(\mathbf{c}) = k$ (since we know that k is the entry of T occupying cell $\mathbf{c}$). Hence, $\mathbf{c} \in (T\mid_{\text{cols}\geq j})^{-1}(k)$, so that the set $(T\mid_{\text{cols}\geq j})^{-1}(k)$ is nonempty. Thus, $\left|(T\mid_{\text{cols}\geq j})^{-1}(k)\right| > 0$.

Recalling the definition of $\operatorname{cont}(T\mid_{\text{cols}\geq j})$, we see that $(\operatorname{cont}(T\mid_{\text{cols}\geq j}))_k = \left|(T\mid_{\text{cols}\geq j})^{-1}(k)\right| > 0$.

But we know that $\operatorname{cont}(T\mid_{\text{cols}\geq j})$ is a partition. Thus,

$$(\operatorname{cont}(T\mid_{\text{cols}\geq j}))_1 \geq (\operatorname{cont}(T\mid_{\text{cols}\geq j}))_2 \geq (\operatorname{cont}(T\mid_{\text{cols}\geq j}))_3 \geq \cdots,$$

so that $(\operatorname{cont}(T\mid_{\text{cols}\geq j}))_1 \geq (\operatorname{cont}(T\mid_{\text{cols}\geq j}))_k > 0$. But the definition of $\operatorname{cont}(T\mid_{\text{cols}\geq j})$ yields

$$(\operatorname{cont}(T\mid_{\text{cols}\geq j}))_1 = \left|(T\mid_{\text{cols}\geq j})^{-1}(1)\right| = (\text{the number of entries of } T\mid_{\text{cols}\geq j} \text{ equal to } 1).$$

Hence, $(\text{the number of entries of } T\mid_{\text{cols}\geq j} \text{ equal to } 1) = (\operatorname{cont}(T\mid_{\text{cols}\geq j}))_1 > 0$. Hence, the skew tableau $T\mid_{\text{cols}\geq j}$ must have at least one entry equal to 1.

But each cell of the skew tableau $T\mid_{\text{cols}\geq j}$ lies in one of the columns $j, j+1, j+2, \ldots$ (by the definition of $T\mid_{\text{cols}\geq j}$) and in one of the rows $1, 2, 3, \ldots$ (obviously). Hence, each cell of the skew tableau $T\mid_{\text{cols}\geq j}$ lies (weakly) southeast of the cell $(1, j) = \mathbf{c}$. As a consequence, each entry of the skew tableau $T\mid_{\text{cols}\geq j}$

(This is because each of the first r columns of T has length k, but the entries in this column must be strictly increasing from top to bottom, and knowing that these entries belong to $\{1, 2, \ldots, k\}$, we see that this is only possible if the column has the form $\begin{array}{c} 1 \\ 2 \\ \vdots \\ k \end{array}$.)

[729] *Proof.* All parts of μ are $\leq m$. Thus, $\mu_1 \leq m$ (since μ_1 is either a part of μ and therefore $\leq m$, or is zero and therefore $\leq m$ as well). Hence, $m \geq \mu_1$. Similarly, $n \geq \nu_1$. Now, $\lambda_1 > \underbrace{m}_{\geq \mu_1} + \underbrace{n}_{\geq \nu_1} \geq \mu_1 + \nu_1$, so that $\lambda_1 - \mu_1 > \nu_1$, qed.

is greater or equal to the entry of $T\mid_{\text{cols}\geq j}$ in the cell $\mathbf{c}$ [730]. Since the entry of $T\mid_{\text{cols}\geq j}$ in the cell $\mathbf{c}$ is $\left(T\mid_{\text{cols}\geq j}\right)(\mathbf{c})=k$, this yields that each entry of the skew tableau $T\mid_{\text{cols}\geq j}$ is greater or equal to k, and thus greater than 1 (since $k>1$). As a consequence, no entry of the skew tableau $T\mid_{\text{cols}\geq j}$ can be equal to 1. This contradicts the fact that the skew tableau $T\mid_{\text{cols}\geq j}$ must have at least one entry equal to 1. This contradiction proves that our assumption was wrong. Hence, (13.92.1) is proven.

Now, Corollary 2.6.12 yields that $c^{\lambda}_{\mu,\nu}$ counts column-strict tableaux T of shape λ/μ with $\operatorname{cont}(T)=\nu$ having the property that each $\operatorname{cont}\left(T\mid_{\text{cols}\geq j}\right)$ is a partition. Since there exists no such T (according to (13.92.1)), this yields that $c^{\lambda}_{\mu,\nu}=0$. This solves Exercise 2.9.16(b).

Remark: An alternative solution of Exercise 2.9.16(b) can be obtained easily from Exercise 2.9.17(c).

(c) Let us forget that λ is fixed.

If $\lambda\in\operatorname{Par}$ is such that $\ell(\lambda)>k$, then

$$s_{\lambda}(x_1,x_2,\ldots,x_k)=0. \tag{13.92.2}$$

[731]

We first notice that

$$s_{\mu}(x_1,x_2,\ldots,x_k)\cdot s_{\nu}(x_1,x_2,\ldots,x_k)=\sum_{\lambda\in\operatorname{Par};\ \ell(\lambda)\leq k}c^{\lambda}_{\mu,\nu}s_{\lambda}(x_1,x_2,\ldots,x_k). \tag{13.92.3}$$

[732] Hence,

$$s_{\mu}(x_1,x_2,\ldots,x_k)\cdot s_{\nu}(x_1,x_2,\ldots,x_k)=\sum_{\substack{\lambda\in\operatorname{Par};\ \ell(\lambda)\leq k;\\ \text{all parts of }\lambda\text{ are }\leq m+n}}c^{\lambda}_{\mu,\nu}s_{\lambda}(x_1,x_2,\ldots,x_k). \tag{13.92.4}$$

[730] Indeed, T is a column-strict tableau, and thus the entries of T increase weakly left-to-right along rows, and increase strictly top-to-bottom in columns. Consequently, the entries of T increase weakly as one moves to the southeast. The same holds for the entries of $T\mid_{\text{cols}\geq j}$ (since $T\mid_{\text{cols}\geq j}$ is a restriction of T), and thus each entry of $T\mid_{\text{cols}\geq j}$ is greater or equal to the entry of $T\mid_{\text{cols}\geq j}$ in the cell $\mathbf{c}$ (because each cell of $T\mid_{\text{cols}\geq j}$ lies (weakly) southeast of the cell $\mathbf{c}$). Qed.

[731] *Proof of (13.92.2):* Let $\lambda\in\operatorname{Par}$ be such that $\ell(\lambda)>k$. Then, the number of parts of λ is $\ell(\lambda)>k$. Hence, Exercise 2.3.8(b) (applied to k instead of n) yields $s_{\lambda}(x_1,x_2,\ldots,x_k)=0$, qed.

[732] *Proof of (13.92.3):* We have $s_\mu s_\nu=\sum_\lambda c^{\lambda}_{\mu,\nu}s_\lambda$, where the sum ranges over all partitions λ (according to the definition of the coefficients $c^{\lambda}_{\mu,\nu}$). In other words, $s_\mu s_\nu=\sum_{\lambda\in\operatorname{Par}}c^{\lambda}_{\mu,\nu}s_\lambda$. Evaluating both sides of this equality at $(x_1,x_2,\ldots,x_k)$, we obtain

$$\begin{aligned}
&s_{\mu}(x_1,x_2,\ldots,x_k)\cdot s_{\nu}(x_1,x_2,\ldots,x_k)\\
&=\sum_{\lambda\in\operatorname{Par}}c^{\lambda}_{\mu,\nu}s_{\lambda}(x_1,x_2,\ldots,x_k)\\
&=\sum_{\lambda\in\operatorname{Par};\ \ell(\lambda)\leq k}c^{\lambda}_{\mu,\nu}s_{\lambda}(x_1,x_2,\ldots,x_k)+\sum_{\lambda\in\operatorname{Par};\ \ell(\lambda)>k}c^{\lambda}_{\mu,\nu}\underbrace{s_{\lambda}(x_1,x_2,\ldots,x_k)}_{\substack{=0\\ \text{(by (13.92.2))}}}\\
&=\sum_{\lambda\in\operatorname{Par};\ \ell(\lambda)\leq k}c^{\lambda}_{\mu,\nu}s_{\lambda}(x_1,x_2,\ldots,x_k)+\underbrace{\sum_{\lambda\in\operatorname{Par};\ \ell(\lambda)>k}c^{\lambda}_{\mu,\nu}0}_{=0}\\
&=\sum_{\lambda\in\operatorname{Par};\ \ell(\lambda)\leq k}c^{\lambda}_{\mu,\nu}s_{\lambda}(x_1,x_2,\ldots,x_k),
\end{aligned}$$

qed.

[733]

Define a set $\mathfrak{A}$ by

$$\mathfrak{A} = \{\alpha \in \operatorname{Par} \ \mid \ \ell(\alpha) \leq k; \text{ all parts of } \alpha \text{ are } \leq m+n\}. \tag{13.92.5}$$

For every partition λ, let $\lambda^{\vee\{m+n\}}$ denote the k-tuple $(m+n-\lambda_k, m+n-\lambda_{k-1}, \ldots, m+n-\lambda_1)$. It is straightforward to see that for every $\lambda \in \mathfrak{A}$, we have

$$\lambda^{\vee\{m+n\}} \in \mathfrak{A}. \tag{13.92.6}$$

Thus, the map

$$\mathfrak{A} \to \mathfrak{A}, \qquad \lambda \mapsto \lambda^{\vee\{m+n\}}$$

is well-defined. It is easy to see that this map is an involution (i.e., every $\lambda \in \mathfrak{A}$ satisfies $\left(\lambda^{\vee\{m+n\}}\right)^{\vee\{m+n\}} = \lambda$), thus a bijection.

Now, the summation sign "$\sum\limits_{\substack{\lambda\in\operatorname{Par};\ \ell(\lambda)\leq k;\\ \text{all parts of } \lambda \text{ are } \leq m+n}}$" in (13.92.4) rewrites as "$\sum_{\lambda\in\mathfrak{A}}$" (because of how we defined $\mathfrak{A}$). Hence, (13.92.4) rewrites as

$$s_\mu(x_1, x_2, \ldots, x_k) \cdot s_\nu(x_1, x_2, \ldots, x_k) = \sum_{\lambda\in\mathfrak{A}} c^\lambda_{\mu,\nu} s_\lambda(x_1, x_2, \ldots, x_k). \tag{13.92.7}$$

Hence, in the Laurent polynomial ring $\mathbf{k}\left[x_1, x_2, \ldots, x_k, x_1^{-1}, x_2^{-1}, \ldots, x_k^{-1}\right]$, we have

$$s_\mu\left(x_1^{-1}, x_2^{-1}, \ldots, x_k^{-1}\right) \cdot s_\nu\left(x_1^{-1}, x_2^{-1}, \ldots, x_k^{-1}\right) = \sum_{\lambda\in\mathfrak{A}} c^\lambda_{\mu,\nu} s_\lambda\left(x_1^{-1}, x_2^{-1}, \ldots, x_k^{-1}\right). \tag{13.92.8}$$

[734]

Now, it is straightforward to see that $\mu^{\vee\{m\}}$ and $\nu^{\vee\{n\}}$ are partitions satisfying $\ell\left(\mu^{\vee\{m\}}\right) \leq k$ and $\ell\left(\nu^{\vee\{n\}}\right) \leq k$; also, all parts of the partition $\mu^{\vee\{m\}}$ are $\leq m$, and all parts of the partition $\nu^{\vee\{n\}}$ are $\leq n$. Moreover,

$$\mu = \left(m - \left(\mu^{\vee\{m\}}\right)_k, m - \left(\mu^{\vee\{m\}}\right)_{k-1}, \ldots, m - \left(\mu^{\vee\{m\}}\right)_1\right) \tag{13.92.9}$$

and

$$\nu = \left(n - \left(\nu^{\vee\{n\}}\right)_k, n - \left(\nu^{\vee\{n\}}\right)_{k-1}, \ldots, n - \left(\nu^{\vee\{n\}}\right)_1\right) \tag{13.92.10}$$

[733] *Proof of (13.92.4):* From (13.92.3), we obtain

$$\begin{aligned}
& s_\mu(x_1, x_2, \ldots, x_k) \cdot s_\nu(x_1, x_2, \ldots, x_k) \\
&= \sum_{\lambda\in\operatorname{Par};\ \ell(\lambda)\leq k} c^\lambda_{\mu,\nu} s_\lambda(x_1, x_2, \ldots, x_k) \\
&= \sum_{\substack{\lambda\in\operatorname{Par};\ \ell(\lambda)\leq k;\\ \text{all parts of } \lambda \text{ are } \leq m+n}} c^\lambda_{\mu,\nu} s_\lambda(x_1, x_2, \ldots, x_k) + \sum_{\substack{\lambda\in\operatorname{Par};\ \ell(\lambda)\leq k;\\ \text{not all parts of } \lambda \text{ are } \leq m+n}} \underbrace{c^\lambda_{\mu,\nu}}_{\substack{=0\\ \text{(by Exercise 2.9.16(b))}}} s_\lambda(x_1, x_2, \ldots, x_k) \\
&= \sum_{\substack{\lambda\in\operatorname{Par};\ \ell(\lambda)\leq k;\\ \text{all parts of } \lambda \text{ are } \leq m+n}} c^\lambda_{\mu,\nu} s_\lambda(x_1, x_2, \ldots, x_k) + \underbrace{\sum_{\substack{\lambda\in\operatorname{Par};\ \ell(\lambda)\leq k;\\ \text{not all parts of } \lambda \text{ are } \leq m+n}} 0 s_\lambda(x_1, x_2, \ldots, x_k)}_{=0} \\
&= \sum_{\substack{\lambda\in\operatorname{Par};\ \ell(\lambda)\leq k;\\ \text{all parts of } \lambda \text{ are } \leq m+n}} c^\lambda_{\mu,\nu} s_\lambda(x_1, x_2, \ldots, x_k).
\end{aligned}$$

This proves (13.92.4).

[734] This follows by substituting the variables $x_1^{-1}, x_2^{-1}, \ldots, x_k^{-1}$ for $x_1, x_2, \ldots, x_k$ in the equality (13.92.7).

(this is still easy to verify). These observations lead us to the conclusion that we can apply (13.92.8) to $\mu^{\vee\{m\}}$, μ, $\nu^{\vee\{n\}}$ and ν instead of μ, $\mu^{\vee\{m\}}$, ν and $\nu^{\vee\{n\}}$. As a result, we obtain

$$\begin{aligned} s_{\mu^{\vee\{m\}}}\left(x_1^{-1}, x_2^{-1}, \ldots, x_k^{-1}\right) \cdot s_{\nu^{\vee\{n\}}}\left(x_1^{-1}, x_2^{-1}, \ldots, x_k^{-1}\right) &= \sum_{\lambda \in \mathfrak{A}} c^{\lambda}_{\mu^{\vee\{m\}}, \nu^{\vee\{n\}}} s_\lambda\left(x_1^{-1}, x_2^{-1}, \ldots, x_k^{-1}\right) \\ &= \sum_{\lambda \in \mathfrak{A}} c^{\lambda^{\vee\{m+n\}}}_{\mu^{\vee\{m\}}, \nu^{\vee\{n\}}} s_{\lambda^{\vee\{m+n\}}}\left(x_1^{-1}, x_2^{-1}, \ldots, x_k^{-1}\right) \qquad (13.92.11) \end{aligned}$$

(here, we have substituted $\lambda^{\vee\{m+n\}}$ for the summation index λ (since the map $\mathfrak{A} \to \mathfrak{A}$, $\lambda \mapsto \lambda^{\vee\{m+n\}}$ is a bijection)).

Now, we can apply Exercise 2.9.15(d) to $\mu^{\vee\{m\}}$, $\mu^{\vee\{m\}}$, μ and μ instead of λ, μ, $\lambda^\vee$ and $\mu^\vee$ (because of (13.92.9)). As a result, we obtain

$$s_\mu\left(x_1, x_2, \ldots, x_k\right) = \left(x_1 x_2 \cdots x_k\right)^m \cdot s_{\mu^{\vee\{m\}}}\left(x_1^{-1}, x_2^{-1}, \ldots, x_k^{-1}\right). \qquad (13.92.12)$$

Also, we can apply Exercise 2.9.15(d) to n, $\nu^{\vee\{n\}}$, $\nu^{\vee\{n\}}$, ν and ν instead of m, λ, μ, $\lambda^\vee$ and $\mu^\vee$ (because of (13.92.10)). As a result, we obtain

$$s_\nu\left(x_1, x_2, \ldots, x_k\right) = \left(x_1 x_2 \cdots x_k\right)^n \cdot s_{\nu^{\vee\{n\}}}\left(x_1^{-1}, x_2^{-1}, \ldots, x_k^{-1}\right). \qquad (13.92.13)$$

Furthermore, every $\lambda \in \mathfrak{A}$ satisfies

$$s_\lambda\left(x_1, x_2, \ldots, x_k\right) = \left(x_1 x_2 \cdots x_k\right)^{m+n} \cdot s_{\lambda^{\vee\{m+n\}}}\left(x_1^{-1}, x_2^{-1}, \ldots, x_k^{-1}\right). \qquad (13.92.14)$$

[735]

Now, (13.92.7) yields

$$\begin{aligned} &\sum_{\lambda \in \mathfrak{A}} c^{\lambda}_{\mu,\nu} s_\lambda\left(x_1, x_2, \ldots, x_k\right) \\ &= \underbrace{s_\mu\left(x_1, x_2, \ldots, x_k\right)}_{\substack{=\left(x_1 x_2 \cdots x_k\right)^m \cdot s_{\mu^{\vee\{m\}}}\left(x_1^{-1}, x_2^{-1}, \ldots, x_k^{-1}\right) \\ \text{(by (13.92.12))}}} \cdot \underbrace{s_\nu\left(x_1, x_2, \ldots, x_k\right)}_{\substack{=\left(x_1 x_2 \cdots x_k\right)^n \cdot s_{\nu^{\vee\{n\}}}\left(x_1^{-1}, x_2^{-1}, \ldots, x_k^{-1}\right) \\ \text{(by (13.92.13))}}} \\ &= \left(x_1 x_2 \cdots x_k\right)^m \cdot s_{\mu^{\vee\{m\}}}\left(x_1^{-1}, x_2^{-1}, \ldots, x_k^{-1}\right) \cdot \left(x_1 x_2 \cdots x_k\right)^n \cdot s_{\nu^{\vee\{n\}}}\left(x_1^{-1}, x_2^{-1}, \ldots, x_k^{-1}\right) \\ &= \underbrace{\left(x_1 x_2 \cdots x_k\right)^m \left(x_1 x_2 \cdots x_k\right)^n}_{=\left(x_1 x_2 \cdots x_k\right)^{m+n}} \cdot \underbrace{s_{\mu^{\vee\{m\}}}\left(x_1^{-1}, x_2^{-1}, \ldots, x_k^{-1}\right) \cdot s_{\nu^{\vee\{n\}}}\left(x_1^{-1}, x_2^{-1}, \ldots, x_k^{-1}\right)}_{\substack{=\sum_{\lambda \in \mathfrak{A}} c^{\lambda^{\vee\{m+n\}}}_{\mu^{\vee\{m\}}, \nu^{\vee\{n\}}} s_{\lambda^{\vee\{m+n\}}}\left(x_1^{-1}, x_2^{-1}, \ldots, x_k^{-1}\right) \\ \text{(by (13.92.11))}}} \\ &= \left(x_1 x_2 \cdots x_k\right)^{m+n} \cdot \sum_{\lambda \in \mathfrak{A}} c^{\lambda^{\vee\{m+n\}}}_{\mu^{\vee\{m\}}, \nu^{\vee\{n\}}} s_{\lambda^{\vee\{m+n\}}}\left(x_1^{-1}, x_2^{-1}, \ldots, x_k^{-1}\right) \\ &= \sum_{\lambda \in \mathfrak{A}} c^{\lambda^{\vee\{m+n\}}}_{\mu^{\vee\{m\}}, \nu^{\vee\{n\}}} \underbrace{\left(x_1 x_2 \cdots x_k\right)^{m+n} \cdot s_{\lambda^{\vee\{m+n\}}}\left(x_1^{-1}, x_2^{-1}, \ldots, x_k^{-1}\right)}_{\substack{=s_\lambda\left(x_1, x_2, \ldots, x_k\right) \\ \text{(by (13.92.14))}}} \\ &= \sum_{\lambda \in \mathfrak{A}} c^{\lambda^{\vee\{m+n\}}}_{\mu^{\vee\{m\}}, \nu^{\vee\{n\}}} s_\lambda\left(x_1, x_2, \ldots, x_k\right). \qquad (13.92.15) \end{aligned}$$

[735] *Proof of (13.92.14):* Let $\lambda \in \mathfrak{A}$. Then, $\lambda \in \mathfrak{A} = \{\alpha \in \operatorname{Par} \mid \ell(\alpha) \leq k;\text{ all parts of } \alpha \text{ are } \leq m+n\}$. In other words, λ is an element of Par such that $\ell(\lambda) \leq k$ and such that all parts of λ are $\leq m+n$. Now, (13.92.6) yields that $\lambda^{\vee\{m+n\}} \in \mathfrak{A} = \{\alpha \in \operatorname{Par} \mid \ell(\alpha) \leq k;\text{ all parts of } \alpha \text{ are } \leq m+n\}$. In other words, $\lambda^{\vee\{m+n\}}$ is an element of Par such that $\ell\left(\lambda^{\vee\{m+n\}}\right) \leq k$ and such that all parts of $\lambda^{\vee\{m+n\}}$ are $\leq m+n$.

Recall that $\left(\lambda^{\vee\{m+n\}}\right)^{\vee\{m+n\}} = \lambda$ (as we said, this follows easily from the definitions). Thus,

$$\lambda = \left(\lambda^{\vee\{m+n\}}\right)^{\vee\{m+n\}} = \left(m+n-\left(\lambda^{\vee\{m+n\}}\right)_k, m+n-\left(\lambda^{\vee\{m+n\}}\right)_{k-1}, \ldots, m+n-\left(\lambda^{\vee\{m+n\}}\right)_1\right)$$

(by the definition of $\left(\lambda^{\vee\{m+n\}}\right)^{\vee\{m+n\}}$). Hence, we can apply Exercise 2.9.15(d) to $m+n$, $\lambda^{\vee\{m+n\}}$, $\lambda^{\vee\{m+n\}}$, λ and λ instead of m, λ, μ, $\lambda^\vee$ and $\mu^\vee$. As a result, we obtain

$$s_\lambda\left(x_1, x_2, \ldots, x_k\right) = \left(x_1 x_2 \cdots x_k\right)^{m+n} \cdot s_{\lambda^{\vee\{m+n\}}}\left(x_1^{-1}, x_2^{-1}, \ldots, x_k^{-1}\right).$$

This proves (13.92.14).

But Remark 2.3.9(d) (applied to $N = k$) yields that the set $\{s_\lambda(x_1, x_2, \ldots, x_k)\}$, as λ runs through all partitions having length $\leq k$, is a basis of the $\mathbf{k}$-module $\Lambda(x_1, x_2, \ldots, x_k)$. In other words, the family $(s_\lambda(x_1, x_2, \ldots, x_k))_{\lambda \in \operatorname{Par};\ \ell(\lambda) \leq k}$ is a basis of the $\mathbf{k}$-module $\Lambda(x_1, x_2, \ldots, x_k)$. In particular, the family $(s_\lambda(x_1, x_2, \ldots, x_k))_{\lambda \in \operatorname{Par};\ \ell(\lambda) \leq k}$ is $\mathbf{k}$-linearly independent.

But the set $\mathfrak{A}$ is a subset of $\{\alpha \in \operatorname{Par} \mid \ell(\alpha) \leq k\}$. Hence, the family $(s_\lambda(x_1, x_2, \ldots, x_k))_{\lambda \in \mathfrak{A}}$ is a subfamily of the family $(s_\lambda(x_1, x_2, \ldots, x_k))_{\lambda \in \{\alpha \in \operatorname{Par} \mid \ell(\alpha) \leq k\}} = (s_\lambda(x_1, x_2, \ldots, x_k))_{\lambda \in \operatorname{Par};\ \ell(\lambda) \leq k}$. Since the family $(s_\lambda(x_1, x_2, \ldots, x_k))_{\lambda \in \operatorname{Par};\ \ell(\lambda) \leq k}$ is $\mathbf{k}$-linearly independent, we thus conclude that its subfamily $(s_\lambda(x_1, x_2, \ldots, x_k))_{\lambda \in \mathfrak{A}}$ is also $\mathbf{k}$-linearly independent. As a consequence, if two $\mathbf{k}$-linear combinations of the family $(s_\lambda(x_1, x_2, \ldots, x_k))_{\lambda \in \mathfrak{A}}$ are equal, then their respective coefficients must be equal. Thus, from (13.92.16), we conclude that

$$c^{\lambda}_{\mu,\nu} = c^{\lambda^{\vee\{m+n\}}}_{\mu^{\vee\{m\}},\nu^{\vee\{n\}}} \qquad \text{for every } \lambda \in \mathfrak{A}. \tag{13.92.16}$$

Now, let λ be a partition such that $\ell(\lambda) \leq k$. Assume that all parts of λ are $\leq m+n$. Then, λ is an element of Par such that $\ell(\lambda) \leq k$ and such that all parts of λ are $\leq m+n$. In other words, $\lambda \in \mathfrak{A}$ (by the definition of $\mathfrak{A}$). Hence, (13.92.16) yields $c^{\lambda}_{\mu,\nu} = c^{\lambda^{\vee\{m+n\}}}_{\mu^{\vee\{m\}},\nu^{\vee\{n\}}}$. This solves Exercise 2.9.16(c).

13.93. **Solution to Exercise 2.9.17.** *Solution to Exercise 2.9.17.* (a) We notice that every partition λ and every positive integer i satisfy

$$\begin{aligned}\left(\lambda^t\right)_i &= |\{j \in \{1,2,3,\ldots\} \mid \lambda_j \geq i\}| \qquad (\text{by } (2.2.7)) \\ &= |\{j \in \{1,2,\ldots,\ell(\lambda)\} \mid \lambda_j \geq i\}| \end{aligned} \tag{13.93.1}$$

(because $\lambda_j \geq i$ can happen only when $j \in \{1,2,\ldots,\ell(\lambda)\}$ (since i is positive)).

Let μ and ν be two partitions. For every positive integer i, we have

$$\begin{aligned}\left((\mu \sqcup \nu)^t\right)_i &= \left|\left\{j \in \{1,2,\ldots,\ell(\mu \sqcup \nu)\} \mid (\mu \sqcup \nu)_j \geq i\right\}\right| \qquad (\text{by } (13.93.1), \text{ applied to } \lambda = \mu \sqcup \nu) \\ &= (\text{the number of entries of } \mu \sqcup \nu \text{ which are } \geq i) \\ &= \left(\text{the number of entries of the list } \left(\mu_1, \mu_2, \ldots, \mu_{\ell(\mu)}, \nu_1, \nu_2, \ldots, \nu_{\ell(\nu)}\right) \text{ which are } \geq i\right) \\ &\qquad \left(\begin{array}{c}\text{since } \mu \sqcup \nu \text{ is the result of sorting the list } \left(\mu_1, \mu_2, \ldots, \mu_{\ell(\mu)}, \nu_1, \nu_2, \ldots, \nu_{\ell(\nu)}\right) \\ \text{in decreasing order, and clearly the procedure of sorting does not change} \\ \text{the number of entries of the list which are } \geq i\end{array}\right) \\ &= \underbrace{\left(\text{the number of entries of the list } \left(\mu_1, \mu_2, \ldots, \mu_{\ell(\mu)}\right) \text{ which are } \geq i\right)}_{\substack{=|\{j \in \{1,2,\ldots,\ell(\mu)\} \mid \mu_j \geq i\}| = \left(\mu^t\right)_i \\ (\text{by } (13.93.1), \text{ applied to } \lambda = \mu)}} \\ &\qquad + \underbrace{\left(\text{the number of entries of the list } \left(\nu_1, \nu_2, \ldots, \nu_{\ell(\nu)}\right) \text{ which are } \geq i\right)}_{\substack{=|\{j \in \{1,2,\ldots,\ell(\nu)\} \mid \nu_j \geq i\}| = \left(\nu^t\right)_i \\ (\text{by } (13.93.1), \text{ applied to } \lambda = \nu)}} \\ &= \left(\mu^t\right)_i + \left(\nu^t\right)_i = \left(\mu^t + \nu^t\right)_i\end{aligned}$$

(since the definition of $\mu^t + \nu^t$ yields $\left(\mu^t + \nu^t\right)_i = \left(\mu^t\right)_i + \left(\nu^t\right)_i$). In other words,

$$(\mu \sqcup \nu)^t = \mu^t + \nu^t. \tag{13.93.2}$$

Applying this to μ^t and ν^t instead of μ and ν, we obtain $\left(\mu^t \sqcup \nu^t\right)^t = \underbrace{\left(\mu^t\right)^t}_{=\mu} + \underbrace{\left(\nu^t\right)^t}_{=\nu} = \mu + \nu$, so that $\mu + \nu = \left(\mu^t \sqcup \nu^t\right)^t$ and thus

$$\left(\underbrace{\mu + \nu}_{=\left(\mu^t \sqcup \nu^t\right)^t}\right)^t = \left(\left(\mu^t \sqcup \nu^t\right)^t\right)^t = \mu^t \sqcup \nu^t.$$

This solves Exercise 2.9.17(a).

(b) Let μ and ν be two partitions. We are going to prove that $c^{\mu+\nu}_{\mu,\nu} = 1$.

A $(\mu+\nu,\mu,\nu)$-*LR-tableau* will mean a column-strict tableau T of shape $(\mu+\nu)/\mu$ with $\operatorname{cont}(T)=\nu$ having the property that each $\operatorname{cont}(T|_{\operatorname{cols}\geq j})$ is a partition (where we are using the notations of Corollary 2.6.12). Corollary 2.6.12 shows that $c^{\mu+\nu}_{\mu,\nu}$ is the number of $(\mu+\nu,\mu,\nu)$-LR-tableaux. Hence, in order to prove that $c^{\mu+\nu}_{\mu,\nu}=1$, it will be enough to show that there exists one and only one $(\mu+\nu,\mu,\nu)$-LR-tableau.

First of all, let T_0 be the filling of the skew shape $(\mu+\nu)/\mu$ which assigns to every cell in row i the number i, for all $i\in\{1,2,3,...\}$. This T_0 is clearly a column-strict tableau, and satisfies $\operatorname{cont}(T_0)=\nu$ (because for every positive integer i, the i-th row of the skew shape $(\mu+\nu)/\mu$ has $\underbrace{(\mu+\nu)_i}_{=\mu_i+\nu_i}-\mu_i=(\mu_i+\nu_i)-\mu_i=\nu_i$ cells). Moreover, for every $j\in\{1,2,3,...\}$, the weak composition $\operatorname{cont}(T_0|_{\operatorname{cols}\geq j})$ is a partition[736]. Therefore, T_0 is a $(\mu+\nu,\mu,\nu)$-LR-tableau. It remains to prove that it is the only $(\mu+\nu,\mu,\nu)$-LR-tableau.

So fix any $(\mu+\nu,\mu,\nu)$-LR-tableau T. Thus, T is a column-strict tableau of shape $(\mu+\nu)/\mu$ with $\operatorname{cont}(T)=\nu$ having the property that each $\operatorname{cont}(T|_{\operatorname{cols}\geq j})$ is a partition.

We shall show that

$$\text{for every } i\in\{1,2,3,...\}\text{, all entries in the } i\text{-th row of } T \text{ are } \leq i. \tag{13.93.3}$$

Proof of (13.93.3): Assume the contrary. Then, there exists an $i\in\{1,2,3,...\}$ such that not all entries in the i-th row of T are $\leq i$. Let p be the **smallest** such i (this is clearly well-defined). Hence, not all entries in the p-th row of T are $\leq p$. But since p is minimal, we know that (13.93.3) holds for every $i<p$.

Not all entries in the p-th row of T are $\leq p$. In other words, at least one entry of the p-th row of T must be $>p$. Since the entries of T weakly increase along rows, this yields that the rightmost entry of the p-th row of T is $>p$. Let q be this entry, and let j be the column in which the rightmost cell of the p-th row of T lies. Thus, q is the entry in cell (p,j) of T. Therefore, the entry q appears in the tableau $T|_{\operatorname{cols}\geq j}$. Hence, $\operatorname{cont}(T|_{\operatorname{cols}\geq j})_q\geq 1$. Note that $q>p$ (by the definition of q).

We know that $\operatorname{cont}(T|_{\operatorname{cols}\geq j})$ is a partition, so that

$$\begin{aligned}(\operatorname{cont}(T|_{\operatorname{cols}\geq j}))_p &\geq \operatorname{cont}(T|_{\operatorname{cols}\geq j})_q \qquad (\text{since } p<q)\\ &\geq 1.\end{aligned}$$

In other words, the entry p appears somewhere in the tableau $T|_{\operatorname{cols}\geq j}$. Where can it appear? It cannot appear in any of the first $p-1$ rows, because all entries in these rows are $<p$ (since (13.93.3) holds for every $i<p$). Hence, it must appear in the p-th row or further down. In other words, this entry p appears in a cell (u,v) of T with $u\geq p$ and $v\geq j$. As a consequence, this entry p is $\geq$ to the entry in cell (p,j) of T (since T is column-strict). Since the entry in cell (p,j) of T is q, this yields $p\geq q$, which contradicts $q>p$. This contradiction shows that our assumption was wrong, and (13.93.3) is proven.

We furthermore claim that

$$\text{for every } i\in\{1,2,3,...\}\text{, all entries in the } i\text{-th row of } T \text{ are } i. \tag{13.93.4}$$

Proof of (13.93.4): Assume the contrary. Then, there exists an $i\in\{1,2,3,...\}$ such that not all entries in the i-th row of T are i. Let p be the **smallest** such i (this is clearly well-defined). Hence, not all entries in the p-th row of T are p. But since p is minimal, we know that (13.93.4) holds for every $i<p$.

All entries in the p-th row of T are $\leq p$ (by (13.93.3)), but not all of them are p. Hence, the p-th row of T has an entry $<p$. Let k be this entry. Thus, $k<p$. We can apply (13.93.4) to $i=k$ (since (13.93.4) holds for every $i<p$), and conclude that all entries in the k-th row of T are k. Thus, in the tableau T, the entry k appears ν_k times in row k (since the length of the k-th row of T is $\underbrace{(\mu+\nu)_k}_{=\mu_k+\nu_k}-\mu_k=(\mu_k+\nu_k)-\mu_k=\nu_k$) and at least 1 time in row p (by the definition of k). In total, k must thus appear at least ν_k+1 times in T, which contradicts the fact that $\operatorname{cont}(T)=\nu$. This contradiction disproves our assumption, and thus (13.93.4) is proven.

[736] *Proof.* In fact, fix $j\in\{1,2,3,...\}$. Then, both ν_i and μ_i decrease with $i\in\{1,2,3,...\}$ (since ν and μ are partitions), and thus $\min\{\mu_i,j\}$ also decreases with $i\in\{1,2,3,...\}$.

Now, for every $i\in\{1,2,3,...\}$, the i-th entry of $\operatorname{cont}(T_0|_{\operatorname{cols}\geq j})$ is the number of boxes in columns $j,j+1,j+2,...$ of the i-th row of the skew shape $(\mu+\nu)/\mu$. This number is easily seen to be $\underbrace{(\mu+\nu)_i}_{=\mu_i+\nu_i}-\underbrace{\max\{\mu_i,j\}}_{=\mu_i+j-\min\{\mu_i,j\}}=\mu_i+\nu_i-(\mu_i+j-\min\{\mu_i,j\})=\nu_i-j+\min\{\mu_i,j\}$, and therefore decreases with i (because both ν_i and $\min\{\mu_i,j\}$ decrease with i). Hence, $\operatorname{cont}(T_0|_{\operatorname{cols}\geq j})$ is a partition, qed.

Now that (13.93.4) is proven, we immediately conclude that $T = T_0$. Now, forget that we fixed T. We thus have shown that every $(\mu+\nu,\mu,\nu)$-LR-tableau T equals T_0. Hence, there exists one and only one $(\mu+\nu,\mu,\nu)$-LR-tableau, namely $T = T_0$ (because we have already seen that T_0 is a $(\mu+\nu,\mu,\nu)$-LR-tableau). As we said above, this proves that

$$c_{\mu,\nu}^{\mu+\nu} = 1. \tag{13.93.5}$$

It remains to show that $c_{\mu,\nu}^{\mu\sqcup\nu} = 1$. Exercise 2.7.11(c) (applied to $\lambda = \mu\sqcup\nu$) shows that

$$\begin{aligned} c_{\mu,\nu}^{\mu\sqcup\nu} &= c_{\mu^t,\nu^t}^{(\mu\sqcup\nu)^t} = c_{\mu^t,\nu^t}^{\mu^t+\nu^t} \qquad \left(\text{since } (\mu\sqcup\nu)^t = \mu^t+\nu^t \text{ (by Exercise 2.9.17(a))}\right) \\ &= 1 \qquad \left(\text{by (13.93.5), applied to } \mu^t \text{ and } \nu^t \text{ instead of } \mu \text{ and } \nu\right). \end{aligned}$$

This solves Exercise 2.9.17(b).

(c) Let $k \in \mathbb{N}$ and $n \in \mathbb{N}$ satisfy $k \le n$, and let $\mu \in \operatorname{Par}_k$, $\nu \in \operatorname{Par}_{n-k}$ and $\lambda \in \operatorname{Par}_n$ be such that $c_{\mu,\nu}^{\lambda} \neq 0$. We need to prove that $\mu+\nu \trianglerighteq \lambda \trianglerighteq \mu\sqcup\nu$.

We have $|\lambda| = n$, $|\mu| = k$ and $|\nu| = n-k$, so that $|\lambda| = n = \underbrace{k}_{=|\mu|} + \underbrace{(n-k)}_{=|\nu|} = |\mu|+|\nu|$. We will first show that

$$\mu+\nu \trianglerighteq \lambda. \tag{13.93.6}$$

A (λ,μ,ν)-*LR-tableau* will mean a column-strict tableau T of shape λ/μ with $\operatorname{cont}(T) = \nu$ having the property that each $\operatorname{cont}(T|_{\text{cols}\ge j})$ is a partition (where we are using the notations of Corollary 2.6.12). Corollary 2.6.12 shows that $c_{\mu,\nu}^{\lambda}$ is the number of (λ,μ,ν)-LR-tableaux. Since $c_{\mu,\nu}^{\lambda} \neq 0$, we thus see that there exists at least one (λ,μ,ν)-LR-tableau. Let this (λ,μ,ν)-LR-tableau be T.

Just as in the solution of Exercise 2.9.17(b), we can prove that (13.93.3) holds. Let k be a positive integer. Applying (13.93.3) to all $i \in \{1,2,...,k\}$, we see that all entries in the first k rows of T (meaning the 1-st row, the 2-nd row, etc., the k-th row) are $\le k$. Hence,

$$\begin{aligned} &(\text{the number of all entries in the first } k \text{ rows of } T) \\ &\le (\text{the number of all entries } \le k \text{ in } T) = \sum_{i=1}^{k} \underbrace{(\text{the number of all entries } i \text{ in } T)}_{\substack{=(\operatorname{cont} T)_i = \nu_i \\ (\text{since } \operatorname{cont} T = \nu)}} = \sum_{i=1}^{k} \nu_i. \end{aligned}$$

Since

$$\begin{aligned} &(\text{the number of all entries in the first } k \text{ rows of } T) \\ &= \sum_{i=1}^{k} \underbrace{(\text{the number of all entries in the } i\text{-th row of } T)}_{=\lambda_i-\mu_i} = \sum_{i=1}^{k} (\lambda_i-\mu_i) = \sum_{i=1}^{k} \lambda_i - \sum_{i=1}^{k} \mu_i, \end{aligned}$$

this rewrites as $\sum_{i=1}^{k} \lambda_i - \sum_{i=1}^{k} \mu_i \le \sum_{i=1}^{k} \nu_i$. Hence,

$$\sum_{i=1}^{k} \lambda_i \le \sum_{i=1}^{k} \mu_i + \sum_{i=1}^{k} \nu_i = \sum_{i=1}^{k} \underbrace{(\mu_i+\nu_i)}_{=(\mu+\nu)_i} = \sum_{i=1}^{k} (\mu+\nu)_i\,,$$

so that $\sum_{i=1}^{k} (\mu+\nu)_i \ge \sum_{i=1}^{k} \lambda_i$. In other words, $(\mu+\nu)_1 + (\mu+\nu)_2 + \cdots + (\mu+\nu)_k \ge \lambda_1+\lambda_2+\cdots+\lambda_k$.

Now, let us forget that we fixed k. We have shown that $(\mu+\nu)_1 + (\mu+\nu)_2 + \cdots + (\mu+\nu)_k \ge \lambda_1+\lambda_2+\cdots+\lambda_k$ for every positive integer k. Combined with $|\lambda| = |\mu|+|\nu| = |\mu+\nu|$, this yields that $\mu+\nu \trianglerighteq \lambda$. This proves (13.93.6).

It now remains to prove that $\lambda \trianglerighteq \mu\sqcup\nu$. To do so, we notice that Exercise 2.7.11(c) yields $c_{\mu,\nu}^{\lambda} = c_{\mu^t,\nu^t}^{\lambda^t}$, so that $c_{\mu^t,\nu^t}^{\lambda^t} = c_{\mu,\nu}^{\lambda} \neq 0$. Hence, we can apply (13.93.6) to λ^t, μ^t and ν^t instead of λ, μ and ν. As a result, we obtain $\mu^t+\nu^t \trianglerighteq \lambda^t$. Since $(\mu\sqcup\nu)^t = \mu^t+\nu^t$ (by Exercise 2.9.17(a)), this rewrites as $(\mu\sqcup\nu)^t \trianglerighteq \lambda^t$. But Exercise 2.2.9 (applied to $\mu\sqcup\nu$ instead of μ) yields that $\lambda \trianglerighteq \mu\sqcup\nu$ if and only if $(\mu\sqcup\nu)^t \trianglerighteq \lambda^t$. Since we already know that $(\mu\sqcup\nu)^t \trianglerighteq \lambda^t$, we can thus conclude that $\lambda \trianglerighteq \mu\sqcup\nu$. Combined with (13.93.6), this yields $\mu+\nu \trianglerighteq \lambda \trianglerighteq \mu\sqcup\nu$. This solves Exercise 2.9.17(c).

(d) Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$ and $\alpha, \beta \in \operatorname{Par}_n$ and $\gamma, \delta \in \operatorname{Par}_m$ be such that $\alpha \rhd \beta$ and $\gamma \rhd \delta$. It is completely straightforward to check that

$$\alpha + \gamma \rhd \beta + \delta \tag{13.93.7}$$

[737]. It remains to prove that $\alpha \sqcup \gamma \rhd \beta \sqcup \delta$.

Exercise 2.2.9 (applied to $\lambda = \alpha$ and $\mu = \beta$) yields that $\alpha \rhd \beta$ if and only if $\beta^t \rhd \alpha^t$. Since $\alpha \rhd \beta$, we thus have $\beta^t \rhd \alpha^t$. Similarly, $\delta^t \rhd \gamma^t$. Hence, we can apply (13.93.7) to β^t, α^t, δ^t and γ^t instead of α, β, γ and δ. As a result, we obtain $\beta^t + \delta^t \rhd \alpha^t + \gamma^t$. Now, (13.93.2) (applied to $\mu = \beta$ and $\nu = \delta$) yields $(\beta \sqcup \delta)^t = \beta^t + \delta^t$. Similarly, $(\alpha \sqcup \gamma)^t = \alpha^t + \gamma^t$. Thus, $(\beta \sqcup \delta)^t = \beta^t + \delta^t \rhd \alpha^t + \gamma^t = (\alpha \sqcup \gamma)^t$.

Finally, Exercise 2.2.9 (applied to $n+m$, $\alpha \sqcup \gamma$ and $\beta \sqcup \delta$ instead of n, λ and μ) yields that $\alpha \sqcup \gamma \rhd \beta \sqcup \delta$ if and only if $(\beta \sqcup \delta)^t \rhd (\alpha \sqcup \gamma)^t$. Since we have $(\beta \sqcup \delta)^t \rhd (\alpha \sqcup \gamma)^t$, we thus obtain $\alpha \sqcup \gamma \rhd \beta \sqcup \delta$. This completes the solution of Exercise 2.9.17(d).

(e) The Ferrers diagram of the partition $\lambda = (m^k)$ is a rectangle. Let C denote the center of this rectangle. For every partition μ satisfying $\mu \subseteq \lambda$, let us define a partition μ^c by $\mu^c = (m - \mu_k, m - \mu_{k-1}, ..., m - \mu_1)$ [738]. The Ferrers diagram of this partition μ^c is obtained from the skew Ferrers diagram of the skew partition [739] λ/μ by the 180° rotation around C. In other words, the skew Ferrers diagram of $\mu^c/\varnothing$ is obtained from the skew Ferrers diagram of the skew partition λ/μ by the 180° rotation around C. Hence, Exercise 2.3.4(b) (applied to $\lambda' = \mu^c$ and $\mu' = \varnothing$) yields that

$$s_{\lambda/\mu} = s_{\mu^c/\varnothing} = s_{\mu^c}. \tag{13.93.8}$$

Now, Proposition 2.3.6(iv) yields

$$\Delta s_\lambda = \sum_{\mu \subseteq \lambda} s_\mu \otimes \underbrace{s_{\lambda/\mu}}_{\substack{=s_{\mu^c} \\ \text{(by (13.93.8))}}} = \sum_{\mu \subseteq \lambda} s_\mu \otimes s_{\mu^c}.$$

Compared with

$$\Delta s_\lambda = \sum_{\mu,\nu} \underbrace{\hat{c}^\lambda_{\mu,\nu}}_{=c^\lambda_{\mu,\nu}} s_\mu \otimes s_\nu \qquad \text{(by (2.5.7))}$$
$$= \sum_{\mu,\nu} c^\lambda_{\mu,\nu} s_\mu \otimes s_\nu,$$

this yields $\sum_{\mu,\nu} c^\lambda_{\mu,\nu} s_\mu \otimes s_\nu = \sum_{\mu \subseteq \lambda} s_\mu \otimes s_{\mu^c}$. Comparing the coefficients in front of $s_\mu \otimes s_\nu$ on both sides of this equality, we obtain: Any two partitions μ and ν satisfy

$$c^\lambda_{\mu,\nu} = \begin{cases} 1, & \text{if } \mu \subseteq \lambda \text{ and } \nu = \mu^c; \\ 0, & \text{otherwise} \end{cases} \in \{0, 1\}.$$

This solves Exercise 2.9.17(e).

(f) We know that $(s_\mu)_{\mu \in \operatorname{Par}}$ is a basis of the $\mathbf{k}$-module Λ. Hence, $(s_\mu \otimes s_\nu)_{\mu,\nu \in \operatorname{Par}}$ is a basis of the $\mathbf{k}$-module $\Lambda \otimes \Lambda$. We will refer to this basis as the *Schur basis* of $\Lambda \otimes \Lambda$.

The equality (2.5.7) yields

$$\Delta s_\lambda = \sum_{\mu,\nu} \underbrace{\hat{c}^\lambda_{\mu,\nu}}_{=c^\lambda_{\mu,\nu}} s_\mu \otimes s_\nu = \sum_{\mu,\nu} c^\lambda_{\mu,\nu} s_\mu \otimes s_\nu.$$

Thus, for every $\mu, \nu \in \operatorname{Par}$, the $s_\mu \otimes s_\nu$-coefficient of Δs_λ with respect to the Schur basis of $\Lambda \otimes \Lambda$ is $c^\lambda_{\mu,\nu}$.

But $\lambda = (a+1, 1^b)$. Hence, Exercise 2.9.14(d) gives a formula for $\Delta s_\lambda = \Delta s_{(a+1,1^b)}$ as a sum of pure tensors of the form $s_\mu \otimes s_\nu$ with $\mu, \nu \in \operatorname{Par}$. Every such pure tensor occurs **at most once** in this formula (as can be easily verified). In other words, for every $\mu, \nu \in \operatorname{Par}$, the $s_\mu \otimes s_\nu$-coefficient of Δs_λ with respect

[737] To prove this, just recall how $\alpha + \gamma$ and $\beta + \delta$ are defined, and recall that two partitions $\lambda, \mu \in \operatorname{Par}_n$ satisfy $\lambda \rhd \mu$ if and only if we have $(\lambda_1 + \lambda_2 + \cdots + \lambda_k \geq \mu_1 + \mu_2 + \cdots + \mu_k$ for every positive integer $k)$.

[738] This is well-defined, since every $i \in \{1, 2, ..., k\}$ satisfies $\mu_i \leq m$ (because $\mu \subseteq \lambda = (m^k)$).

[739] A *skew partition* shall mean a pair (α, β) of partitions satisfying $\beta \subseteq \alpha$. We write such a skew partition (α, β) as α/β. For every skew partition α/β, we define the skew Ferrers diagram $Y(\alpha/\beta)$ of α/β by $Y(\alpha/\beta) = Y(\alpha) \setminus Y(\beta)$, where $Y(\kappa)$ means the Ferrers diagram of a partition κ.

to the Schur basis of $\Lambda \otimes \Lambda$ is either 0 or 1. Since the $s_\mu \otimes s_\nu$-coefficient of Δs_λ with respect to the Schur basis of $\Lambda \otimes \Lambda$ is $c^\lambda_{\mu,\nu}$, this rewrites as follows: For every $\mu, \nu \in \operatorname{Par}$, the scalar $c^\lambda_{\mu,\nu}$ is either 0 or 1. This solves Exercise 2.9.17(f).

(g) The following solution is inspired by [211, proof of Thm. 2.1(iv)].

Before we start with the solution, we recall two formulas. Firstly, any two partitions μ and ν satisfy

$$\begin{aligned} s_\mu s_\nu &= \sum_{\lambda \in \operatorname{Par}} c^\lambda_{\mu,\nu} s_\lambda && \text{(this is just a restatement of (2.5.6))} \\ &= \sum_{\tau \in \operatorname{Par}} c^\tau_{\mu,\nu} s_\tau && \text{(here, we renamed the summation index } \lambda \text{ as } \tau\text{)} . \end{aligned} \tag{13.93.9}$$

Secondly, any two partitions λ and μ satisfy

$$\begin{aligned} s_{\lambda/\mu} &= \sum_{\nu \in \operatorname{Par}} c^\lambda_{\mu,\nu} s_\nu && \text{(this is one of the identities in Remark 2.5.9)} \\ &= \sum_{\tau \in \operatorname{Par}} c^\lambda_{\mu,\tau} s_\tau && \text{(here, we renamed the summation index } \nu \text{ as } \tau\text{)} . \end{aligned} \tag{13.93.10}$$

Now, let λ be any partition, and let μ and ν be two rectangular partitions. We need to show that $c^\lambda_{\mu,\nu} \in \{0,1\}$.

Assume the contrary. Thus, $c^\lambda_{\mu,\nu} \notin \{0,1\}$, so that $c^\lambda_{\mu,\nu} \neq 0$.

We have $c^\lambda_{\mu,\nu} = c^\lambda_{\nu,\mu}$. Hence, we can WLOG assume that $\ell(\mu) \geq \ell(\nu)$ (since otherwise, we can just switch μ and ν). Assume this. Assume WLOG that $\mu \neq \varnothing$ (since otherwise, $\mu = \varnothing$ and thus $\nu = \varnothing$ (because $\ell(\mu) \geq \ell(\nu)$), which makes the claim $c^\lambda_{\mu,\nu} \in \{0,1\}$ a rather obvious fact).

The partition μ is rectangular, i.e., has the form $\left(m^k\right) = \left(\underbrace{m, m, \ldots, m}_{k \text{ times}}\right)$ for some $m \in \mathbb{N}$ and $k \in \mathbb{N}$. Consider these m and k. Both m and k are positive (since $\left(m^k\right) = \mu \neq \varnothing$). Thus, $k = \ell(\mu)$ (since $\mu = \left(m^k\right)$), so that $k = \ell(\mu) \geq \ell(\nu)$, thus $\ell(\nu) \leq k$.

Corollary 2.6.12 shows that $c^\lambda_{\mu,\nu}$ counts column-strict tableaux T of shape λ/μ with $\operatorname{cont}(T) = \nu$ having the property that each $\operatorname{cont}(T|_{\operatorname{cols} \geq j})$ is a partition (where we are using the notations of Corollary 2.6.12). Since $c^\lambda_{\mu,\nu} \neq 0$, we see that there exists at least one such tableau T. Consider this T. All entries of the tableau T are $\leq \ell(\nu)$ (because $\operatorname{cont}(T) = \nu$). This quickly yields that $\mu_k \geq \lambda_{k+1}$ [740]. Of course, we also have $\mu \subseteq \lambda$ (since T is a tableau of shape λ/μ), and thus $\lambda_i \geq \mu_i$ for every $i \in \{1,2,3,\ldots\}$. In particular, every $i \in \{1,2,\ldots,k\}$ satisfies $\lambda_i \geq \mu_i = m$ (since $\mu = \left(m^k\right)$).

Now, define F, $F_{\text{rows} \leq k}$ and $F_{\text{rows} > k}$ as in Exercise 2.3.5. Define four partitions α, β, γ and δ by

$$\alpha = (\lambda_1 - m, \lambda_2 - m, \ldots, \lambda_k - m), \qquad \beta = \varnothing,$$
$$\gamma = (\lambda_{k+1}, \lambda_{k+2}, \lambda_{k+3}, \ldots), \qquad \delta = \varnothing$$

(notice that α is well-defined because every $i \in \{1,2,\ldots,k\}$ satisfies $\lambda_i \geq m$). It is now easy to see that the skew Ferrers diagram α/β can be obtained from $F_{\text{rows} \leq k}$ by parallel translation (namely, the translation by m steps to the west), and that the skew Ferrers diagram γ/δ can be obtained from $F_{\text{rows} > k}$ by parallel

[740] *Proof.* Assume the contrary. Thus, $\mu_k < \lambda_{k+1}$. Since $\mu = \left(m^k\right)$, we have $\mu_k = m$, so that $m = \mu_k < \lambda_{k+1}$. Thus, $m \leq \lambda_{k+1} - 1$ (since m and λ_{k+1} are integers), so that $m + 1 \leq \lambda_{k+1}$. Now, for every $i \in \{1,2,\ldots,k+1\}$, we have

$$\begin{aligned} \mu_i &\leq m && \left(\text{since } \mu = \left(m^k\right)\right) \\ &< m + 1 \leq \lambda_{k+1} \leq \lambda_i && (\text{since } k+1 \geq i) . \end{aligned}$$

Hence, $(i, m+1)$ is a cell of the skew Ferrers diagram λ/μ for every $i \in \{1,2,\ldots,k+1\}$. Altogether, the $(m+1)$-th column of the skew Ferrers diagram λ/μ contains at least $k+1$ different cells (namely, the cells $(i, m+1)$ for all $i \in \{1,2,\ldots,k+1\}$). In the tableau T, these $k+1$ different cells must be filled with $k+1$ distinct values (because the entries of T are strictly decreasing top-to-bottom in columns). As a consequence, there must be at least $k+1$ distinct values among the entries of T; but this is impossible, because all entries of T are $\leq \ell(\nu) \leq k$. This contradiction proves that our assumption was wrong, qed.

translation (namely, the translation by k steps to the north). Hence, Exercise 2.3.5 yields

$$\begin{aligned} s_{\lambda/\mu} &= s_{\alpha/\beta} s_{\gamma/\delta} = \underbrace{s_{\alpha/\varnothing}}_{=s_\alpha} \underbrace{s_{\gamma/\varnothing}}_{=s_\gamma} \qquad (\text{since } \beta = \varnothing \text{ and } \delta = \varnothing) \\ &= s_\alpha s_\gamma = \sum_{\tau \in \operatorname{Par}} c^{\tau}_{\alpha,\gamma} s_\tau \qquad (\text{by (13.93.9)}) . \end{aligned}$$

Compared with (13.93.10), this yields $\sum_{\tau \in \operatorname{Par}} c^{\lambda}_{\mu,\tau} s_\tau = \sum_{\tau \in \operatorname{Par}} c^{\tau}_{\alpha,\gamma} s_\tau$. Since $(s_\tau)_{\tau \in \operatorname{Par}}$ is a $\mathbf{k}$-basis of Λ, we can compare coefficients before s_ν in this equality, and thus obtain $c^{\lambda}_{\mu,\nu} = c^{\nu}_{\alpha,\gamma}$. But ν is a rectangular partition, and thus has the form $\left(\widetilde{m}^{\widetilde{k}}\right)$ for some $\widetilde{m} \in \mathbb{N}$ and $\widetilde{k} \in \mathbb{N}$. Hence, $c^{\nu}_{\alpha,\gamma} \in \{0, 1\}$ (by Exercise 2.9.17(e), applied to $\widetilde{m}$, $\widetilde{k}$, ν, α and γ instead of m, k, λ, μ and ν). Thus, $c^{\lambda}_{\mu,\nu} = c^{\nu}_{\alpha,\gamma} \in \{0, 1\}$, which contradicts $c^{\lambda}_{\mu,\nu} \notin \{0, 1\}$. This contradiction proves that our assumption was wrong, and Exercise 2.9.17(g) is solved.

13.94. **Solution to Exercise 2.9.18.** *Solution to Exercise 2.9.18.* (a) We shall prove the implications $\mathcal{A} \Longrightarrow \mathcal{B}$ and $\mathcal{B} \Longrightarrow \mathcal{A}$.

Proof of the implication $\mathcal{A} \Longrightarrow \mathcal{B}$: Assume that Assertion $\mathcal{A}$ holds. That is, there exist a partition λ and a column-strict tableau T of shape λ/μ such that all $(i, j) \in \{1, 2, 3, \ldots\}^2$ satisfy (2.9.13). Consider this λ and this T. Since T is column-strict, the entries of T increase weakly left-to-right along rows, and increase strictly top-to-bottom along columns.

For every $u \in \mathbb{N}$ and $j \in \{1, 2, 3, ...\}$, we have

$$\begin{aligned} &(\text{the number of all entries } \leq u \text{ in the } j\text{-th row of } T) \\ &= \sum_{i=1}^{u} \underbrace{(\text{the number of all entries } i \text{ in the } j\text{-th row of } T)}_{\substack{=b_{i,j} \\ (\text{by (2.9.13)})}} = \sum_{i=1}^{u} b_{i,j} \\ &= b_{1,j} + b_{2,j} + \cdots + b_{u,j} . \end{aligned} \tag{13.94.1}$$

Let $(i, j) \in \mathbb{N} \times \{1, 2, 3, \ldots\}$. We are going to prove that

$$\mu_{j+1} + (b_{1,j+1} + b_{2,j+1} + \cdots + b_{i+1,j+1}) \leq \mu_j + (b_{1,j} + b_{2,j} + \cdots + b_{i,j}) . \tag{13.94.2}$$

Indeed, assume the contrary. Then,

$$\mu_{j+1} + (b_{1,j+1} + b_{2,j+1} + \cdots + b_{i+1,j+1}) > \mu_j + (b_{1,j} + b_{2,j} + \cdots + b_{i,j}) . \tag{13.94.3}$$

Hence,

$$\begin{aligned} &\mu_{j+1} + \underbrace{(\text{the number of all entries } \leq i+1 \text{ in the } (j+1)\text{-th row of } T)}_{\substack{=b_{1,j+1}+b_{2,j+1}+\cdots+b_{i+1,j+1} \\ (\text{by (13.94.1), applied to } i+1 \text{ and } j+1 \text{ instead of } u \text{ and } j)}} \\ &= \mu_{j+1} + (b_{1,j+1} + b_{2,j+1} + \cdots + b_{i+1,j+1}) \\ &> \mu_j + \underbrace{(b_{1,j} + b_{2,j} + \cdots + b_{i,j})}_{\geq 0} > \mu_j \geq \mu_{j+1} \qquad (\text{since } \mu \text{ is a partition}) . \end{aligned}$$

Subtracting μ_{j+1} from both sides of this inequality, we obtain

$$(\text{the number of all entries } \leq i+1 \text{ in the } (j+1)\text{-th row of } T) > 0.$$

In other words, there exists at least one entry $\leq i + 1$ in the $(j + 1)$-th row of T. Let $\mathbf{c}$ be the rightmost cell of the $(j + 1)$-th row of T which contains such an entry. That is, $\mathbf{c}$ is the rightmost cell of the $(j + 1)$-th row of T which contains an entry $\leq i + 1$.

The cell $\mathbf{c}$ lies in the $(j + 1)$-th row of T. Hence, we can write the cell $\mathbf{c}$ in the form $\mathbf{c} = (j + 1, y)$ for some positive integer y. Consider this y.

The entries of T increase weakly left-to-right along rows. Thus, the cells of the $(j + 1)$-th row of T which contain entries $\leq i + 1$ form a contiguous segment of the $(j + 1)$-th row of T. This segment begins in cell

$(j+1, \mu_{j+1}+1)$ (since the entries of the $(j+1)$-th row of T begin in cell $(j+1, \mu_{j+1}+1)$ (because T has shape λ/μ)), and ends in cell $(j+1, y)$ (since $(j+1, y) = \mathbf{c}$ is the rightmost cell of the $(j+1)$-th row of T which contains an entry $\leq i+1$). Hence, this segment contains precisely $y - \mu_{j+1}$ cells. In other words, there exist exactly $y - \mu_{j+1}$ cells of the $(j+1)$-th row of T which contain entries $\leq i+1$. In other words, the number of all cells of the $(j+1)$-th row of T which contain entries $\leq i+1$ is $y - \mu_{j+1}$. In other words,

$$(\text{the number of all entries } \leq i+1 \text{ in the } (j+1)\text{-th row of } T) = y - \mu_{j+1}.$$

Solving this for y, we obtain

$$y = \mu_{j+1} + (\text{the number of all entries } \leq i+1 \text{ in the } (j+1)\text{-th row of } T).$$

Since $\mathbf{c}$ is a cell of the tableau T, it is clear that the cell $\mathbf{c}$ lies inside the Ferrers diagram of λ, and therefore (due to $j \in \{1, 2, 3, ...\}$) the cell (j, y) must also lie inside the Ferrers diagram of λ (because the cell (j, y) is the northern neighbor of the cell $(j+1, y) = \mathbf{c}$).

We have

$$\begin{aligned} y &= \mu_{j+1} + (\text{the number of all entries } \leq i+1 \text{ in the } (j+1)\text{-th row of } T) \\ &= \mu_{j+1} + (b_{1,j+1} + b_{2,j+1} + \cdots + b_{i+1,j+1}) > \mu_j. \end{aligned}$$

Therefore, the cell (j, y) lies outside the Ferrers diagram of μ. Since (j, y) lies inside the Ferrers diagram of λ but outside the Ferrers diagram of μ, we see that (j, y) is a cell of T.

But recall that the entries of T increase strictly top-to-bottom along columns. Hence, the entry of T in the cell (j, y) must be strictly smaller than the entry of T in the cell $\mathbf{c}$ (since the cell (j, y) is the northern neighbor of the cell $(j+1, y) = \mathbf{c}$). Since the latter entry is $\leq i+1$ (by the definition of $\mathbf{c}$), this shows that the entry of T in the cell (j, y) must be strictly smaller than $i+1$. Hence, this entry must be $\leq i$. As a consequence, all cells in the j-th row of T which lie weakly to the left of the cell (j, y) must also have entries $\leq i$ (because the entries of T increase weakly left-to-right along rows). The number of such cells is $y - \mu_j$ (because the entries in the j-th row of T begin in cell $(j, \mu_j + 1)$ (since T has shape λ/μ)). Thus, there are at least $y - \mu_j$ cells in the j-th row of T which have entries $\leq i$; in other words, the number of all entries $\leq i$ in the j-th row of T is at least $y - \mu_j$. In other words,

$$(\text{the number of all entries } \leq i \text{ in the } j\text{-th row of } T) \geq y - \mu_j.$$

Since

$$(\text{the number of all entries } \leq i \text{ in the } j\text{-th row of } T) = b_{1,j} + b_{2,j} + \cdots + b_{i,j}$$

(by (13.94.1), applied to $u = i$), this rewrites as follows:

$$b_{1,j} + b_{2,j} + \cdots + b_{i,j} \geq y - \mu_j.$$

Now, (13.94.3) becomes

$$\begin{aligned} \mu_{j+1} + (b_{1,j+1} + b_{2,j+1} + \cdots + b_{i+1,j+1}) > \mu_j + \underbrace{(b_{1,j} + b_{2,j} + \cdots + b_{i,j})}_{\geq y - \mu_j} \geq \mu_j + (y - \mu_j) \\ = y = \mu_{j+1} + (b_{1,j+1} + b_{2,j+1} + \cdots + b_{i+1,j+1}). \end{aligned}$$

This is absurd. This contradiction proves that our assumption was wrong. Hence, (13.94.2) holds.

Now, forget that we have fixed (i, j). We thus have proven that the inequality (13.94.2) holds for all $(i, j) \in \mathbb{N} \times \{1, 2, 3, \ldots\}$. In other words, Assertion $\mathcal{B}$ holds. We thus have proven the implication $\mathcal{A} \Longrightarrow \mathcal{B}$.

Proof of the implication $\mathcal{B} \Longrightarrow \mathcal{A}$: Assume that Assertion $\mathcal{B}$ holds. That is, the inequality (2.9.14) holds for all $(i, j) \in \mathbb{N} \times \{1, 2, 3, \ldots\}$.

For every $j \in \{1, 2, 3, ...\}$, the sum $b_{1,j} + b_{2,j} + b_{3,j} + \cdots$ has only finitely many nonzero addends (since $b_{i,j} = 0$ for all but finitely many pairs (i, j)), and can be computed as the following limit with respect to the discrete topology:

$$b_{1,j} + b_{2,j} + b_{3,j} + \cdots = \lim_{i \to \infty} (b_{1,j} + b_{2,j} + \cdots + b_{i,j}). \tag{13.94.4}$$

For every $j \in \{1, 2, 3, ...\}$, we have

$$\mu_j + (b_{1,j} + b_{2,j} + b_{3,j} + \cdots) \geq \mu_{j+1} + (b_{1,j+1} + b_{2,j+1} + b_{3,j+1} + \cdots) \tag{13.94.5}$$

[741]. Hence, we can define a partition λ by

$$(\lambda_j = \mu_j + (b_{1,j} + b_{2,j} + b_{3,j} + \cdots) \qquad \text{for every } j \in \{1,2,3,...\}) \tag{13.94.6}$$

[742]. Consider this partition λ.

We have $\mu \subseteq \lambda$ (since every $j \in \{1,2,3,...\}$ satisfies $\lambda_j = \mu_j + \underbrace{(b_{1,j} + b_{2,j} + b_{3,j} + \cdots)}_{\geq 0} \geq \mu_j$). We also have

$$\lambda_j - \mu_j = b_{1,j} + b_{2,j} + b_{3,j} + \cdots \qquad \text{for every } j \in \{1,2,3,...\}$$

(because of (13.94.6)).

Now, we construct a filling T of the Ferrers diagram of λ/μ with positive integers as follows: For every $j \in \{1,2,3,...\}$, the j-th row of this Ferrers diagram of λ/μ has $\lambda_j - \mu_j = b_{1,j} + b_{2,j} + b_{3,j} + \cdots$ cells. We fill in the leftmost $b_{1,j}$ of these cells with 1's, the leftmost $b_{2,j}$ of the remaining cells with 2's, the leftmost $b_{3,j}$ of the still remaining cells with 3's, and so on. Once this has been done for all positive integers j (of course, for all sufficiently high j, the j-th row of the Ferrers diagram of λ/μ has no cells, and therefore nothing has to be filled), we are left with a filling of the Ferrers diagram of λ/μ with positive integers. Denote this filling by T. It is clear that the entries of T increase weakly left-to-right in rows (by the construction of T). We shall soon show that the entries of T increase strictly top-to-bottom in columns.

First, however, let us observe that (2.9.13) holds for all $(i,j) \in \{1,2,3,\ldots\}^2$ (by the construction of T). Hence, every $u \in \mathbb{N}$ and $j \in \{1,2,3,...\}$ satisfy (13.94.1) (this is proven just as in our proof of the implication $\mathcal{A} \Longrightarrow \mathcal{B}$).

Now, we are going to prove that the entries of T increase strictly top-to-bottom in columns.

Indeed, assume the contrary. Then, there exists at least one column of T in which the entries don't increase strictly top-to-bottom. Let this be the k-th column. So the entries in the k-th column of T don't increase strictly top-to-bottom. As a consequence, there exist two cells $\mathbf{c}$ and $\mathbf{d}$ in the k-th column of T such that $\mathbf{d}$ is the northern neighbor of $\mathbf{c}$, but the entry of T in cell $\mathbf{d}$ is not smaller than the entry of T in cell $\mathbf{c}$. Consider these cells $\mathbf{c}$ and $\mathbf{d}$.

Write the cell $\mathbf{c}$ as $\mathbf{c} = (x,y)$. Then, $\mathbf{d} = (x-1,y)$ (since $\mathbf{d}$ is the northern neighbor of $\mathbf{c}$), so that $x - 1 \geq 1$.

Let p be the entry of T in cell $\mathbf{c}$. Then, the entry of T in cell $\mathbf{d}$ is not smaller than p (since the entry of T in cell $\mathbf{d}$ is not smaller than the entry of T in cell $\mathbf{c}$). In other words, the entry of T in cell $(x-1,y)$ is not

[741] *Proof.* Let $j \in \{1,2,3,...\}$. Then,

$$\begin{aligned}
&\mu_{j+1} + \underbrace{(b_{1,j+1} + b_{2,j+1} + b_{3,j+1} + \cdots)}_{\substack{=\lim_{i\to\infty}(b_{1,j+1}+b_{2,j+1}+\cdots+b_{i,j+1}) \\ \text{(by (13.94.4), applied to } j+1 \text{ instead of } j)}} \\
&= \mu_{j+1} + \underbrace{\lim_{i\to\infty}(b_{1,j+1} + b_{2,j+1} + \cdots + b_{i,j+1})}_{\substack{=\lim_{i\to\infty}(b_{1,j+1}+b_{2,j+1}+\cdots+b_{i+1,j+1}) \\ \text{(here, we substituted } i+1 \text{ for } i \text{ in the limit)}}} \\
&= \mu_{j+1} + \lim_{i\to\infty}(b_{1,j+1} + b_{2,j+1} + \cdots + b_{i+1,j+1}) = \lim_{i\to\infty}\left(\underbrace{\mu_{j+1} + (b_{1,j+1} + b_{2,j+1} + \cdots + b_{i+1,j+1})}_{\substack{\leq \mu_j + (b_{1,j}+b_{2,j}+\cdots+b_{i,j}) \\ \text{(by (2.9.14))}}}\right) \\
&\leq \lim_{i\to\infty}(\mu_j + (b_{1,j} + b_{2,j} + \cdots + b_{i,j})) = \mu_j + \underbrace{\lim_{i\to\infty}(b_{1,j} + b_{2,j} + \cdots + b_{i,j})}_{\substack{=b_{1,j}+b_{2,j}+b_{3,j}+\cdots \\ \text{(by (13.94.4))}}} = \mu_j + (b_{1,j} + b_{2,j} + b_{3,j} + \cdots),
\end{aligned}$$

and this proves (13.94.5).

[742] Here, we are using the fact that $\mu_j + (b_{1,j} + b_{2,j} + b_{3,j} + \cdots) = 0$ for all sufficiently high positive integers j. The proof of this is easy: We have $\mu_j = 0$ for all sufficiently high positive integers j (since μ is a partition), and we also have $b_{1,j} + b_{2,j} + b_{3,j} + \cdots = 0$ for all sufficiently high positive integers j (since $b_{i,j} = 0$ for all but finitely many pairs (i,j)). Thus, if a positive integer j is sufficiently high, we have both $\mu_j = 0$ and $b_{1,j} + b_{2,j} + b_{3,j} + \cdots = 0$, and therefore $\mu_j + (b_{1,j} + b_{2,j} + b_{3,j} + \cdots) = 0 + 0 = 0$, qed.

smaller than p (since $\mathbf{d} = (x-1, y)$). In other words,

$$\text{(the entry of } T \text{ in cell } (x-1,y)) \geq p. \tag{13.94.7}$$

But let $i = p-1$. Then, $i+1 = p$ is the entry of T in cell $\mathbf{c} = (x,y)$. Thus, the cell (x,y) of T has the entry $i+1$. Hence, all cells in the x-th row of T which lie weakly to the left of the cell (x,y) must have entries $\leq i+1$ (since the entries of T increase weakly left-to-right in rows). The number of such cells is $y - \mu_x$ (since the entries in the x-th row of T begin in cell $(x, \mu_x + 1)$ (since T has shape λ/μ)). Thus, there are at least $y - \mu_x$ cells in the x-th row of T which have entries $\leq i+1$; in other words, the number of all entries $\leq i+1$ in the x-th row of T is at least $y - \mu_x$. In other words,

$$\text{(the number of all entries } \leq i+1 \text{ in the } x\text{-th row of } T) \geq y - \mu_x .$$

Since (the number of all entries $\leq i+1$ in the x-th row of T) $= b_{1,x} + b_{2,x} + \cdots + b_{i+1,x}$ (by (13.94.1), applied to $u = i+1$ and $j = x$), this rewrites as follows:

$$b_{1,x} + b_{2,x} + \cdots + b_{i+1,x} \geq y - \mu_x .$$

Now, recall that $x - 1 \geq 1$, so that $x - 1 \in \{1, 2, 3, ...\}$. Hence, (2.9.14) (applied to $j = x-1$) yields

$$\mu_{(x-1)+1} + \left(b_{1,(x-1)+1} + b_{2,(x-1)+1} + \cdots + b_{i+1,(x-1)+1}\right) \leq \mu_{x-1} + \left(b_{1,x-1} + b_{2,x-1} + \cdots + b_{i,x-1}\right),$$

so that

$$\begin{aligned} \mu_{x-1} + \left(b_{1,x-1} + b_{2,x-1} + \cdots + b_{i,x-1}\right) &\geq \underbrace{\mu_{(x-1)+1}}_{=\mu_x} + \underbrace{\left(b_{1,(x-1)+1} + b_{2,(x-1)+1} + \cdots + b_{i+1,(x-1)+1}\right)}_{=b_{1,x}+b_{2,x}+\cdots+b_{i+1,x} \geq y-\mu_x} \\ &\geq \mu_x + y - \mu_x = y. \end{aligned} \tag{13.94.8}$$

But every entry $\leq i$ in the $(x-1)$-st row of T must lie in a cell strictly left of the cell $(x-1, y)$ [743]. Since the number of such cells[744] is $y - 1 - \mu_{x-1}$ (because the entries in the $(x-1)$-st row of T begin in cell $(x-1, \mu_{x-1}+1)$ (since T has shape λ/μ)), this yields that there are at most $y - 1 - \mu_{x-1}$ entries $\leq i$ in the $(x-1)$-st row of T. In other words, the number of all entries $\leq i$ in the $(x-1)$-st row of T is at most $y - 1 - \mu_{x-1}$. In other words,

$$\text{(the number of all entries } \leq i \text{ in the } (x-1)\text{-st row of } T) \leq y - 1 - \mu_{x-1}.$$

Since (the number of all entries $\leq i$ in the $(x-1)$-st row of T) $= b_{1,x-1} + b_{2,x-1} + \cdots + b_{i,x-1}$ (by (13.94.1), applied to $u = i$ and $j = x-1$), this rewrites as follows:

$$b_{1,x-1} + b_{2,x-1} + \cdots + b_{i,x-1} \leq y - \underbrace{1}_{>0} - \mu_{x-1} < y - \mu_{x-1}.$$

Hence, $\mu_{x-1} + \left(b_{1,x-1} + b_{2,x-1} + \cdots + b_{i,x-1}\right) < y$. This contradicts (13.94.8). This contradiction shows that our assumption was wrong. Hence, the entries of T increase strictly top-to-bottom in columns. Since we also know that the entries of T increase weakly left-to-right in rows, and that T is a filling of the Ferrers diagram of λ/μ, this yields that T is a column-strict tableau of shape λ/μ. Besides, we already know that all $(i,j) \in \{1, 2, 3, \ldots\}^2$ satisfy (2.9.13). Hence, Assertion $\mathcal{A}$ holds (with the λ and T that we have constructed above). We thus have proven the implication $\mathcal{B} \Longrightarrow \mathcal{A}$.

Now that both implications $\mathcal{A} \Longrightarrow \mathcal{B}$ and $\mathcal{B} \Longrightarrow \mathcal{A}$ are proven, we conclude that Assertions $\mathcal{A}$ and $\mathcal{B}$ are equivalent. Exercise 2.9.18(a) is solved.

[743] *Proof.* Assume the contrary. Then, there exists an entry $\leq i$ in the $(x-1)$-st row of T which lies in a cell not strictly left of the cell $(x-1, y)$. Let c be this entry. Since c lies in a cell not strictly left of the cell $(x-1,y)$, the entry c must lie in a cell weakly to the right of the cell $(x-1, y)$, and therefore this entry c must be $\geq$ to the entry of T in cell $(x-1, y)$ (because the entries of T increase weakly left-to-right in rows). Hence,

$$\begin{aligned} c &\geq \text{(the entry of } T \text{ in cell } (x-1,y)) \geq p \qquad \text{(by (13.94.7))} \\ &> i \qquad \text{(since } i = p-1 < p\text{)}, \end{aligned}$$

which contradicts the fact that $c \leq i$ (by definition of c). This contradiction shows that our assumption was wrong, qed.

[744] Here, "such cells" means cells of the $(x-1)$-st row of T which lie strictly left of the cell $(x-1, y)$.

(b) Since T is a column-strict tableau, we know that the entries of T increase weakly left-to-right in rows and increase strictly top-to-bottom along columns. Hence, if $\mathbf{c}$ and $\mathbf{d}$ are two cells of T such that the cell $\mathbf{c}$ lies southeast[745] of the cell $\mathbf{d}$, then

$$(\text{the entry of } T \text{ in cell } \mathbf{c}) \geq (\text{the entry of } T \text{ in cell } \mathbf{d}) . \tag{13.94.9}$$

For the same reason, if $\mathbf{c}$ and $\mathbf{d}$ are two cells of T such that the cell $\mathbf{c}$ lies southeast of the cell $\mathbf{d}$ but not on the same row as $\mathbf{d}$, then

$$(\text{the entry of } T \text{ in cell } \mathbf{c}) > (\text{the entry of } T \text{ in cell } \mathbf{d}) . \tag{13.94.10}$$

We shall first prove the equivalence of Assertions $\mathcal{D}$ and $\mathcal{G}$:

Proof of the equivalence $\mathcal{D} \Longleftrightarrow \mathcal{G}$: For every $(i,j) \in \{1,2,3,...\}^2$, let $b_{i,j}$ be the number of all entries j in the i-th row of T. (This is not a typo; we don't want the number of all entries i in the j-th row of T.)

It is clear that $b_{i,j} = 0$ for all but finitely many pairs (i,j). Hence, we can apply Exercise 2.9.18(a) to $\varnothing$ instead of μ. We conclude that the following two assertions $\mathcal{A}'$ and $\mathcal{B}'$ are equivalent[746]:

- *Assertion $\mathcal{A}'$:* There exist a partition ν and a column-strict tableau S of shape $\nu/\varnothing$ such that all $(i,j) \in \{1,2,3,\ldots\}^2$ satisfy
$$b_{i,j} = (\text{the number of all entries } i \text{ in the } j\text{-th row of } S) .$$
- *Assertion $\mathcal{B}'$:* The inequality
$$\varnothing_{j+1} + (b_{1,j+1} + b_{2,j+1} + \cdots + b_{i+1,j+1}) \leq \varnothing_j + (b_{1,j} + b_{2,j} + \cdots + b_{i,j})$$
holds for all $(i,j) \in \mathbb{N} \times \{1,2,3,\ldots\}$.

Now, it is easy to see that Assertion $\mathcal{A}'$ is equivalent to Assertion $\mathcal{G}$ [747]. Also, Assertion $\mathcal{B}'$ is equivalent to Assertion $\mathcal{D}$ [748]. Altogether, we have obtained the chain of equivalences $\mathcal{D} \Longleftrightarrow \mathcal{B}' \Longleftrightarrow \mathcal{A}' \Longleftrightarrow \mathcal{G}$.

[745] A cell (r,c) is said to lie *southeast* of a cell (r',c') if and only if we have $r \geq r'$ and $c \geq c'$.

[746] Note that we denote by ν and S the variables that have been called λ and T in Exercise 2.9.18(a), since in our current situation the letters λ and T already have different meanings.

[747] *Proof.* Assertion $\mathcal{A}'$ is equivalent to the following Assertion $\mathcal{A}''$:

- *Assertion $\mathcal{A}''$:* There exists a column-strict tableau S whose shape is a partition such that all $(i,j) \in \{1,2,3,\ldots\}^2$ satisfy
$$b_{i,j} = (\text{the number of all entries } i \text{ in the } j\text{-th row of } S) .$$

(Indeed, Assertion $\mathcal{A}'$ is equivalent to Assertion $\mathcal{A}''$ because a column-strict tableau whose shape is a partition is the same thing as a column-strict tableau of shape $\nu/\varnothing$ with ν being a partition.)

Recall that $b_{i,j} = (\text{the number of all entries } j \text{ in the } i\text{-th row of } T)$ for all $(i,j) \in \{1,2,3,...\}^2$ (by the definition of $b_{i,j}$). Hence, Assertion $\mathcal{A}''$ is equivalent to the following Assertion $\mathcal{A}'''$:

- *Assertion $\mathcal{A}'''$:* There exists a column-strict tableau S whose shape is a partition such that all $(i,j) \in \{1,2,3,\ldots\}^2$ satisfy
$$(\text{the number of all entries } j \text{ in the } i\text{-th row of } T) = (\text{the number of all entries } i \text{ in the } j\text{-th row of } S) .$$

Assertion $\mathcal{A}'''$ is equivalent to the following Assertion $\mathcal{A}''''$:

- *Assertion $\mathcal{A}''''$:* There exists a column-strict tableau S whose shape is a partition such that all $(i,j) \in \{1,2,3,\ldots\}^2$ satisfy
$$(\text{the number of all entries } i \text{ in the } j\text{-th row of } T) = (\text{the number of all entries } j \text{ in the } i\text{-th row of } S) .$$

(Indeed, Assertion $\mathcal{A}''''$ is obtained from Assertion $\mathcal{A}'''$ by substituting (j,i) for the index (i,j).)

Assertion $\mathcal{A}''''$ is obviously equivalent to Assertion $\mathcal{G}$.

Altogether, we have obtained the chain of equivalences $\mathcal{A}' \Longleftrightarrow \mathcal{A}'' \Longleftrightarrow \mathcal{A}''' \Longleftrightarrow \mathcal{A}'''' \Longleftrightarrow \mathcal{G}$. Thus, we know that Assertion $\mathcal{A}'$ is equivalent to Assertion $\mathcal{G}$, qed.

[748] *Proof.* Every $(i,j) \in \mathbb{N} \times \{1,2,3,\ldots\}$ satisfies

$$\begin{aligned}
\underbrace{\varnothing_j}_{=0} + (b_{1,j} + b_{2,j} + \cdots + b_{i,j}) &= b_{1,j} + b_{2,j} + \cdots + b_{i,j} = \sum_{u=1}^{i} \underbrace{b_{u,j}}_{\substack{=(\text{the number of all entries } j \text{ in the } u\text{-th row of } T) \\ (\text{by the definition of } b_{u,j})}} \\
&= \sum_{u=1}^{i} (\text{the number of all entries } j \text{ in the } u\text{-th row of } T) \\
&= (\text{the number of all entries } j \text{ in the first } i \text{ rows of } T) \qquad (13.94.11)
\end{aligned}$$

and

$$\varnothing_{j+1} + (b_{1,j+1} + b_{2,j+1} + \cdots + b_{i+1,j+1}) = (\text{the number of all entries } j+1 \text{ in the first } i+1 \text{ rows of } T) \tag{13.94.12}$$

Hence, we have proven the equivalence $\mathcal{D} \Longleftrightarrow \mathcal{G}$. In order to complete the solution of Exercise 2.9.18(b), it thus remains to prove the equivalence $\mathcal{C} \Longleftrightarrow \mathcal{D} \Longleftrightarrow \mathcal{E} \Longleftrightarrow \mathcal{F}$.

We will achieve this by splitting each of the assertions $\mathcal{C}$, $\mathcal{D}$, $\mathcal{E}$ and $\mathcal{F}$ into many sub-assertions. Namely, for every $i \in \{1,2,3,...\}$, let us define four assertions $\mathcal{C}_i$, $\mathcal{D}_i$, $\mathcal{E}_i$ and $\mathcal{F}_i$ as follows:

- *Assertion $\mathcal{C}_i$:* For every positive integer j, we have $\left(\operatorname{cont}\left(T|_{\text{cols}\geq j}\right)\right)_i \geq \left(\operatorname{cont}\left(T|_{\text{cols}\geq j}\right)\right)_{i+1}$.
- *Assertion $\mathcal{D}_i$:* For every positive integer j, the number of entries $i+1$ in the first j rows of T is $\leq$ to the number of entries i in the first $j-1$ rows of T.
- *Assertion $\mathcal{E}_i$:* For every NE-set S of T, we have $\left(\operatorname{cont}\left(T|_S\right)\right)_i \geq \left(\operatorname{cont}\left(T|_S\right)\right)_{i+1}$.
- *Assertion $\mathcal{F}_i$:* For every prefix v of the Semitic reading word of T, there are at least as many i's among the letters of v as there are $(i+1)$'s among them.

We have the following equivalences:

$$\mathcal{C} \Longleftrightarrow \left(\mathcal{C}_i \text{ holds for every } i \in \{1,2,3,...\}\right)$$

(since $\operatorname{cont}\left(T|_{\text{cols}\geq j}\right)$ is a partition if and only if every $i \in \{1,2,3,...\}$ satisfies $\left(\operatorname{cont}\left(T|_{\text{cols}\geq j}\right)\right)_i \geq \left(\operatorname{cont}\left(T|_{\text{cols}\geq j}\right)\right)_{i+1}$), and

$$\mathcal{D} \Longleftrightarrow \left(\mathcal{D}_i \text{ holds for every } i \in \{1,2,3,...\}\right)$$

(obviously), and

$$\mathcal{E} \Longleftrightarrow \left(\mathcal{E}_i \text{ holds for every } i \in \{1,2,3,...\}\right)$$

(since $\operatorname{cont}\left(T|_S\right)$ is a partition if and only if every $i \in \{1,2,3,...\}$ satisfies $\left(\operatorname{cont}\left(T|_S\right)\right)_i \geq \left(\operatorname{cont}\left(T|_S\right)\right)_{i+1}$), and

$$\mathcal{F} \Longleftrightarrow \left(\mathcal{F}_i \text{ holds for every } i \in \{1,2,3,...\}\right)$$

(by the definition of a Yamanouchi word). Hence, in order to prove the equivalence $\mathcal{C} \Longleftrightarrow \mathcal{D} \Longleftrightarrow \mathcal{E} \Longleftrightarrow \mathcal{F}$, it is enough to show that for every $i \in \{1,2,3,...\}$, we have an equivalence $\mathcal{C}_i \Longleftrightarrow \mathcal{D}_i \Longleftrightarrow \mathcal{E}_i \Longleftrightarrow \mathcal{F}_i$. So let us do this now.

Let $i \in \{1,2,3,...\}$. We need to prove the equivalence $\mathcal{C}_i \Longleftrightarrow \mathcal{D}_i \Longleftrightarrow \mathcal{E}_i \Longleftrightarrow \mathcal{F}_i$. We shall achieve this by proving the implications $\mathcal{C}_i \Longrightarrow \mathcal{E}_i$, $\mathcal{E}_i \Longrightarrow \mathcal{F}_i$, $\mathcal{F}_i \Longrightarrow \mathcal{D}_i$ and $\mathcal{D}_i \Longrightarrow \mathcal{C}_i$.

Proof of the implication $\mathcal{C}_i \Longrightarrow \mathcal{E}_i$: Assume that Assertion $\mathcal{C}_i$ holds. Let S be an NE-set of T. We will prove that $\left(\operatorname{cont}\left(T|_S\right)\right)_i \geq \left(\operatorname{cont}\left(T|_S\right)\right)_{i+1}$.

Assume the contrary. Thus, $\left(\operatorname{cont}\left(T|_S\right)\right)_i < \left(\operatorname{cont}\left(T|_S\right)\right)_{i+1}$. In other words, $\left(\operatorname{cont}\left(T|_S\right)\right)_{i+1} > \left(\operatorname{cont}\left(T|_S\right)\right)_i$. Recall that

$$\left(\operatorname{cont}\left(T|_S\right)\right)_i = \left|\left(T|_S\right)^{-1}(i)\right| = \left(\text{the number of entries } i \text{ in } T|_S\right)$$

and

$$\left(\operatorname{cont}\left(T|_S\right)\right)_{i+1} = \left|\left(T|_S\right)^{-1}(i+1)\right| = \left(\text{the number of entries } i+1 \text{ in } T|_S\right).$$

Hence,

$$\begin{aligned}\left(\text{the number of entries } i+1 \text{ in } T|_S\right) &= \left(\operatorname{cont}\left(T|_S\right)\right)_{i+1} > \left(\operatorname{cont}\left(T|_S\right)\right)_i \\ &= \left(\text{the number of entries } i \text{ in } T|_S\right) \geq 0.\end{aligned}$$

Hence, there exists at least one entry $i+1$ in $T|_S$. In other words, there exists at least one cell $\mathbf{c} \in S$ such that the entry of T in $\mathbf{c}$ equals $i+1$. Let $\mathbf{d}$ be the **leftmost** such cell $\mathbf{c}$ (or one of the leftmost, if there are

(by (13.94.11), applied to $(i+1, j+1)$ instead of (i,j)). Hence, Assertion $\mathcal{B}'$ is equivalent to the following Assertion $\mathcal{B}''$:

- *Assertion $\mathcal{B}''$:* The inequality

$$(\text{the number of all entries } j+1 \text{ in the first } i+1 \text{ rows of } T) \leq (\text{the number of all entries } j \text{ in the first } i \text{ rows of } T)$$

holds for all $(i,j) \in \mathbb{N} \times \{1,2,3,\ldots\}$.

This Assertion $\mathcal{B}''$, in turn, is equivalent to the following Assertion $\mathcal{B}'''$:

- *Assertion $\mathcal{B}'''$:* The inequality

$$(\text{the number of all entries } i+1 \text{ in the first } j \text{ rows of } T) \leq (\text{the number of all entries } i \text{ in the first } j-1 \text{ rows of } T)$$

holds for all $(i,j) \in \{1,2,3,\ldots\} \times \{1,2,3,\ldots\}$.

(Indeed, Assertion $\mathcal{B}'''$ is obtained from Assertion $\mathcal{B}''$ by substituting $(j-1, i)$ for the index (i,j).) But Assertion $\mathcal{B}'''$ is clearly equivalent to Assertion $\mathcal{D}$.

We thus have found the chain of equivalences $\mathcal{B}' \Longleftrightarrow \mathcal{B}'' \Longleftrightarrow \mathcal{B}''' \Longleftrightarrow \mathcal{D}$. Thus, Assertion $\mathcal{B}'$ is equivalent to Assertion $\mathcal{D}$, qed.

several of them[749]). Thus, we have $\mathbf{d} \in S$, and the entry of T in $\mathbf{d}$ equals $i+1$. Let j be the column in which this cell $\mathbf{d}$ lies. Assertion $\mathcal{C}_i$ then yields $\left(\operatorname{cont}\left(T|_{\operatorname{cols}\geq j}\right)\right)_i \geq \left(\operatorname{cont}\left(T|_{\operatorname{cols}\geq j}\right)\right)_{i+1}$. It is rather clear that $\left(\operatorname{cont}\left(T|_{\operatorname{cols}\geq j}\right)\right)_{i+1} \geq \left(\operatorname{cont}\left(T|_S\right)\right)_{i+1}$ [750].

We shall now show that $\left(\operatorname{cont}\left(T|_S\right)\right)_i \geq \left(\operatorname{cont}\left(T|_{\operatorname{cols}\geq j}\right)\right)_i$. Indeed, let $\mathbf{c}$ be a cell of $T|_{\operatorname{cols}\geq j}$ such that the entry of T in $\mathbf{c}$ equals i. We shall show that $\mathbf{c} \in S$.

Write the cell $\mathbf{d}$ as $\mathbf{d} = (x, j)$ for some positive integer x (this is possible, since $\mathbf{d}$ lies in column j). Write the cell $\mathbf{c}$ as $\mathbf{c} = (x', y')$ for two positive integers x' and y'. Then, y' is the column in which the cell $\mathbf{c}$ lies. Since this column is one of the columns $j, j+1, j+2, \ldots$ (because $\mathbf{c}$ is a cell of $T|_{\operatorname{cols}\geq j}$), this yields that $y' \in \{j, j+1, j+2, \ldots\}$, so that $y' \geq j$.

If $x' \geq x$, then the cell (x', y') lies southeast of the cell (x, j) (since $x' \geq x$ and $y' \geq j$). Since $(x', y') = \mathbf{c}$ and $(x, j) = \mathbf{d}$, this rewrites as follows: If $x' \geq x$, then the cell $\mathbf{c}$ lies southeast of the cell $\mathbf{d}$. Hence, if $x' \geq x$, then (13.94.9) yields

$$(\text{the entry of } T \text{ in cell } \mathbf{c}) \geq (\text{the entry of } T \text{ in cell } \mathbf{d}) = i+1,$$

which contradicts $(\text{the entry of } T \text{ in cell } \mathbf{c}) = i < i+1$. Therefore, we cannot have $x' \geq x$. We thus have $x' < x$.

Since $x' < x$ and $y' \geq j$, the cell (x', y') lies northeast of the cell (x, j). In other words, the cell $\mathbf{c}$ lies northeast of the cell $\mathbf{d}$ (since $(x', y') = \mathbf{c}$ and $(x, j) = \mathbf{d}$). Since $\mathbf{c}$ is a cell of T and since $\mathbf{d} \in S$, this yields that $\mathbf{c} \in S$ as well (since S is an NE-set).

Now, forget that we fixed $\mathbf{c}$. We thus have shown that if $\mathbf{c}$ is a cell of $T|_{\operatorname{cols}\geq j}$ such that the entry of T in $\mathbf{c}$ equals i, then $\mathbf{c} \in S$. Hence, the set

$$\left\{\mathbf{c} \text{ is a cell of } T|_{\operatorname{cols}\geq j} \mid \text{the entry of } T \text{ in } \mathbf{c} \text{ equals } i\right\}$$

is a subset of the set

$$\{\mathbf{c} \in S \mid \text{the entry of } T \text{ in } \mathbf{c} \text{ equals } i\}.$$

As a consequence,

$$\begin{aligned}
&\left|\left\{\mathbf{c} \text{ is a cell of } T|_{\operatorname{cols}\geq j} \mid \text{the entry of } T \text{ in } \mathbf{c} \text{ equals } i\right\}\right| \\
&\leq |\{\mathbf{c} \in S \mid \text{the entry of } T \text{ in } \mathbf{c} \text{ equals } i\}| \\
&= (\text{the number of entries } i \text{ in } T|_S) = \left(\operatorname{cont}\left(T|_S\right)\right)_i .
\end{aligned}$$

Since

$$\begin{aligned}
&\left|\left\{\mathbf{c} \text{ is a cell of } T|_{\operatorname{cols}\geq j} \mid \text{the entry of } T \text{ in } \mathbf{c} \text{ equals } i\right\}\right| \\
&= \left(\text{the number of entries } i \text{ in } T|_{\operatorname{cols}\geq j}\right) = \left|\left(T|_{\operatorname{cols}\geq j}\right)^{-1}(i)\right| = \left(\operatorname{cont}\left(T|_{\operatorname{cols}\geq j}\right)\right)_i
\end{aligned}$$

[749] A moment's thought reveals that there cannot be several of them, but we don't actually need to think about this.

[750] *Proof.* We have defined $\mathbf{d}$ as the leftmost cell $\mathbf{c} \in S$ such that the entry of T in $\mathbf{c}$ equals $i+1$. Hence, if $\mathbf{c} \in S$ is any cell such that the entry of T in $\mathbf{c}$ equals $i+1$, then $\mathbf{c}$ lies in the same column as $\mathbf{d}$ or in some column further right. Since $\mathbf{d}$ lies in column j, this rewrites as follows: If $\mathbf{c} \in S$ is any cell such that the entry of T in $\mathbf{c}$ equals $i+1$, then $\mathbf{c}$ lies in column j or in some column further right. In other words, if $\mathbf{c} \in S$ is any cell such that the entry of T in $\mathbf{c}$ equals $i+1$, then $\mathbf{c}$ lies in one of the columns $j, j+1, j+2, \ldots$. In other words, if $\mathbf{c} \in S$ is any cell such that the entry of T in $\mathbf{c}$ equals $i+1$, then $\mathbf{c}$ is a cell of $T|_{\operatorname{cols}\geq j}$. Thus, the set

$$\{\mathbf{c} \in S \mid \text{the entry of } T \text{ in } \mathbf{c} \text{ equals } i+1\}$$

is a subset of the set

$$\left\{\mathbf{c} \text{ is a cell of } T|_{\operatorname{cols}\geq j} \mid \text{the entry of } T \text{ in } \mathbf{c} \text{ equals } i+1\right\}.$$

As a consequence,

$$\begin{aligned}
&|\{\mathbf{c} \in S \mid \text{the entry of } T \text{ in } \mathbf{c} \text{ equals } i+1\}| \\
&\leq \left|\left\{\mathbf{c} \text{ is a cell of } T|_{\operatorname{cols}\geq j} \mid \text{the entry of } T \text{ in } \mathbf{c} \text{ equals } i+1\right\}\right| \\
&= \left(\text{the number of entries } i+1 \text{ in } T|_{\operatorname{cols}\geq j}\right) = \left|\left(T|_{\operatorname{cols}\geq j}\right)^{-1}(i+1)\right| = \left(\operatorname{cont}\left(T|_{\operatorname{cols}\geq j}\right)\right)_{i+1}
\end{aligned}$$

(because $\left(\operatorname{cont}\left(T|_{\operatorname{cols}\geq j}\right)\right)_{i+1} = \left|\left(T|_{\operatorname{cols}\geq j}\right)^{-1}(i+1)\right|$ (by the definition of $\operatorname{cont}\left(T|_{\operatorname{cols}\geq j}\right)$)). Since

$$\begin{aligned}
&|\{\mathbf{c} \in S \mid \text{the entry of } T \text{ in } \mathbf{c} \text{ equals } i+1\}| \\
&= (\text{the number of entries } i+1 \text{ in } T|_S) = \left(\operatorname{cont}\left(T|_S\right)\right)_{i+1},
\end{aligned}$$

this rewrites as $\left(\operatorname{cont}\left(T|_S\right)\right)_{i+1} \leq \left(\operatorname{cont}\left(T|_{\operatorname{cols}\geq j}\right)\right)_{i+1}$. Hence, $\left(\operatorname{cont}\left(T|_{\operatorname{cols}\geq j}\right)\right)_{i+1} \geq \left(\operatorname{cont}\left(T|_S\right)\right)_{i+1}$, qed.

(because the definition of $\operatorname{cont}\left(T|_{\operatorname{cols}\geq j}\right)$ shows that $\left(\operatorname{cont}\left(T|_{\operatorname{cols}\geq j}\right)\right)_i = \left|\left(T|_{\operatorname{cols}\geq j}\right)^{-1}(i)\right|$), this rewrites as $\left(\operatorname{cont}\left(T|_{\operatorname{cols}\geq j}\right)\right)_i \leq \left(\operatorname{cont}\left(T|_S\right)\right)_i$. In other words, $\left(\operatorname{cont}\left(T|_S\right)\right)_i \geq \left(\operatorname{cont}\left(T|_{\operatorname{cols}\geq j}\right)\right)_i$. Altogether,

$$\left(\operatorname{cont}\left(T|_S\right)\right)_i \geq \left(\operatorname{cont}\left(T|_{\operatorname{cols}\geq j}\right)\right)_i \geq \left(\operatorname{cont}\left(T|_{\operatorname{cols}\geq j}\right)\right)_{i+1} \geq \left(\operatorname{cont}\left(T|_S\right)\right)_{i+1},$$

which contradicts $\left(\operatorname{cont}\left(T|_S\right)\right)_i < \left(\operatorname{cont}\left(T|_S\right)\right)_{i+1}$. This contradiction shows that our assumption was wrong. Hence, $\left(\operatorname{cont}\left(T|_S\right)\right)_i \geq \left(\operatorname{cont}\left(T|_S\right)\right)_{i+1}$.

Now, forget that we fixed S. We thus have proven that for every NE-set S of T, we have $\left(\operatorname{cont}\left(T|_S\right)\right)_i \geq \left(\operatorname{cont}\left(T|_S\right)\right)_{i+1}$. Thus, Assertion $\mathcal{E}_i$ holds. This proves the implication $\mathcal{C}_i \Longrightarrow \mathcal{E}_i$.

Proof of the implication $\mathcal{E}_i \Longrightarrow \mathcal{F}_i$: Assume that Assertion $\mathcal{E}_i$ holds. Let v be a prefix of the Semitic reading word of T. We shall prove that there are at least as many i's among the letters of v as there are $(i+1)$'s among them.

For every $i \in \{1,2,3,...\}$, let r_i be the word obtained by reading the i-th row of T from right to left. Then, the Semitic reading word of T is the concatenation $r_1 r_2 r_3 \cdots$ (according to its definition). Hence, every prefix of this Semitic reading word must have the form $r_1 r_2 \cdots r_k s$ for some $k \in \{0,1,2,...\}$ and some prefix s of r_{k+1}. In particular, v must have this form (since v is a prefix of the Semitic reading word of T). In other words, there exists some $k \in \{0,1,2,...\}$ and some prefix s of r_{k+1} such that $v = r_1 r_2 \cdots r_k s$. Consider this k and this s.

Let ℓ be the length of s. Since s is a prefix of r_{k+1} and has length ℓ, it is evident that the word s consists of the first ℓ letters of r_{k+1}. These first ℓ letters of r_{k+1} are the rightmost ℓ entries of the $(k+1)$-st row of T (since r_{k+1} is the word obtained by reading the $(k+1)$-st row of T from right to left). Thus, the word s consists of the rightmost ℓ entries of the $(k+1)$-st row of T. Hence, the word $r_1 r_2 \cdots r_k s$ consists of all entries of the first k rows of T and the rightmost ℓ entries of the $(k+1)$-st row of T. In other words, the letters of the word $r_1 r_2 \cdots r_k s$ are precisely all entries of the first k rows of T and the rightmost ℓ entries of the $(k+1)$-st row of T.

Let S be the set which consists of all cells of the first k rows of T and the rightmost ℓ cells of the $(k+1)$-st row of T. Then, S is an NE-set of T, and therefore we have $\left(\operatorname{cont}\left(T|_S\right)\right)_i \geq \left(\operatorname{cont}\left(T|_S\right)\right)_{i+1}$ (by Assertion $\mathcal{E}_i$).

Now, $\left(\operatorname{cont}\left(T|_S\right)\right)_i = \left|\left(T|_S\right)^{-1}(i)\right|$ (by the definition of $\operatorname{cont}\left(T|_S\right)$), so that $\left(\operatorname{cont}\left(T|_S\right)\right)_i$ is the number of entries i in $T|_S$. In other words, $\left(\operatorname{cont}\left(T|_S\right)\right)_i$ is the number of i's among the entries of $T|_S$. But since the entries of $T|_S$ are precisely the letters of v [751], this rewrites as follows: $\left(\operatorname{cont}\left(T|_S\right)\right)_i$ is the number of i's among the letters of v. So we have

$$\left(\operatorname{cont}\left(T|_S\right)\right)_i = (\text{the number of } i\text{'s among the letters of } v).$$

The same argument, with $i+1$ in place of i, shows that

$$\left(\operatorname{cont}\left(T|_S\right)\right)_{i+1} = (\text{the number of } (i+1)\text{'s among the letters of } v).$$

Hence,

$$\begin{aligned}(\text{the number of } i\text{'s among the letters of } v) &= \left(\operatorname{cont}\left(T|_S\right)\right)_i \geq \left(\operatorname{cont}\left(T|_S\right)\right)_{i+1} \\ &= (\text{the number of } (i+1)\text{'s among the letters of } v).\end{aligned}$$

In other words, there are at least as many i's among the letters of v as there are $(i+1)$'s among them.

Now, forget that we fixed v. We thus have shown that for every prefix v of the Semitic reading word of T, there are at least as many i's among the letters of v as there are $(i+1)$'s among them. In other words, Assertion $\mathcal{F}_i$ holds. We thus have shown the implication $\mathcal{E}_i \Longrightarrow \mathcal{F}_i$.

Proof of the implication $\mathcal{F}_i \Longrightarrow \mathcal{D}_i$: Assume that Assertion $\mathcal{F}_i$ holds.

For every $i \in \{1,2,3,...\}$, let r_i be the word obtained by reading the i-th row of T from right to left. Then, the Semitic reading word of T is the concatenation $r_1 r_2 r_3 \cdots$ (according to its definition).

Now, let j be a positive integer. The word r_j is the word obtained by reading the j-th row of T from right to left. Hence, the letters of this word r_j are in (weakly) decreasing order (since the entries of T increase

[751] *Proof.* The set S consists of all cells of the first k rows of T and the rightmost ℓ cells of the $(k+1)$-st row of T. Hence, the entries of $T|_S$ are precisely all entries of the first k rows of T and the rightmost ℓ entries of the $(k+1)$-st row of T. Comparing this with the fact that the letters of the word $r_1 r_2 \cdots r_k s$ are precisely all entries of the first k rows of T and the rightmost ℓ entries of the $(k+1)$-st row of T, we conclude the following: The entries of $T|_S$ are precisely the letters of $r_1 r_2 \cdots r_k s$. In other words, the entries of $T|_S$ are precisely the letters of v (since $v = r_1 r_2 \cdots r_k s$), qed.

weakly left-to-right in rows). Thus, the subword of r_j consisting of all letters $> i$ in r_j is a prefix of r_j. Denote this prefix by s. By the definition of s, all letters of s are $> i$, so that

$$(\text{the number of } i\text{'s among the letters of } s) = 0.$$

Also, by the definition of s, the word s consists of all letters $> i$ in r_j. As a consequence,

$$(\text{the number of } (i+1)\text{'s among the letters of } s) = (\text{the number of } (i+1)\text{'s among the letters of } r_j).$$

Since s is a prefix of r_j, it is clear that the word $r_1 r_2 \cdots r_{j-1} s$ is a prefix of the word $r_1 r_2 r_3 \cdots$. In other words, the word $r_1 r_2 \cdots r_{j-1} s$ is a prefix of the Semitic reading word of T (since the Semitic reading word of T is the concatenation $r_1 r_2 r_3 \cdots$). Hence, there are at least as many i's among the letters of $r_1 r_2 \cdots r_{j-1} s$ as there are $(i+1)$'s among them (by Assertion $\mathcal{F}_i$, applied to $v = r_1 r_2 \cdots r_{j-1} s$). In other words,

$$\begin{aligned}&(\text{the number of } i\text{'s among the letters of } r_1 r_2 \cdots r_{j-1} s)\\ &\geq (\text{the number of } (i+1)\text{'s among the letters of } r_1 r_2 \cdots r_{j-1} s).\end{aligned}$$

Since

$$\begin{aligned}&(\text{the number of } i\text{'s among the letters of } r_1 r_2 \cdots r_{j-1} s)\\ &= \sum_{k=1}^{j-1} (\text{the number of } i\text{'s among the letters of } r_k) + \underbrace{(\text{the number of } i\text{'s among the letters of } s)}_{=0}\\ &= \sum_{k=1}^{j-1} \underbrace{(\text{the number of } i\text{'s among the letters of } r_k)}_{\substack{=(\text{the number of entries } i \text{ in the } k\text{-th row of } T)\\ (\text{since } r_k \text{ is the word obtained by reading the } k\text{-th row of } T \text{ from right to left})}}\\ &= \sum_{k=1}^{j-1} (\text{the number of entries } i \text{ in the } k\text{-th row of } T)\\ &= (\text{the number of entries } i \text{ in the first } j-1 \text{ rows of } T)\end{aligned}$$

and

$$\begin{aligned}&(\text{the number of } (i+1)\text{'s among the letters of } r_1 r_2 \cdots r_{j-1} s)\\ &= \sum_{k=1}^{j-1} (\text{the number of } (i+1)\text{'s among the letters of } r_k) + \underbrace{(\text{the number of } (i+1)\text{'s among the letters of } s)}_{=(\text{the number of } (i+1)\text{'s among the letters of } r_j)}\\ &= \sum_{k=1}^{j-1} (\text{the number of } (i+1)\text{'s among the letters of } r_k) + (\text{the number of } (i+1)\text{'s among the letters of } r_j)\\ &= \sum_{k=1}^{j} \underbrace{(\text{the number of } (i+1)\text{'s among the letters of } r_k)}_{\substack{=(\text{the number of entries } i+1 \text{ in the } k\text{-th row of } T)\\ (\text{since } r_k \text{ is the word obtained by reading the } k\text{-th row of } T \text{ from right to left})}}\\ &= \sum_{k=1}^{j} (\text{the number of entries } i+1 \text{ in the } k\text{-th row of } T)\\ &= (\text{the number of entries } i+1 \text{ in the first } j \text{ rows of } T),\end{aligned}$$

this rewrites as follows:

$$\begin{aligned}&(\text{the number of entries } i \text{ in the first } j-1 \text{ rows of } T)\\ &\geq (\text{the number of entries } i+1 \text{ in the first } j \text{ rows of } T).\end{aligned}$$

In other words, the number of entries $i+1$ in the first j rows of T is $\leq$ to the number of entries i in the first $j-1$ rows of T.

Now, forget that we fixed j. We thus have shown that for every positive integer j, the number of entries $i+1$ in the first j rows of T is $\leq$ to the number of entries i in the first $j-1$ rows of T. In other words, Assertion $\mathcal{D}_i$ holds. Thus, we have shown the implication $\mathcal{F}_i \Longrightarrow \mathcal{D}_i$.

Proof of the implication $\mathcal{D}_i \Longrightarrow \mathcal{C}_i$: Assume that Assertion $\mathcal{D}_i$ holds. We want to prove that Assertion $\mathcal{C}_i$ holds. In other words, we want to prove that, for every positive integer j, we have $\left(\operatorname{cont}\left(T|_{\operatorname{cols}\geq j}\right)\right)_i \geq \left(\operatorname{cont}\left(T|_{\operatorname{cols}\geq j}\right)\right)_{i+1}$.

In order to prove this, we assume the contrary. That is, there exists a positive integer j such that we don't have $\left(\operatorname{cont}\left(T|_{\operatorname{cols}\geq j}\right)\right)_i \geq \left(\operatorname{cont}\left(T|_{\operatorname{cols}\geq j}\right)\right)_{i+1}$. Consider the **minimal** such j. Then, we don't have $\left(\operatorname{cont}\left(T|_{\operatorname{cols}\geq j}\right)\right)_i \geq \left(\operatorname{cont}\left(T|_{\operatorname{cols}\geq j}\right)\right)_{i+1}$; but

$$\text{for every } k \in \{1,2,...,j-1\}\text{, we do have } \left(\operatorname{cont}\left(T|_{\operatorname{cols}\geq k}\right)\right)_i \geq \left(\operatorname{cont}\left(T|_{\operatorname{cols}\geq k}\right)\right)_{i+1}. \tag{13.94.13}$$

Applying (13.94.13) to $k = j-1$, we obtain:

$$\text{if } j-1 \text{ is positive, then we do have } \left(\operatorname{cont}\left(T|_{\operatorname{cols}\geq j-1}\right)\right)_i \geq \left(\operatorname{cont}\left(T|_{\operatorname{cols}\geq j-1}\right)\right)_{i+1}. \tag{13.94.14}$$

We don't have $\left(\operatorname{cont}\left(T|_{\operatorname{cols}\geq j}\right)\right)_i \geq \left(\operatorname{cont}\left(T|_{\operatorname{cols}\geq j}\right)\right)_{i+1}$. Thus,

$$\left(\operatorname{cont}\left(T|_{\operatorname{cols}\geq j}\right)\right)_i < \left(\operatorname{cont}\left(T|_{\operatorname{cols}\geq j}\right)\right)_{i+1}. \tag{13.94.15}$$

In other words, $\left(\operatorname{cont}\left(T|_{\operatorname{cols}\geq j}\right)\right)_{i+1} > \left(\operatorname{cont}\left(T|_{\operatorname{cols}\geq j}\right)\right)_i$. Recall that

$$\left(\operatorname{cont}\left(T|_{\operatorname{cols}\geq j}\right)\right)_i = \left(\text{the number of entries } i \text{ in } T|_{\operatorname{cols}\geq j}\right) \tag{13.94.16}$$

and

$$\left(\operatorname{cont}\left(T|_{\operatorname{cols}\geq j}\right)\right)_{i+1} = \left(\text{the number of entries } i+1 \text{ in } T|_{\operatorname{cols}\geq j}\right).$$

Hence,

$$\begin{aligned}\left(\text{the number of entries } i+1 \text{ in } T|_{\operatorname{cols}\geq j}\right) &= \left(\operatorname{cont}\left(T|_{\operatorname{cols}\geq j}\right)\right)_{i+1} > \left(\operatorname{cont}\left(T|_{\operatorname{cols}\geq j}\right)\right)_i \\ &= \left(\text{the number of entries } i \text{ in } T|_{\operatorname{cols}\geq j}\right) \geq 0.\end{aligned}$$

Hence, there exists at least one entry $i+1$ in $T|_{\operatorname{cols}\geq j}$. In other words, there exists at least one cell $\mathbf{c}$ of $T|_{\operatorname{cols}\geq j}$ such that the entry of $T|_{\operatorname{cols}\geq j}$ in $\mathbf{c}$ equals $i+1$. In other words, there exists at least one cell $\mathbf{c}$ of $T|_{\operatorname{cols}\geq j}$ such that the entry of T in $\mathbf{c}$ equals $i+1$ (since the entry of $T|_{\operatorname{cols}\geq j}$ in $\mathbf{c}$, when it is defined, equals the entry of T in $\mathbf{c}$). Let r be the **bottommost** row which contains such a cell $\mathbf{c}$, and let (x,y) be the **leftmost** such cell $\mathbf{c}$ on this row. Thus, (x,y) is a cell of $T|_{\operatorname{cols}\geq j}$ such that the entry of T in (x,y) equals $i+1$. Also, the cell (x,y) lies in the r-th row; that is, $r = x$.

The cell (x,y) must lie in one of the columns $j, j+1, j+2, ...$ (since it is a cell of $T|_{\operatorname{cols}\geq j}$), so that we have $y \geq j$. Also, (x,y) is a cell of T (since it is a cell of $T|_{\operatorname{cols}\geq j}$).

Applying Assertion $\mathcal{D}_i$ to x instead of j, we see that the number of entries $i+1$ in the first x rows of T is $\leq$ to the number of entries i in the first $x-1$ rows of T. In other words,

$$\begin{aligned}&\left(\text{the number of entries } i+1 \text{ in the first } x \text{ rows of } T\right) \\ &\leq \left(\text{the number of entries } i \text{ in the first } x-1 \text{ rows of } T\right).\end{aligned} \tag{13.94.17}$$

It is easy to see that

$$\left(\operatorname{cont}\left(T|_{\operatorname{cols}\geq j}\right)\right)_{i+1} \leq \left(\text{the number of entries } i+1 \text{ in the first } x \text{ rows of } T\right) \tag{13.94.18}$$

[752].

We shall now show that

$$\left(\text{the number of entries } i \text{ in the first } x-1 \text{ rows of } T\right) \leq \left(\operatorname{cont}\left(T|_{\operatorname{cols}\geq j}\right)\right)_i. \tag{13.94.19}$$

[752] *Proof.* Let $\mathbf{c}$ be a cell of $T|_{\operatorname{cols}\geq j}$ such that the entry of T in $\mathbf{c}$ equals $i+1$. Since r is the bottommost row which contains such a cell $\mathbf{c}$, it is clear that the cell $\mathbf{c}$ must be in the r-th row or in a row further north. In other words, the cell $\mathbf{c}$ must lie in one of the first r rows of T. Since $r = x$, this rewrites as follows: The cell $\mathbf{c}$ must lie in one of the first x rows of T.

Now, let us forget that we fixed $\mathbf{c}$. We thus have proven that if $\mathbf{c}$ is a cell of $T|_{\operatorname{cols}\geq j}$ such that the entry of T in $\mathbf{c}$ equals $i+1$, then the cell $\mathbf{c}$ must lie in one of the first x rows of T. Hence, the set

$$\left\{\mathbf{c} \text{ is a cell of } T|_{\operatorname{cols}\geq j} \mid \text{the entry of } T \text{ in } \mathbf{c} \text{ equals } i+1\right\}$$

is a subset of the set

$$\{\mathbf{c} \text{ is a cell of } T \mid \text{the entry of } T \text{ in } \mathbf{c} \text{ equals } i+1;\ \text{the cell } \mathbf{c} \text{ lies in one of the first } x \text{ rows of } T\}.$$

Indeed, let $\mathbf{c}$ be a cell of T such that the entry of T in $\mathbf{c}$ equals i and such that $\mathbf{c}$ lies in one of the first $x-1$ rows of T. We shall show that $\mathbf{c}$ is a cell of $T|_{\text{cols}\geq j}$.

Assume the contrary. Thus, $\mathbf{c}$ is not a cell of $T|_{\text{cols}\geq j}$.

Let us write the cell $\mathbf{c}$ as $\mathbf{c} = (x', y')$ for some integers x' and y'. Then, the cell $\mathbf{c}$ lies in row x', and thus this row x' must be one of the first $x-1$ rows of T (since we know that $\mathbf{c}$ lies in one of the first $x-1$ rows of T). In other words, $x' \leq x-1$. Hence, $x' \leq x-1 < x$ and thus $x > x'$ and $x \neq x'$ and $x \geq x'$.

The cell $\mathbf{c}$ lies in the y'-th column (since $\mathbf{c} = (x', y')$). Hence, if $y' \geq j$, then the cell $\mathbf{c}$ lies in one of the columns $j, j+1, j+2, \ldots$ and therefore is a cell of $T|_{\text{cols}\geq j}$, which contradicts our assumption that $\mathbf{c}$ is not a cell of $T|_{\text{cols}\geq j}$. Thus, we cannot have $y' \geq j$. We therefore have $y' < j$. That is, $y' \leq j-1$ (since y' and j are integers), so that $j-1 \geq y'$.

Now, the cell (x, y) lies southeast of the cell $(x, j-1)$ (since $x \geq x$ and $y \geq j \geq j-1$), which in turn lies southeast of the cell (x', y') (since $x \geq x'$ and $j-1 \geq y'$). Since both cells (x, y) and (x', y') are cells of T (indeed, (x, y) is known to be a cell of T, and $(x', y') = \mathbf{c}$ also is a cell of T), this yields that the intermediate cell $(x, j-1)$ is also a cell of T [753]. Thus, $j-1$ is a positive integer. Moreover, since the cell (x, y) lies southeast of the cell $(x, j-1)$, we can apply (13.94.9) to (x, y) and $(x, j-1)$ instead of $\mathbf{c}$ and $\mathbf{d}$. We thus obtain

$$(\text{the entry of } T \text{ in cell } (x, y)) \geq (\text{the entry of } T \text{ in cell } (x, j-1)).$$

Since $(\text{the entry of } T \text{ in cell } (x, y)) = i+1$, this rewrites as

$$i+1 \geq (\text{the entry of } T \text{ in cell } (x, j-1)). \tag{13.94.20}$$

But the cell $(x, j-1)$ lies southeast of the cell (x', y') and not on the same row as (x', y') (since $x \neq x'$). Since $(x', y') = \mathbf{c}$, this rewrites as follows: The cell $(x, j-1)$ lies southeast of the cell $\mathbf{c}$ and not on the same row as $\mathbf{c}$. Hence, (13.94.10) (applied to $(x, j-1)$ and $\mathbf{c}$ instead of $\mathbf{c}$ and $\mathbf{d}$) yields

$$(\text{the entry of } T \text{ in cell } (x, j-1)) > (\text{the entry of } T \text{ in cell } \mathbf{c}) = i,$$

so that

$$(\text{the entry of } T \text{ in cell } (x, j-1)) \geq i+1$$

(since the entry of T in cell $(x, j-1)$ and the number i are integers). Combined with (13.94.20), this yields

$$(\text{the entry of } T \text{ in cell } (x, j-1)) = i+1.$$

Thus, the number $i+1$ appears in the $(j-1)$-th column of T (since the cell $(x, j-1)$ lies in the $(j-1)$-th column of T). In other words,

$$(\text{the number of entries } i+1 \text{ in the } (j-1)\text{-th column of } T) \geq 1. \tag{13.94.21}$$

On the other hand, the entries of T increase strictly top-to-bottom along columns. In particular, in every given column of T, the entries are distinct. Applied to the $(j-1)$-th column, this shows that the entries of the $(j-1)$-th column of T are distinct. In particular, for every $k \in \{1, 2, 3, \ldots\}$, the number k appears at most once in the $(j-1)$-th column of T. Applying this to $k = i$, we conclude that the number i appears at most once in the $(j-1)$-th column of T. That is,

$$(\text{the number of entries } i \text{ in the } (j-1)\text{-th column of } T) \leq 1. \tag{13.94.22}$$

Hence,

$$\begin{aligned}
&\left|\left\{\mathbf{c} \text{ is a cell of } T|_{\text{cols}\geq j} \ \mid\ \text{the entry of } T \text{ in } \mathbf{c} \text{ equals } i+1\right\}\right| \\
&\leq |\{\mathbf{c} \text{ is a cell of } T \ \mid\ \text{the entry of } T \text{ in } \mathbf{c} \text{ equals } i+1; \text{ the cell } \mathbf{c} \text{ lies in one of the first } x \text{ rows of } T\}| \\
&= (\text{the number of entries } i+1 \text{ in the first } x \text{ rows of } T).
\end{aligned}$$

But

$$\begin{aligned}
\left(\text{cont}\left(T|_{\text{cols}\geq j}\right)\right)_{i+1} &= \left(\text{the number of entries } i+1 \text{ in } T|_{\text{cols}\geq j}\right) \\
&= \left|\left\{\mathbf{c} \text{ is a cell of } T|_{\text{cols}\geq j} \ \mid\ \text{the entry of } T \text{ in } \mathbf{c} \text{ equals } i+1\right\}\right| \\
&\leq (\text{the number of entries } i+1 \text{ in the first } x \text{ rows of } T),
\end{aligned}$$

qed.

[753] Here, we are using the following fact: If α, β and γ are three cells such that α lies southeast of the cell β, which in turn lies southeast of the cell γ, and if α and γ are cells of T, then β is also a cell of T. This can be easily derived from the fact that T has shape λ/μ.

Now, for every $k \in \{1, 2, 3, ...\}$ and $h \in \{1, 2, 3, ...\}$, we have

$$\begin{aligned}\left(\operatorname{cont}\left(T|_{\operatorname{cols}\geq k}\right)\right)_h &= \left|\left(T|_{\operatorname{cols}\geq k}\right)^{-1}(h)\right| = \left(\text{the number of entries } h \text{ in } T|_{\operatorname{cols}\geq k}\right) \\ &= \left(\text{the number of entries } h \text{ in the columns } k, k+1, k+2, ... \text{ of } T\right). \qquad (13.94.23)\end{aligned}$$

Applying this to $k = j$ and $h = i$, we obtain

$$\left(\operatorname{cont}\left(T|_{\operatorname{cols}\geq j}\right)\right)_i = \left(\text{the number of entries } i \text{ in the columns } j, j+1, j+2, ... \text{ of } T\right). \qquad (13.94.24)$$

But applying (13.94.23) to $k = j-1$ and $h = i$, we obtain

$$\begin{aligned}\left(\operatorname{cont}\left(T|_{\operatorname{cols}\geq j-1}\right)\right)_i &= \left(\text{the number of entries } i \text{ in the columns } j-1, j, j+1, ... \text{ of } T\right) \\ &= \underbrace{\left(\text{the number of entries } i \text{ in the } (j-1)\text{-th column of } T\right)}_{\substack{\leq 1 \\ \text{(by (13.94.22))}}} \\ &\qquad + \underbrace{\left(\text{the number of entries } i \text{ in the columns } j, j+1, j+2, ... \text{ of } T\right)}_{\substack{=\left(\operatorname{cont}\left(T|_{\operatorname{cols}\geq j}\right)\right)_i \\ \text{(by (13.94.24))}}} \\ &\leq 1 + \underbrace{\left(\operatorname{cont}\left(T|_{\operatorname{cols}\geq j}\right)\right)_i}_{\substack{<\left(\operatorname{cont}\left(T|_{\operatorname{cols}\geq j}\right)\right)_{i+1} \\ \text{(by (13.94.15))}}} < 1 + \left(\operatorname{cont}\left(T|_{\operatorname{cols}\geq j}\right)\right)_{i+1}.\end{aligned}$$

Since $\left(\operatorname{cont}\left(T|_{\operatorname{cols}\geq j-1}\right)\right)_i$ and $1 + \left(\operatorname{cont}\left(T|_{\operatorname{cols}\geq j}\right)\right)_{i+1}$ are integers, this yields

$$\left(\operatorname{cont}\left(T|_{\operatorname{cols}\geq j-1}\right)\right)_i \leq \left(1 + \left(\operatorname{cont}\left(T|_{\operatorname{cols}\geq j}\right)\right)_{i+1}\right) - 1 = \left(\operatorname{cont}\left(T|_{\operatorname{cols}\geq j}\right)\right)_{i+1},$$

so that

$$\begin{aligned}\left(\operatorname{cont}\left(T|_{\operatorname{cols}\geq j}\right)\right)_{i+1} &\geq \left(\operatorname{cont}\left(T|_{\operatorname{cols}\geq j-1}\right)\right)_i \geq \left(\operatorname{cont}\left(T|_{\operatorname{cols}\geq j-1}\right)\right)_{i+1} \qquad \text{(by (13.94.14))} \\ &= \left(\text{the number of entries } i+1 \text{ in the columns } j-1, j, j+1, ... \text{ of } T\right) \\ &\qquad \text{(by (13.94.23), applied to } k = j-1 \text{ and } h = i+1\text{)} \\ &= \underbrace{\left(\text{the number of entries } i+1 \text{ in the } (j-1)\text{-th column of } T\right)}_{\substack{\geq 1 \\ \text{(by (13.94.21))}}} \\ &\qquad + \underbrace{\left(\text{the number of entries } i+1 \text{ in the columns } j, j+1, j+2, ... \text{ of } T\right)}_{\substack{=\left(\operatorname{cont}\left(T|_{\operatorname{cols}\geq j}\right)\right)_{i+1} \\ \text{(since } \left(\operatorname{cont}\left(T|_{\operatorname{cols}\geq j}\right)\right)_{i+1} = \text{(the number of entries } i+1 \text{ in the columns } j,j+1,j+2,... \text{ of } T\text{)} \\ \text{(by (13.94.23), applied to } k=j \text{ and } h=i+1\text{))}}} \\ &\geq 1 + \left(\operatorname{cont}\left(T|_{\operatorname{cols}\geq j}\right)\right)_{i+1} > \left(\operatorname{cont}\left(T|_{\operatorname{cols}\geq j}\right)\right)_{i+1},\end{aligned}$$

which is absurd. This contradiction proves that our assumption was wrong, and thus we have proven that $\mathbf{c}$ is a cell of $T|_{\operatorname{cols}\geq j}$.

Now, let us forget that we fixed $\mathbf{c}$. We thus have proven that if $\mathbf{c}$ is a cell of T such that the entry of T in $\mathbf{c}$ equals i and such that $\mathbf{c}$ lies in one of the first $x-1$ rows of T, then $\mathbf{c}$ is a cell of $T|_{\operatorname{cols}\geq j}$. Hence, the set

$$\{\mathbf{c} \text{ is a cell of } T \mid \text{the entry of } T \text{ in } \mathbf{c} \text{ equals } i\text{; the cell } \mathbf{c} \text{ lies in one of the first } x-1 \text{ rows of } T\}$$

is a subset of the set

$$\{\mathbf{c} \text{ is a cell of } T|_{\operatorname{cols}\geq j} \mid \text{the entry of } T \text{ in } \mathbf{c} \text{ equals } i\}.$$

Hence,

$$\begin{aligned}&|\{\mathbf{c} \text{ is a cell of } T \mid \text{the entry of } T \text{ in } \mathbf{c} \text{ equals } i\text{; the cell } \mathbf{c} \text{ lies in one of the first } x-1 \text{ rows of } T\}| \\ &\leq |\{\mathbf{c} \text{ is a cell of } T|_{\operatorname{cols}\geq j} \mid \text{the entry of } T \text{ in } \mathbf{c} \text{ equals } i\}| \\ &= \left(\text{the number of entries } i \text{ in } T|_{\operatorname{cols}\geq j}\right) = \left(\operatorname{cont}\left(T|_{\operatorname{cols}\geq j}\right)\right)_i \qquad \text{(by (13.94.16))}.\end{aligned}$$

Since

$$\begin{aligned}&\left|\left\{\mathbf{c} \text{ is a cell of } T \ \mid \ \text{the entry of } T \text{ in } \mathbf{c} \text{ equals } i\text{; the cell } \mathbf{c} \text{ lies in one of the first } x-1 \text{ rows of } T\right\}\right| \\ &= (\text{the number of entries } i \text{ in the first } x-1 \text{ rows of } T),\end{aligned}$$

this rewrites as

$$(\text{the number of entries } i \text{ in the first } x-1 \text{ rows of } T) \leq \left(\operatorname{cont}\left(T|_{\operatorname{cols}\geq j}\right)\right)_i .$$

Thus, (13.94.19) is proven. Now, (13.94.19) becomes

$$\begin{aligned}&(\text{the number of entries } i \text{ in the first } x-1 \text{ rows of } T) \\ &\leq \left(\operatorname{cont}\left(T|_{\operatorname{cols}\geq j}\right)\right)_i < \left(\operatorname{cont}\left(T|_{\operatorname{cols}\geq j}\right)\right)_{i+1} \qquad (\text{by (13.94.15)}) \\ &\leq (\text{the number of entries } i+1 \text{ in the first } x \text{ rows of } T) \qquad (\text{by (13.94.18)}),\end{aligned}$$

which contradicts (13.94.17). This contradiction proves that our assumption was wrong. Hence, Assertion $\mathcal{C}_i$ must hold. We thus have proven the implication $\mathcal{D}_i \Longrightarrow \mathcal{C}_i$.

We have now proven the four implications $\mathcal{C}_i \Longrightarrow \mathcal{E}_i$, $\mathcal{E}_i \Longrightarrow \mathcal{F}_i$, $\mathcal{F}_i \Longrightarrow \mathcal{D}_i$ and $\mathcal{D}_i \Longrightarrow \mathcal{C}_i$. Combining them, we obtain the equivalence $\mathcal{C}_i \Longleftrightarrow \mathcal{D}_i \Longleftrightarrow \mathcal{E}_i \Longleftrightarrow \mathcal{F}_i$.

Now, let us forget that we fixed i. We thus have proven the equivalence $\mathcal{C}_i \Longleftrightarrow \mathcal{D}_i \Longleftrightarrow \mathcal{E}_i \Longleftrightarrow \mathcal{F}_i$ for every $i \in \{1, 2, 3, ...\}$. As we know, this shows that we have the equivalence $\mathcal{C} \Longleftrightarrow \mathcal{D} \Longleftrightarrow \mathcal{E} \Longleftrightarrow \mathcal{F}$, and this finishes the solution of Exercise 2.9.18(b).

13.95. **Solution to Exercise 2.9.20.** *Solution to Exercise 2.9.20.* (a) The solution to Exercise 2.9.20(a) is a rather straightforward adaptation of the solution of Exercise 2.9.18(b) that we gave above. We leave the details to the reader (who can also look them up in the LaTeX source code of this file, where they appear in a "commentedout" environment starting after this sentence).

(b) Exercise 2.9.20(a) yields the equivalence $\mathcal{C}^{(\kappa)} \Longleftrightarrow \mathcal{D}^{(\kappa)} \Longleftrightarrow \mathcal{E}^{(\kappa)} \Longleftrightarrow \mathcal{F}^{(\kappa)} \Longleftrightarrow \mathcal{G}^{(\kappa)}$. It thus remains to prove the equivalence $\mathcal{G}^{(\kappa)} \Longleftrightarrow \mathcal{H}^{(\kappa)}$. In order to do so, we must prove the implications $\mathcal{G}^{(\kappa)} \Longrightarrow \mathcal{H}^{(\kappa)}$ and $\mathcal{H}^{(\kappa)} \Longrightarrow \mathcal{G}^{(\kappa)}$.

The implication $\mathcal{H}^{(\kappa)} \Longrightarrow \mathcal{G}^{(\kappa)}$ is obvious (since we can just take the S whose existence is guaranteed by Assertion $\mathcal{H}^{(\kappa)}$, and set $\zeta = \tau$). It thus remains to prove the implication $\mathcal{G}^{(\kappa)} \Longrightarrow \mathcal{H}^{(\kappa)}$.

Proof of the implication $\mathcal{G}^{(\kappa)} \Longrightarrow \mathcal{H}^{(\kappa)}$*:* Assume that Assertion $\mathcal{G}^{(\kappa)}$ holds. We want to prove that Assertion $\mathcal{H}^{(\kappa)}$ holds.

We have assumed that Assertion $\mathcal{G}^{(\kappa)}$ holds. In other words, there exist a partition ζ and a column-strict tableau S of shape ζ/κ which satisfies the following property: For any positive integers i and j,

$$\left(\begin{array}{c}\text{the number of entries } i \text{ in the } j\text{-th row of } T \text{ equals} \\ \text{the number of entries } j \text{ in the } i\text{-th row of } S\end{array}\right). \tag{13.95.1}$$

Consider this ζ and this S.

For every $i \in \{1, 2, 3, \ldots\}$, we have

$$\begin{aligned}(\operatorname{cont} T)_i &= \left|T^{-1}(i)\right| \qquad (\text{by the definition of } \operatorname{cont} T) \\ &= (\text{the number of entries } i \text{ in } T) \\ &= \sum_{j=1}^{\infty} \underbrace{(\text{the number of entries } i \text{ in the } j\text{-th row of } T)}_{\substack{=(\text{the number of entries } j \text{ in the } i\text{-th row of } S) \\ (\text{by (13.95.1)})}} \\ &= \sum_{j=1}^{\infty} (\text{the number of entries } j \text{ in the } i\text{-th row of } S) \\ &= (\text{the number of entries in the } i\text{-th row of } S) \\ &= (\text{the length of the } i\text{-th row of the skew partition } \zeta/\kappa) \qquad (\text{since the tableau } S \text{ has shape } \zeta/\kappa) \\ &= \zeta_i - \kappa_i.\end{aligned}$$

Hence, for every $i \in \{1, 2, 3, \ldots\}$, we have

$$(\kappa + \operatorname{cont} T)_i = \kappa_i + \underbrace{(\operatorname{cont} T)_i}_{=\zeta_i - \kappa_i} = \kappa_i + (\zeta_i - \kappa_i) = \zeta_i.$$

In other words, $\kappa + \operatorname{cont} T = \zeta$, so that $\zeta = \kappa + \operatorname{cont} T = \tau$. But S is a column-strict tableau of shape ζ/κ. In other words, S is a column-strict tableau of shape τ/κ (since $\zeta = \tau$).

We thus have constructed a column-strict tableau S of shape τ/κ such that for any positive integers i and j, the property (13.95.1) holds. Therefore, Assertion $\mathcal{H}^{(\kappa)}$ holds. We thus have proven the implication $\mathcal{G}^{(\kappa)} \Longrightarrow \mathcal{H}^{(\kappa)}$. As we have said, this finishes the solution of Exercise 2.9.20(b).

13.96. **Solution to Exercise 2.9.21.** *Solution to Exercise 2.9.21.* (a) If T is a column-strict tableau of shape λ/μ, then we have the following chain of logical equivalences:

$$\begin{aligned} & \left(\text{for all } j \in \{1,2,3,\ldots\}, \text{ the weak composition } \kappa + \operatorname{cont}\left(T|_{\operatorname{cols}\geq j}\right) \text{ is a partition}\right) \\ \Longleftrightarrow\ & \left(\text{Assertion } \mathcal{C}^{(\kappa)} \text{ holds}\right) \qquad \left(\text{because this is how we defined Assertion } \mathcal{C}^{(\kappa)}\right) \\ \Longleftrightarrow\ & \left(\text{the five equivalent assertions } \mathcal{C}^{(\kappa)}, \mathcal{D}^{(\kappa)}, \mathcal{E}^{(\kappa)}, \mathcal{F}^{(\kappa)} \text{ and } \mathcal{G}^{(\kappa)} \text{ hold}\right) \end{aligned} \tag{13.96.1}$$

(because the five assertions $\mathcal{C}^{(\kappa)}$, $\mathcal{D}^{(\kappa)}$, $\mathcal{E}^{(\kappa)}$, $\mathcal{F}^{(\kappa)}$ and $\mathcal{G}^{(\kappa)}$ are equivalent (by Exercise 2.9.20(a))).

We can apply (2.6.3) to $\nu = \kappa$. As a result, we see that

$$s_\kappa s_{\lambda/\mu} = \sum_T s_{\kappa + \operatorname{cont} T},$$

where T runs through all column-strict tableaux of shape λ/μ with the property that for each $j = 1, 2, 3, \ldots$, the weak composition $\kappa + \operatorname{cont}(T|_{\operatorname{cols}\geq j})$ is a partition. In other words,

$$\begin{aligned} s_\kappa s_{\lambda/\mu} &= \underbrace{\sum_{\substack{T \text{ is a column-strict tableau} \\ \text{of shape } \lambda/\mu; \\ \text{for all } j \in \{1,2,3,\ldots\}, \text{ the weak} \\ \text{composition } \kappa + \operatorname{cont}\left(T|_{\operatorname{cols}\geq j}\right) \\ \text{is a partition}}}_{\substack{= \sum \\ T \text{ is a column-strict tableau} \\ \text{of shape } \lambda/\mu; \\ \text{the five equivalent assertions} \\ \mathcal{C}^{(\kappa)},\, \mathcal{D}^{(\kappa)},\, \mathcal{E}^{(\kappa)},\, \mathcal{F}^{(\kappa)} \text{ and } \mathcal{G}^{(\kappa)} \text{ hold} \\ \text{(because of the equivalence (13.96.1))}}} s_{\kappa + \operatorname{cont} T} && (13.96.2) \\ &= \sum_{\substack{T \text{ is a column-strict tableau} \\ \text{of shape } \lambda/\mu; \\ \text{the five equivalent assertions} \\ \mathcal{C}^{(\kappa)},\, \mathcal{D}^{(\kappa)},\, \mathcal{E}^{(\kappa)},\, \mathcal{F}^{(\kappa)} \text{ and } \mathcal{G}^{(\kappa)} \text{ hold}}} s_{\kappa + \operatorname{cont} T}. && (13.96.3) \end{aligned}$$

In other words,

$$s_\kappa s_{\lambda/\mu} = \sum_T s_{\kappa + \operatorname{cont} T},$$

where the sum ranges over all column-strict tableaux T of shape λ/μ satisfying the five equivalent assertions $\mathcal{C}^{(\kappa)}$, $\mathcal{D}^{(\kappa)}$, $\mathcal{E}^{(\kappa)}$, $\mathcal{F}^{(\kappa)}$ and $\mathcal{G}^{(\kappa)}$ introduced in Exercise 2.9.20(a). This solves Exercise 2.9.21(a).

(b) If T is a column-strict tableau of shape λ/μ such that the five equivalent assertions $\mathcal{C}^{(\kappa)}$, $\mathcal{D}^{(\kappa)}$, $\mathcal{E}^{(\kappa)}$, $\mathcal{F}^{(\kappa)}$ and $\mathcal{G}^{(\kappa)}$ hold, then we have

$$\left(s_{\kappa + \operatorname{cont} T}, s_\tau\right)_\Lambda = \delta_{\tau, \kappa + \operatorname{cont} T} \tag{13.96.4}$$

[754].

[754] *Proof of (13.96.4):* We know that the basis $(s_\lambda)_{\lambda \in \operatorname{Par}}$ of Λ is orthonormal with respect to the Hall inner product. In other words, we have

$$\left(s_\alpha, s_\beta\right)_\Lambda = \delta_{\alpha,\beta} \tag{13.96.5}$$

But if T is a column-strict tableau of shape λ/μ satisfying $\tau = \kappa + \operatorname{cont} T$, then we have the following chain of logical equivalences:

$$\begin{aligned} &\left(\text{the five equivalent assertions } \mathcal{C}^{(\kappa)}, \mathcal{D}^{(\kappa)}, \mathcal{E}^{(\kappa)}, \mathcal{F}^{(\kappa)} \text{ and } \mathcal{G}^{(\kappa)} \text{ hold}\right) \\ \iff &\left(\text{the six equivalent assertions } \mathcal{C}^{(\kappa)}, \mathcal{D}^{(\kappa)}, \mathcal{E}^{(\kappa)}, \mathcal{F}^{(\kappa)}, \mathcal{G}^{(\kappa)} \text{ and } \mathcal{H}^{(\kappa)} \text{ hold}\right) \end{aligned} \tag{13.96.6}$$

(because the six assertions $\mathcal{C}^{(\kappa)}$, $\mathcal{D}^{(\kappa)}$, $\mathcal{E}^{(\kappa)}$, $\mathcal{F}^{(\kappa)}$, $\mathcal{G}^{(\kappa)}$ and $\mathcal{H}^{(\kappa)}$ are equivalent (by Exercise 2.9.20(b))).

Now, recall that $s_\mu^\perp (s_\lambda) = s_{\lambda/\mu}$. Applying this to κ and τ instead of μ and λ, we obtain $s_\kappa^\perp (s_\tau) = s_{\tau/\kappa}$.

Now, Proposition 2.8.2(i) (applied to $A = \Lambda$) shows that every $a \in \Lambda$, $f \in \Lambda$ and $g \in \Lambda$ satisfy $\left(g, f^\perp (a)\right)_\Lambda = (fg, a)_\Lambda$. We can apply this to $f = s_\kappa$, $g = s_{\lambda/\mu}$ and $a = s_\tau$. As a result, we obtain $\left(s_{\lambda/\mu}, s_\kappa^\perp (s_\tau)\right)_\Lambda = \left(s_\kappa s_{\lambda/\mu}, s_\tau\right)_\Lambda$. Since $s_\kappa^\perp (s_\tau) = s_{\tau/\kappa}$, this rewrites as follows:

$$\left(s_{\lambda/\mu}, s_{\tau/\kappa}\right)_\Lambda = \left(s_\kappa s_{\lambda/\mu}, s_\tau\right)_\Lambda . \tag{13.96.7}$$

for any $\alpha \in \operatorname{Par}$ and $\beta \in \operatorname{Par}$.

Now, let T be a column-strict tableau of shape λ/μ such that the five equivalent assertions $\mathcal{C}^{(\kappa)}$, $\mathcal{D}^{(\kappa)}$, $\mathcal{E}^{(\kappa)}$, $\mathcal{F}^{(\kappa)}$ and $\mathcal{G}^{(\kappa)}$ hold. Then, Assertion $\mathcal{C}^{(\kappa)}$ holds (since the five equivalent assertions $\mathcal{C}^{(\kappa)}$, $\mathcal{D}^{(\kappa)}$, $\mathcal{E}^{(\kappa)}$, $\mathcal{F}^{(\kappa)}$ and $\mathcal{G}^{(\kappa)}$ hold). In other words, for every positive integer j, the weak composition $\kappa + \operatorname{cont}\left(T|_{\operatorname{cols} \geq j}\right)$ is a partition. Applying this to $j = 1$, we conclude that the weak composition $\kappa + \operatorname{cont}\left(T|_{\operatorname{cols} \geq 1}\right)$ is a partition.

But $T|_{\operatorname{cols} \geq 1}$ is the subtableau which is the restriction of T to the union of its columns $1, 2, 3, \ldots$. In other words, $T|_{\operatorname{cols} \geq 1}$ is the whole tableau T. In other words, $T|_{\operatorname{cols} \geq 1} = T$.

Now, recall that $\kappa + \operatorname{cont}\left(T|_{\operatorname{cols} \geq 1}\right)$ is a partition. In other words, $\kappa + \operatorname{cont} T$ is a partition (since $T|_{\operatorname{cols} \geq 1} = T$). In other words, $\kappa + \operatorname{cont} T \in \operatorname{Par}$. Also, $\tau \in \operatorname{Par}$ (since τ is a partition). Thus, (13.96.5) (applied to $\alpha = \kappa + \operatorname{cont} T$ and $\beta = \tau$) shows that $\left(s_{\kappa + \operatorname{cont} T}, s_\tau\right)_\Lambda = \delta_{\kappa + \operatorname{cont} T, \tau} = \delta_{\tau, \kappa + \operatorname{cont} T}$. This proves (13.96.4).

Thus,

$$\begin{aligned}
\left(s_{\lambda/\mu}, s_{\tau/\kappa}\right)_\Lambda &= \left(s_\kappa s_{\lambda/\mu}, s_\tau\right)_\Lambda = \left(\sum_{\substack{T \text{ is a column-strict tableau} \\ \text{of shape } \lambda/\mu; \\ \text{the five equivalent assertions} \\ \mathcal{C}^{(\kappa)},\ \mathcal{D}^{(\kappa)},\ \mathcal{E}^{(\kappa)},\ \mathcal{F}^{(\kappa)} \text{ and } \mathcal{G}^{(\kappa)} \text{ hold}}} s_{\kappa+\operatorname{cont} T}, s_\tau \right)_\Lambda \qquad \text{(by (13.96.3))} \\
&= \sum_{\substack{T \text{ is a column-strict tableau} \\ \text{of shape } \lambda/\mu; \\ \text{the five equivalent assertions} \\ \mathcal{C}^{(\kappa)},\ \mathcal{D}^{(\kappa)},\ \mathcal{E}^{(\kappa)},\ \mathcal{F}^{(\kappa)} \text{ and } \mathcal{G}^{(\kappa)} \text{ hold}}} \underbrace{\left(s_{\kappa+\operatorname{cont} T}, s_\tau\right)_\Lambda}_{\substack{=\delta_{\tau,\kappa+\operatorname{cont} T} \\ \text{(by (13.96.4))}}} \qquad \text{(since the Hall inner product is } \mathbf{k}\text{-bilinear)} \\
&= \sum_{\substack{T \text{ is a column-strict tableau} \\ \text{of shape } \lambda/\mu; \\ \text{the five equivalent assertions} \\ \mathcal{C}^{(\kappa)},\ \mathcal{D}^{(\kappa)},\ \mathcal{E}^{(\kappa)},\ \mathcal{F}^{(\kappa)} \text{ and } \mathcal{G}^{(\kappa)} \text{ hold}}} \delta_{\tau,\kappa+\operatorname{cont} T} \\
&= \sum_{\substack{T \text{ is a column-strict tableau} \\ \text{of shape } \lambda/\mu; \\ \text{the five equivalent assertions} \\ \mathcal{C}^{(\kappa)},\ \mathcal{D}^{(\kappa)},\ \mathcal{E}^{(\kappa)},\ \mathcal{F}^{(\kappa)} \text{ and } \mathcal{G}^{(\kappa)} \text{ hold}; \\ \tau=\kappa+\operatorname{cont} T}} \underbrace{\delta_{\tau,\kappa+\operatorname{cont} T}}_{\substack{=1 \\ \text{(since } \tau=\kappa+\operatorname{cont} T)}} + \sum_{\substack{T \text{ is a column-strict tableau} \\ \text{of shape } \lambda/\mu; \\ \text{the five equivalent assertions} \\ \mathcal{C}^{(\kappa)},\ \mathcal{D}^{(\kappa)},\ \mathcal{E}^{(\kappa)},\ \mathcal{F}^{(\kappa)} \text{ and } \mathcal{G}^{(\kappa)} \text{ hold}; \\ \tau\neq\kappa+\operatorname{cont} T}} \underbrace{\delta_{\tau,\kappa+\operatorname{cont} T}}_{\substack{=0 \\ \text{(since } \tau\neq\kappa+\operatorname{cont} T)}} \\
&= \sum_{\substack{T \text{ is a column-strict tableau} \\ \text{of shape } \lambda/\mu; \\ \text{the five equivalent assertions} \\ \mathcal{C}^{(\kappa)},\ \mathcal{D}^{(\kappa)},\ \mathcal{E}^{(\kappa)},\ \mathcal{F}^{(\kappa)} \text{ and } \mathcal{G}^{(\kappa)} \text{ hold}; \\ \tau=\kappa+\operatorname{cont} T}} 1 + \underbrace{\sum_{\substack{T \text{ is a column-strict tableau} \\ \text{of shape } \lambda/\mu; \\ \text{the five equivalent assertions} \\ \mathcal{C}^{(\kappa)},\ \mathcal{D}^{(\kappa)},\ \mathcal{E}^{(\kappa)},\ \mathcal{F}^{(\kappa)} \text{ and } \mathcal{G}^{(\kappa)} \text{ hold}; \\ \tau\neq\kappa+\operatorname{cont} T}} 0}_{=0} \\
&= \underbrace{\sum_{\substack{T \text{ is a column-strict tableau} \\ \text{of shape } \lambda/\mu; \\ \text{the five equivalent assertions} \\ \mathcal{C}^{(\kappa)},\ \mathcal{D}^{(\kappa)},\ \mathcal{E}^{(\kappa)},\ \mathcal{F}^{(\kappa)} \text{ and } \mathcal{G}^{(\kappa)} \text{ hold}; \\ \tau=\kappa+\operatorname{cont} T}}}_{\substack{= \sum_{\substack{T \text{ is a column-strict tableau} \\ \text{of shape } \lambda/\mu; \\ \tau=\kappa+\operatorname{cont} T; \\ \text{the five equivalent assertions} \\ \mathcal{C}^{(\kappa)},\ \mathcal{D}^{(\kappa)},\ \mathcal{E}^{(\kappa)},\ \mathcal{F}^{(\kappa)} \text{ and } \mathcal{G}^{(\kappa)} \text{ hold}}} = \sum_{\substack{T \text{ is a column-strict tableau} \\ \text{of shape } \lambda/\mu; \\ \tau=\kappa+\operatorname{cont} T; \\ \text{the six equivalent assertions} \\ \mathcal{C}^{(\kappa)},\ \mathcal{D}^{(\kappa)},\ \mathcal{E}^{(\kappa)},\ \mathcal{F}^{(\kappa)},\ \mathcal{G}^{(\kappa)} \text{ and } \mathcal{H}^{(\kappa)} \text{ hold}}} \\ \text{(by the equivalence (13.96.6))}}} 1 \\
&= \sum_{\substack{T \text{ is a column-strict tableau} \\ \text{of shape } \lambda/\mu; \\ \tau=\kappa+\operatorname{cont} T; \\ \text{the six equivalent assertions} \\ \mathcal{C}^{(\kappa)},\ \mathcal{D}^{(\kappa)},\ \mathcal{E}^{(\kappa)},\ \mathcal{F}^{(\kappa)},\ \mathcal{G}^{(\kappa)} \text{ and } \mathcal{H}^{(\kappa)} \text{ hold}}} 1 \\
&= \Big(\text{the number of all column-strict tableaux } T \text{ of shape } \lambda/\mu \text{ such that} \\
&\qquad \tau = \kappa + \operatorname{cont} T \text{ and such that the six equivalent} \\
&\qquad \text{assertions } \mathcal{C}^{(\kappa)},\ \mathcal{D}^{(\kappa)},\ \mathcal{E}^{(\kappa)},\ \mathcal{F}^{(\kappa)},\ \mathcal{G}^{(\kappa)} \text{ and } \mathcal{H}^{(\kappa)} \text{ hold}\Big) \\
&= \Big(\text{the number of all column-strict tableaux } T \text{ of shape } \lambda/\mu \text{ satisfying} \\
&\qquad \tau = \kappa + \operatorname{cont} T \text{ and also satisfying the six equivalent} \\
&\qquad \text{assertions } \mathcal{C}^{(\kappa)},\ \mathcal{D}^{(\kappa)},\ \mathcal{E}^{(\kappa)},\ \mathcal{F}^{(\kappa)},\ \mathcal{G}^{(\kappa)} \text{ and } \mathcal{H}^{(\kappa)}\Big).
\end{aligned}$$

This solves Exercise 2.9.21(b).

13.97. **Solution to Exercise 2.9.22.** *Solution to Exercise 2.9.22.* (a) We shall first prove that

$$\text{if } N \text{ has Jordan type } \lambda\text{, then every } k \in \mathbb{N} \text{ satisfies } \dim\left(\ker\left(N^k\right)\right) = \left(\lambda^t\right)_1 + \left(\lambda^t\right)_2 + \cdots + \left(\lambda^t\right)_k . \tag{13.97.1}$$

Proof of (13.97.1): Assume that N has Jordan type λ. Let $k \in \mathbb{N}$.

For each $m \in \mathbb{N}$, let J_m be $m \times m$-matrix $\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix} \in \mathbb{K}^{m \times m}$. This is a Jordan block of size m corresponding to the eigenvalue 0 (when m is positive). For future use, we record the following simple fact:

$$\dim\left(\ker\left(J_m^k\right)\right) = \min\{m, k\} \tag{13.97.2}$$

for all $m \in \mathbb{N}$. [755].

Since N has Jordan type λ, the Jordan normal form of N has Jordan blocks of sizes $\lambda_1, \lambda_2, \lambda_3, \ldots$. Since the only eigenvalue of N is 0, this shows that the Jordan blocks of N are $J_{\lambda_1}, J_{\lambda_2}, J_{\lambda_3}, \ldots$ (or, more precisely, the nonempty matrices among $J_{\lambda_1}, J_{\lambda_2}, J_{\lambda_3}, \ldots$). In other words, N is similar to the block-diagonal matrix $J_\lambda := \begin{pmatrix} J_{\lambda_1} & 0 & 0 & \cdots \\ 0 & J_{\lambda_2} & 0 & \cdots \\ 0 & 0 & J_{\lambda_3} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$. (This block-diagonal matrix J_λ is finite, since only finitely many λ_p are nonzero.) We can thus WLOG assume that N **is** this matrix J_λ (because replacing N by a matrix similar to N changes neither the dimension $\dim\left(\ker\left(N^k\right)\right)$ nor the Jordan type of N). Assume this. Thus, $N = J_\lambda = \begin{pmatrix} J_{\lambda_1} & 0 & 0 & \cdots \\ 0 & J_{\lambda_2} & 0 & \cdots \\ 0 & 0 & J_{\lambda_3} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$, so that

$$N^k = \begin{pmatrix} J_{\lambda_1} & 0 & 0 & \cdots \\ 0 & J_{\lambda_2} & 0 & \cdots \\ 0 & 0 & J_{\lambda_3} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}^k = \begin{pmatrix} J_{\lambda_1}^k & 0 & 0 & \cdots \\ 0 & J_{\lambda_2}^k & 0 & \cdots \\ 0 & 0 & J_{\lambda_3}^k & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

and therefore

$$\ker\left(N^k\right) = \ker \begin{pmatrix} J_{\lambda_1}^k & 0 & 0 & \cdots \\ 0 & J_{\lambda_2}^k & 0 & \cdots \\ 0 & 0 & J_{\lambda_3}^k & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \cong \bigoplus_{p \geq 1} \ker\left(J_{\lambda_p}^k\right),$$

whence

$$\dim\left(\ker\left(N^k\right)\right) = \dim\left(\bigoplus_{p\geq 1} \ker\left(J_{\lambda_p}^k\right)\right) = \sum_{p \geq 1} \underbrace{\dim\left(\ker\left(J_{\lambda_p}^k\right)\right)}_{\substack{=\min\{\lambda_p, k\} \\ \text{(by (13.97.2) (applied to } m=\lambda_p))}} = \sum_{p\geq 1} \min\{\lambda_p, k\}. \tag{13.97.3}$$

[755] *Proof sketch.* Let $m \in \mathbb{N}$. Let $(e_1, e_2, \ldots, e_m)$ be the standard basis of the $\mathbb{K}$-vector space $\mathbb{K}^m$. Now, it is easy to check that J_m^k is the $m \times m$-matrix whose (i,j)-th entry is $\begin{cases} 1, & \text{if } j = i + k; \\ 0, & \text{if } j \neq i + k \end{cases}$ for all $(i,j) \in \{1, 2, \ldots, m\}^2$. Hence, it is easy to see that $\ker\left(J_m^k\right)$ is the $\mathbb{K}$-linear span of the basis vectors e_j with $j \leq k$. These basis vectors are linearly independent and their number is $\min\{m, k\}$; therefore, (13.97.2) follows.

We are now going to prove that the right hand side of this equality is $(\lambda^t)_1 + (\lambda^t)_2 + \cdots + (\lambda^t)_k$. Indeed, let us use the *Iverson bracket notation*: For every assertion $\mathcal{A}$, we let $[\mathcal{A}]$ denote the integer $\begin{cases} 1, & \text{if } \mathcal{A} \text{ is true;} \\ 0, & \text{if } \mathcal{A} \text{ is false} \end{cases}$. (This integer is called the *truth value* of $\mathcal{A}$.)

Any two nonnegative integers u and v satisfy $\min\{u, v\} = \sum_{i=1}^{v} [i \leq u]$. Thus, $\min\{\lambda_p, k\} = \sum_{i=1}^{k} [i \leq \lambda_p]$ for every $p \in \{1, 2, 3, \ldots\}$. Hence, (13.97.3) becomes

$$\dim\left(\ker\left(N^k\right)\right) = \sum_{p\geq 1} \underbrace{\min\{\lambda_p, k\}}_{=\sum_{i=1}^{k}[i\leq\lambda_p]} = \sum_{p\geq 1}\sum_{i=1}^{k} [i \leq \lambda_p] = \sum_{i=1}^{k} \underbrace{\sum_{p\geq 1} [i \leq \lambda_p]}_{\substack{=|\{p\geq 1 \mid i\leq\lambda_p\}| \\ =|\{p\geq 1 \mid \lambda_p\geq i\}| \\ =|\{j\geq 1 \mid \lambda_j\geq i\}|=\left(\lambda^t\right)_i \\ (\text{by (2.2.7)})}}$$

$$= \sum_{i=1}^{k} \left(\lambda^t\right)_i = \left(\lambda^t\right)_1 + \left(\lambda^t\right)_2 + \cdots + \left(\lambda^t\right)_k .$$

This proves (13.97.1).

With (13.97.1), we have proven one direction of the equivalence that Exercise 2.9.22(a) requires us to prove. To prove the other direction, we need to show that

$$\text{if every } k \in \mathbb{N} \text{ satisfies } \dim\left(\ker\left(N^k\right)\right) = \left(\lambda^t\right)_1 + \left(\lambda^t\right)_2 + \cdots + \left(\lambda^t\right)_k, \text{ then } N \text{ has Jordan type } \lambda. \tag{13.97.4}$$

Proof of (13.97.4): Assume that every $k \in \mathbb{N}$ satisfies $\dim\left(\ker\left(N^k\right)\right) = (\lambda^t)_1 + (\lambda^t)_2 + \cdots + (\lambda^t)_k$. Let μ be the Jordan type of N. Then, (13.97.1) (applied to μ instead of λ) shows that every $k \in \mathbb{N}$ satisfies $\dim\left(\ker\left(N^k\right)\right) = (\mu^t)_1 + (\mu^t)_2 + \cdots + (\mu^t)_k$. Hence, every $k \in \mathbb{N}$ satisfies

$$\left(\mu^t\right)_1 + \left(\mu^t\right)_2 + \cdots + \left(\mu^t\right)_k = \dim\left(\ker\left(N^k\right)\right) = \left(\lambda^t\right)_1 + \left(\lambda^t\right)_2 + \cdots + \left(\lambda^t\right)_k . \tag{13.97.5}$$

Applying (13.97.5) to $k-1$ instead of k, and subtracting the result from (13.97.5), we obtain

$$\left(\mu^t\right)_k = \left(\lambda^t\right)_k \qquad \text{for every positive integer } k.$$

Hence, $\mu^t = \lambda^t$, so that $\mu = \lambda$. Thus, N has Jordan type $\mu = \lambda$. This proves (13.97.4), and thus the solution of Exercise 2.9.22(a) is complete.

Before we come to the solution of Exercise 2.9.22(b), let us show a linear-algebraic lemma:

Lemma 13.97.1. *Let W be a finite-dimensional $\mathbb{K}$-vector space. Let A and B be $\mathbb{K}$-vector subspaces of W. Let $f \in \operatorname{End} W$ be such that $f(A) \subset A$ and $f(B) \subset B$. For any $(i, j) \in \mathbb{N}^2$, let us define a nonnegative integer $w_{i,j}$ by $w_{i,j} = \dim\left(\left(f^i\right)^{-1}(A) \cap \left(f^j\right)^{-1}(B)\right)$.*

(a) For any $(i, j) \in \{1, 2, 3, \ldots\}^2$, we have $w_{i,j} + w_{i-1,j-1} \geq w_{i,j-1} + w_{i-1,j}$.

(b) For any $(i, j) \in \mathbb{N} \times \{1, 2, 3, \ldots\}$, we have $w_{i+1,j+1} - w_{i+1,j} \leq w_{i,j} - w_{i,j-1}$.

(c) For any $(i, j) \in \{1, 2, 3, \ldots\} \times \mathbb{N}$, we have $w_{i+1,j+1} - w_{i,j+1} \leq w_{i,j} - w_{i-1,j}$.

(d) If $i \in \{1, 2, 3, \ldots\}$ and $j \in \mathbb{N}$ are such that $i > \dim W$, then $w_{i,j} = w_{i-1,j}$.

(e) If $j \in \{1, 2, 3, \ldots\}$ and $i \in \mathbb{N}$ are such that $j > \dim W$, then $w_{i,j} = w_{i,j-1}$.

(f) Assume that f is nilpotent. For every $j \in \mathbb{N}$, we have $\sum_{i=1}^{\infty} (w_{i,j} - w_{i-1,j}) = \dim\left(\left(f^j\right)^{-1}(B)\right) - \dim\left(A \cap \left(f^j\right)^{-1}(B)\right)$. (In particular, the sum $\sum_{i=1}^{\infty} (w_{i,j} - w_{i-1,j})$ converges with respect to the discrete topology, i.e., all but finitely many of its terms are 0.)

(g) Assume that f is nilpotent. For every $i \in \mathbb{N}$, we have $\sum_{j=1}^{\infty} (w_{i,j} - w_{i,j-1}) = \dim\left(\left(f^i\right)^{-1}(A)\right) - \dim\left(\left(f^i\right)^{-1}(A) \cap B\right)$. (In particular, the sum $\sum_{j=1}^{\infty} (w_{i,j} - w_{i,j-1})$ converges with respect to the discrete topology, i.e., all but finitely many of its terms are 0.)

Proof of Lemma 13.97.1. Notice first that $\Big(\underbrace{f^0}_{=\mathrm{id}}\Big)^{-1}(A) = \mathrm{id}^{-1}(A) = A$ and $A = \left(f^0\right)^{-1}(A) \subset \left(f^1\right)^{-1}(A) \subset \left(f^2\right)^{-1}(A) \subset \cdots$ (since $f(A) \subset A$) and $B = \left(f^0\right)^{-1}(B) \subset \left(f^1\right)^{-1}(B) \subset \left(f^2\right)^{-1}(B) \subset \cdots$ (similarly).

(a) Let $(i,j) \in \{1,2,3,\ldots\}^2$. We need to prove that $w_{i,j} + w_{i-1,j-1} \geq w_{i,j-1} + w_{i-1,j}$. In other words, we need to prove that $w_{i,j} - w_{i,j-1} \geq w_{i-1,j} - w_{i-1,j-1}$. Since

$$\begin{aligned} w_{i,j} - w_{i,j-1} &= \dim\left(\left(f^i\right)^{-1}(A) \cap \left(f^j\right)^{-1}(B)\right) - \dim\left(\left(f^i\right)^{-1}(A) \cap \left(f^{j-1}\right)^{-1}(B)\right) \\ &\qquad \left(\text{by the definitions of } w_{i,j} \text{ and } w_{i,j-1}\right) \\ &= \dim\left(\left(\left(f^i\right)^{-1}(A) \cap \left(f^j\right)^{-1}(B)\right) / \left(\left(f^i\right)^{-1}(A) \cap \left(f^{j-1}\right)^{-1}(B)\right)\right) \\ &\qquad \left(\text{since } \left(f^i\right)^{-1}(A) \cap \left(f^{j-1}\right)^{-1}(B) \subset \left(f^i\right)^{-1}(A) \cap \left(f^j\right)^{-1}(B)\right) \end{aligned} \tag{13.97.6}$$

and

$$w_{i-1,j} - w_{i-1,j-1} = \dim\left(\left(\left(f^{i-1}\right)^{-1}(A) \cap \left(f^j\right)^{-1}(B)\right) / \left(\left(f^{i-1}\right)^{-1}(A) \cap \left(f^{j-1}\right)^{-1}(B)\right)\right)$$

(by the same argument as (13.97.6), only with $i-1$ instead of i), this is equivalent to showing that

$$\begin{aligned} &\dim\left(\left(\left(f^i\right)^{-1}(A) \cap \left(f^j\right)^{-1}(B)\right) / \left(\left(f^i\right)^{-1}(A) \cap \left(f^{j-1}\right)^{-1}(B)\right)\right) \\ &\geq \dim\left(\left(\left(f^{i-1}\right)^{-1}(A) \cap \left(f^j\right)^{-1}(B)\right) / \left(\left(f^{i-1}\right)^{-1}(A) \cap \left(f^{j-1}\right)^{-1}(B)\right)\right). \end{aligned}$$

This will clearly be achieved if we can construct a $\mathbb{K}$-linear injection

$$\begin{aligned} &\left(\left(f^{i-1}\right)^{-1}(A) \cap \left(f^j\right)^{-1}(B)\right) / \left(\left(f^{i-1}\right)^{-1}(A) \cap \left(f^{j-1}\right)^{-1}(B)\right) \\ &\to \left(\left(f^i\right)^{-1}(A) \cap \left(f^j\right)^{-1}(B)\right) / \left(\left(f^i\right)^{-1}(A) \cap \left(f^{j-1}\right)^{-1}(B)\right). \end{aligned}$$

Here is how to construct it: Let ι denote the canonical inclusion $\left(f^{i-1}\right)^{-1}(A) \cap \left(f^j\right)^{-1}(B) \to \left(f^i\right)^{-1}(A) \cap \left(f^j\right)^{-1}(B)$. Then, ι restricts to an inclusion $\left(f^{i-1}\right)^{-1}(A) \cap \left(f^{j-1}\right)^{-1}(B) \to \left(f^i\right)^{-1}(A) \cap \left(f^{j-1}\right)^{-1}(B)$, and so gives rise to a map

$$\begin{aligned} &\left(\left(f^{i-1}\right)^{-1}(A) \cap \left(f^j\right)^{-1}(B)\right) / \left(\left(f^{i-1}\right)^{-1}(A) \cap \left(f^{j-1}\right)^{-1}(B)\right) \\ &\to \left(\left(f^i\right)^{-1}(A) \cap \left(f^j\right)^{-1}(B)\right) / \left(\left(f^i\right)^{-1}(A) \cap \left(f^{j-1}\right)^{-1}(B)\right). \end{aligned}$$

This map is injective because $\iota^{-1}\left(\left(f^i\right)^{-1}(A) \cap \left(f^{j-1}\right)^{-1}(B)\right) = \left(f^{i-1}\right)^{-1}(A) \cap \left(f^{j-1}\right)^{-1}(B)$, and thus we have found our $\mathbb{K}$-linear injection. Lemma 13.97.1(a) is proven.

(b) Let $(i,j) \in \mathbb{N} \times \{1,2,3,\ldots\}$. We need to prove that $w_{i+1,j+1} - w_{i+1,j} \leq w_{i,j} - w_{i,j-1}$. Due to (13.97.6) and due to

$$w_{i+1,j+1} - w_{i+1,j} = \dim\left(\left(\left(f^{i+1}\right)^{-1}(A) \cap \left(f^{j+1}\right)^{-1}(B)\right) / \left(\left(f^{i+1}\right)^{-1}(A) \cap \left(f^j\right)^{-1}(B)\right)\right)$$

(this follows by the same arguments as (13.97.6), only with i and j replaced by $i+1$ and $j+1$), this is equivalent to showing that

$$\begin{aligned} &\dim\left(\left(\left(f^{i+1}\right)^{-1}(A) \cap \left(f^{j+1}\right)^{-1}(B)\right) / \left(\left(f^{i+1}\right)^{-1}(A) \cap \left(f^j\right)^{-1}(B)\right)\right) \\ &\leq \dim\left(\left(\left(f^i\right)^{-1}(A) \cap \left(f^j\right)^{-1}(B)\right) / \left(\left(f^i\right)^{-1}(A) \cap \left(f^{j-1}\right)^{-1}(B)\right)\right). \end{aligned}$$

This will clearly be achieved if we can construct a $\mathbb{K}$-linear injection

$$\begin{aligned} &\left(\left(f^{i+1}\right)^{-1}(A) \cap \left(f^{j+1}\right)^{-1}(B)\right) / \left(\left(f^{i+1}\right)^{-1}(A) \cap \left(f^j\right)^{-1}(B)\right) \\ &\to \left(\left(f^i\right)^{-1}(A) \cap \left(f^j\right)^{-1}(B)\right) / \left(\left(f^i\right)^{-1}(A) \cap \left(f^{j-1}\right)^{-1}(B)\right). \end{aligned}$$

Here is how this can be done: The map f restricts to a map $\varphi : \left(f^{i+1}\right)^{-1}(A) \cap \left(f^{j+1}\right)^{-1}(B) \to \left(f^i\right)^{-1}(A) \cap \left(f^j\right)^{-1}(B)$, which further restricts to a map $\left(f^{i+1}\right)^{-1}(A) \cap \left(f^j\right)^{-1}(B) \to \left(f^i\right)^{-1}(A) \cap \left(f^{j-1}\right)^{-1}(B)$. Hence, φ gives rise to a map

$$\begin{aligned} &\left(\left(f^{i+1}\right)^{-1}(A) \cap \left(f^{j+1}\right)^{-1}(B)\right) / \left(\left(f^{i+1}\right)^{-1}(A) \cap \left(f^j\right)^{-1}(B)\right) \\ &\to \left(\left(f^i\right)^{-1}(A) \cap \left(f^j\right)^{-1}(B)\right) / \left(\left(f^i\right)^{-1}(A) \cap \left(f^{j-1}\right)^{-1}(B)\right), \end{aligned}$$

which is injective because

$$\begin{aligned}\varphi^{-1}\left(\left(f^{i}\right)^{-1}(A)\cap\left(f^{j-1}\right)^{-1}(B)\right) &= f^{-1}\left(\left(f^{i}\right)^{-1}(A)\cap\left(f^{j-1}\right)^{-1}(B)\right)\\ &= \underbrace{f^{-1}\left(\left(f^{i}\right)^{-1}(A)\right)}_{=\left(f^{i+1}\right)^{-1}(A)}\cap\underbrace{f^{-1}\left(\left(f^{j-1}\right)^{-1}(B)\right)}_{=\left(f^{j}\right)^{-1}(B)} = \left(f^{i+1}\right)^{-1}(A)\cap\left(f^{j}\right)^{-1}(B).\end{aligned}$$

Thus, we have found our $\mathbb{K}$-linear injection. Lemma 13.97.1(b) is proven.

(c) The proof of Lemma 13.97.1(c) is analogous to our above proof of Lemma 13.97.1(b) (one merely has to interchange A with B and i with j).

(d) We have

$$\left(f^{0}\right)^{-1}(A)\subset\left(f^{1}\right)^{-1}(A)\subset\cdots\subset\left(f^{\dim W+1}\right)^{-1}(A)$$

and therefore

$$\dim\left(\left(f^{0}\right)^{-1}(A)\right)\leq\dim\left(\left(f^{1}\right)^{-1}(A)\right)\leq\cdots\leq\dim\left(\left(f^{\dim W+1}\right)^{-1}(A)\right).$$

This latter chain of inequalities contains $\dim W+1$ inequality signs, but only at most $\dim W$ of them can be strict (because each $\dim\left(\left(f^{k}\right)^{-1}(A)\right)$ is an integer between 0 and $\dim W$ inclusive[756], and a sequence of $\dim W+2$ integers between 0 and $\dim W$ cannot strictly increase). Thus, at least one of the inequality signs is an equality. That is, there exists an $I\in\{1,2,\ldots,\dim W+1\}$ such that $\dim\left(\left(f^{I-1}\right)^{-1}(A)\right)=\dim\left(\left(f^{I}\right)^{-1}(A)\right)$. Consider this I.

Since $\dim\left(\left(f^{I-1}\right)^{-1}(A)\right)=\dim\left(\left(f^{I}\right)^{-1}(A)\right)$ and $\left(f^{I-1}\right)^{-1}(A)\subset\left(f^{I}\right)^{-1}(A)$, we must have

$$\left(f^{I-1}\right)^{-1}(A)=\left(f^{I}\right)^{-1}(A). \qquad (13.97.7)$$

From here, it is easy to see that

$$\left(f^{i-1}\right)^{-1}(A)=\left(f^{i}\right)^{-1}(A)\qquad\text{for every } i\in\{1,2,3,\ldots\} \text{ satisfying } i>\dim W.$$

[757]

Thus, for every $i\in\{1,2,3,\ldots\}$ satisfying $i>\dim W$, we have

$$\begin{aligned}w_{i-1,j} &= \dim\left(\underbrace{\left(f^{i-1}\right)^{-1}(A)}_{=\left(f^{i}\right)^{-1}(A)}\cap\left(f^{j}\right)^{-1}(B)\right)\qquad\text{(by the definition of } w_{i-1,j}\text{)}\\ &= \dim\left(\left(f^{i}\right)^{-1}(A)\cap\left(f^{j}\right)^{-1}(B)\right)=w_{i,j}\qquad\text{for all } j\in\mathbb{N}.\end{aligned}$$

Thus, if $i\in\{1,2,3,\ldots\}$ and $j\in\mathbb{N}$ are such that $i>\dim W$, then $w_{i-1,j}=w_{i-1,j}$. This proves Lemma 13.97.1(d).

(e) The proof of Lemma 13.97.1(e) is analogous to our above proof of Lemma 13.97.1(d) (one merely has to interchange A with B and i with j).

[756] since $\left(f^{k}\right)^{-1}(A)$ is a subspace of the finite-dimensional $\mathbb{K}$-vector space W

[757] *Proof.* Let $i\in\{1,2,3,\ldots\}$ be such that $i>\dim W$. Thus, $i\geq\dim W+1$. But $I\leq\dim W+1$ (since $I\in\{1,2,\ldots,\dim W+1\}$) and thus $i\geq\dim W+1\geq I$. Hence, there exists some $k\in\mathbb{N}$ such that $i=I+k$. Consider this k. Since $i=I+k$, we have $i-1=I+k-1=(I-1)+k$, so that $f^{i-1}=f^{(I-1)+k}=f^{I-1}\circ f^{k}$ and thus $\left(f^{i-1}\right)^{-1}(A)=\left(f^{I-1}\circ f^{k}\right)^{-1}(A)=\left(f^{k}\right)^{-1}\left(\left(f^{I-1}\right)^{-1}(A)\right)$. Similarly, $\left(f^{i}\right)^{-1}(A)=\left(f^{k}\right)^{-1}\left(\left(f^{I}\right)^{-1}(A)\right)$. Now,

$$\left(f^{i-1}\right)^{-1}(A)=\left(f^{k}\right)^{-1}\left(\underbrace{\left(f^{I-1}\right)^{-1}(A)}_{\substack{=\left(f^{I}\right)^{-1}(A)\\ \text{(by (13.97.7))}}}\right)=\left(f^{k}\right)^{-1}\left(\left(f^{I}\right)^{-1}(A)\right)=\left(f^{i}\right)^{-1}(A),$$

qed.

(f) Let $j \in \mathbb{N}$. We have

$$\sum_{i=1}^{\infty} (w_{i,j} - w_{i-1,j}) = \sum_{i=1}^{\dim W} (w_{i,j} - w_{i-1,j}) + \sum_{i=\dim W+1}^{\infty} \left(\underbrace{w_{i,j}}_{\substack{=w_{i-1,j} \\ \text{(by Lemma 13.97.1(d),} \\ \text{since } i>\dim W)}} - w_{i-1,j} \right)$$

$$= \sum_{i=1}^{\dim W} (w_{i,j} - w_{i-1,j}) + \sum_{i=\dim W+1}^{\infty} \underbrace{(w_{i-1,j} - w_{i-1,j})}_{=0} = \sum_{i=1}^{\dim W} (w_{i,j} - w_{i-1,j}) + \underbrace{\sum_{i=\dim W+1}^{\infty} 0}_{=0}$$

$$= \sum_{i=1}^{\dim W} (w_{i,j} - w_{i-1,j}) = w_{\dim W,j} - w_{0,j} \qquad \text{(by the telescope principle)}. \tag{13.97.8}$$

Now, we are going to prove that $w_{\dim W,j} = \dim \left(\left(f^j\right)^{-1}(B) \right)$ and $w_{0,j} = \dim \left(A \cap \left(f^j\right)^{-1}(B) \right)$.

Indeed, it is well-known that if g is any nilpotent endomorphism of a finite-dimensional vector space V, then $g^{\dim V} = 0$. Applied to $g = f$ and $V = W$, this yields $f^{\dim W} = 0$, and thus $f^{\dim W}(W) = 0(W) = 0 \subset A$, whence $W \subset \left(f^{\dim W}\right)^{-1}(A)$, so that $\left(f^{\dim W}\right)^{-1}(A) = W$. Hence,

$$\underbrace{\left(f^{\dim W}\right)^{-1}(A)}_{=W} \cap \left(f^j\right)^{-1}(B) = W \cap \left(f^j\right)^{-1}(B) = \left(f^j\right)^{-1}(B)$$

(since $\left(f^j\right)^{-1}(B) \subset W$). Now, the definition of $w_{\dim W,j}$ yields

$$w_{\dim W,j} = \dim \left(\underbrace{\left(f^{\dim W}\right)^{-1}(A) \cap \left(f^j\right)^{-1}(B)}_{=(f^j)^{-1}(B)} \right) = \dim \left(\left(f^j\right)^{-1}(B) \right).$$

Also, the definition of $w_{0,j}$ yields

$$w_{0,j} = \dim \left(\left(\underbrace{f^0}_{=\mathrm{id}} \right)^{-1}(A) \cap \left(f^j\right)^{-1}(B) \right) = \dim \left(\underbrace{\mathrm{id}^{-1}(A)}_{=A} \cap \left(f^j\right)^{-1}(B) \right) = \dim \left(A \cap \left(f^j\right)^{-1}(B) \right).$$

Hence, (13.97.8) becomes

$$\sum_{i=1}^{\infty} (w_{i,j} - w_{i-1,j}) = \underbrace{w_{\dim W,j}}_{=\dim\left((f^j)^{-1}(B)\right)} - \underbrace{w_{0,j}}_{=\dim\left(A\cap(f^j)^{-1}(B)\right)} = \dim \left(\left(f^j\right)^{-1}(B) \right) - \dim \left(A \cap \left(f^j\right)^{-1}(B) \right).$$

This proves Lemma 13.97.1(f).

(g) The proof of Lemma 13.97.1(g) is analogous to our above proof of Lemma 13.97.1(f) (one merely has to interchange A with B and i with j). □

(b) Assume that $\mathbb{Z}$ is a subring of $\mathbf{k}$.

For any $(i,j) \in \mathbb{N}^2$, let us define a nonnegative integer $a_{i,j}$ by

$$a_{i,j} = \dim \left(\left(f^i\right)^{-1}(U) \cap \ker\left(f^j\right) \right).$$

Furthermore, for every $(i,j) \in \{1,2,3,\ldots\}^2$, let us define an integer $b_{i,j}$ by

$$b_{i,j} = a_{i,j} - a_{i,j-1} - a_{i-1,j} + a_{i-1,j-1}.$$

We first observe that

$$a_{i,0} = 0 \qquad \text{for every } i \in \mathbb{N}. \tag{13.97.9}$$

[758] Furthermore,

$$a_{0,j} = \left(\mu^t\right)_1 + \left(\mu^t\right)_2 + \cdots + \left(\mu^t\right)_j \qquad \text{for every } j \in \mathbb{N}. \tag{13.97.10}$$

[759] Thus,

$$a_{0,j} - a_{0,j-1} = \left(\mu^t\right)_j \qquad \text{for every } j \in \{1,2,3,\ldots\}. \tag{13.97.11}$$

[760] Moreover, every $(i,j) \in \mathbb{N}^2$ satisfies

$$a_{i,j} = \dim\left(\left(f^i\right)^{-1}(U) \cap \underbrace{\ker\left(f^j\right)}_{=\left(f^j\right)^{-1}(0)} \right) = \dim\left(\left(f^i\right)^{-1}(U) \cap \left(f^j\right)^{-1}(0) \right).$$

Hence, Lemma 13.97.1(a) (applied to $W = V$, $A = U$, $B = 0$ and $w_{i,j} = a_{i,j}$) shows that for any $(i,j) \in \{1,2,3,\ldots\}^2$, we have $a_{i,j} + a_{i-1,j-1} \geq a_{i,j-1} + a_{i-1,j}$. In other words, for any $(i,j) \in \{1,2,3,\ldots\}^2$, we have $a_{i,j} - a_{i,j-1} - a_{i-1,j} + a_{i-1,j-1} \geq 0$. In other words, for any $(i,j) \in \{1,2,3,\ldots\}^2$, we have

$$b_{i,j} \geq 0 \tag{13.97.12}$$

(since $b_{i,j} = a_{i,j} - a_{i,j-1} - a_{i-1,j} + a_{i-1,j-1}$). Thus, $b_{i,j}$ is a nonnegative integer for every two positive integers i and j. Moreover,

$$b_{i,j} = 0 \qquad \text{for all but finitely many pairs } (i,j). \tag{13.97.13}$$

[758] *Proof of (13.97.9):* For every $i \in \mathbb{N}$, the definition of $a_{i,0}$ yields $a_{i,0} = \dim\left(\left(f^i\right)^{-1}(U) \cap \underbrace{\ker\left(f^0\right)}_{=\ker(\mathrm{id})=0} \right) = \dim \underbrace{\left(\left(f^i\right)^{-1}(U) \cap 0 \right)}_{=0} = \dim 0 = 0$, qed.

[759] *Proof of (13.97.10):* Let $j \in \mathbb{N}$. We can represent the endomorphism $f \mid U$ of U as a $k \times k$-matrix $G \in \mathbb{K}^{k \times k}$ for $k = \dim U$. This $k \times k$-matrix G is nilpotent (since $f \mid U$ is nilpotent) and has Jordan type μ (since $f \mid U$ has Jordan type μ). But Exercise 2.9.22(a) (applied to k, G and μ instead of n, N and λ) shows that G has Jordan type μ if and only if every $k \in \mathbb{N}$ satisfies $\dim\left(\ker\left(G^k\right)\right) = \left(\mu^t\right)_1 + \left(\mu^t\right)_2 + \cdots + \left(\mu^t\right)_k$. Since we know that G has Jordan type μ, we thus conclude that every $k \in \mathbb{N}$ satisfies $\dim\left(\ker\left(G^k\right)\right) = \left(\mu^t\right)_1 + \left(\mu^t\right)_2 + \cdots + \left(\mu^t\right)_k$. Applied to $k = j$, this yields $\dim\left(\ker\left(G^j\right)\right) = \left(\mu^t\right)_1 + \left(\mu^t\right)_2 + \cdots + \left(\mu^t\right)_j$.

Since the matrix G represents the endomorphism $f \mid U$, we have

$$\dim\left(\ker\left(G^j\right)\right) = \dim\left(\underbrace{\ker\left((f \mid U)^j \right)}_{=U \cap \ker\left(f^j\right)} \right) = \dim\left(U \cap \ker\left(f^j\right)\right).$$

Compared with $\dim\left(\ker\left(G^j\right)\right) = \left(\mu^t\right)_1 + \left(\mu^t\right)_2 + \cdots + \left(\mu^t\right)_j$, this yields

$$\dim\left(U \cap \ker\left(f^j\right)\right) = \left(\mu^t\right)_1 + \left(\mu^t\right)_2 + \cdots + \left(\mu^t\right)_j.$$

Now, the definition of $a_{0,j}$ yields

$$a_{0,j} = \dim\left(\left(\underbrace{f^0}_{=\mathrm{id}} \right)^{-1} (U) \cap \ker\left(f^j\right) \right) = \dim\left(\underbrace{\mathrm{id}^{-1}(U)}_{=U} \cap \ker\left(f^j\right) \right) = \dim\left(U \cap \ker\left(f^j\right)\right) = \left(\mu^t\right)_1 + \left(\mu^t\right)_2 + \cdots + \left(\mu^t\right)_j.$$

This proves (13.97.10).

[760] *Proof of (13.97.11):* For every $j \in \{1,2,3,\ldots\}$, we have

$$\underbrace{a_{0,j}}_{\substack{=\left(\mu^t\right)_1+\left(\mu^t\right)_2+\cdots+\left(\mu^t\right)_j \\ \text{(by (13.97.10))}}} - \underbrace{a_{0,j-1}}_{\substack{=\left(\mu^t\right)_1+\left(\mu^t\right)_2+\cdots+\left(\mu^t\right)_{j-1} \\ \text{(by (13.97.10), applied to} \\ j-1 \text{ instead of } j)}} = \left(\left(\mu^t\right)_1 + \left(\mu^t\right)_2 + \cdots + \left(\mu^t\right)_j \right) - \left(\left(\mu^t\right)_1 + \left(\mu^t\right)_2 + \cdots + \left(\mu^t\right)_{j-1} \right) = \left(\mu^t\right)_j,$$

qed.

[761] Also,

$$\left(\mu^t\right)_j + \left(b_{1,j} + b_{2,j} + \cdots + b_{i,j}\right) = a_{i,j} - a_{i,j-1} \qquad \text{for every } (i,j) \in \mathbb{N} \times \{1,2,3,\ldots\}. \tag{13.97.15}$$

[762]

Now, Exercise 2.9.18(a) (applied to μ^t instead of μ) yields that the following two assertions are equivalent[763]:

- *Assertion $\mathcal{A}'$:* There exist a partition γ and a column-strict tableau T of shape γ/μ^t such that all $(i,j) \in \{1,2,3,\ldots\}^2$ satisfy

$$b_{i,j} = (\text{the number of all entries } i \text{ in the } j\text{-th row of } T). \tag{13.97.16}$$

- *Assertion $\mathcal{B}'$:* The inequality

$$\left(\mu^t\right)_{j+1} + \left(b_{1,j+1} + b_{2,j+1} + \cdots + b_{i+1,j+1}\right) \leq \left(\mu^t\right)_j + \left(b_{1,j} + b_{2,j} + \cdots + b_{i,j}\right) \tag{13.97.17}$$

  holds for all $(i,j) \in \mathbb{N} \times \{1,2,3,\ldots\}$.

We shall now prove that Assertion $\mathcal{B}'$ holds. Indeed, any $(i,j) \in \mathbb{N} \times \{1,2,3,\ldots\}$ satisfies

$$\begin{aligned}
&\left(\mu^t\right)_{j+1} + \left(b_{1,j+1} + b_{2,j+1} + \cdots + b_{i+1,j+1}\right) \\
&= a_{i+1,j+1} - a_{i+1,j} \qquad (\text{by (13.97.15), applied to } i+1 \text{ and } j+1 \text{ instead of } i \text{ and } j) \\
&\leq a_{i,j} - a_{i,j-1} \qquad (\text{by Lemma 13.97.1(b) (applied to } W = V,\ A = U,\ B = 0 \text{ and } w_{i,j} = a_{i,j})) \\
&= \left(\mu^t\right)_j + \left(b_{1,j} + b_{2,j} + \cdots + b_{i,j}\right) \qquad (\text{by (13.97.15)}).
\end{aligned}$$

Thus, Assertion $\mathcal{B}'$ holds. Since Assertions $\mathcal{A}'$ and $\mathcal{B}'$ are equivalent, this yields that Assertion $\mathcal{A}'$ also holds. In other words, there exist a partition γ and a column-strict tableau T of shape γ/μ^t such that all $(i,j) \in \{1,2,3,\ldots\}^2$ satisfy (13.97.16). Let us consider this γ and this T.

We shall soon see that $\gamma = \lambda^t$ and $\operatorname{cont} T = \nu^t$. Let us first prepare for this. We have

$$\dim\left(\left(f^j\right)^{-1}(0)\right) = \left(\lambda^t\right)_1 + \left(\lambda^t\right)_2 + \cdots + \left(\lambda^t\right)_j \qquad \text{for every } j \in \mathbb{N}. \tag{13.97.18}$$

[761] *Proof of (13.97.13):* Lemma 13.97.1(d) (applied to $W = V$, $A = U$, $B = 0$ and $w_{i,j} = a_{i,j}$) shows that if $i \in \{1,2,3,\ldots\}$ and $j \in \mathbb{N}$ are such that $i > \dim V$, then

$$a_{i,j} = a_{i-1,j}. \tag{13.97.14}$$

Lemma 13.97.1(e) (applied to $W = V$, $A = U$, $B = 0$ and $w_{i,j} = a_{i,j}$) shows that if $j \in \{1,2,3,\ldots\}$ and $i \in \mathbb{N}$ satisfy $j > \dim V$, then $a_{i,j} = a_{i,j-1}$.

Now, let $(i,j) \in \{1,2,3,\ldots\}^2$. If $i > \dim V$, then

$$b_{i,j} = \underbrace{a_{i,j}}_{\substack{=a_{i-1,j} \\ \text{(by (13.97.14),} \\ \text{since } i>\dim V)}} - \underbrace{a_{i,j-1}}_{\substack{=a_{i-1,j-1} \\ \text{(by (13.97.14) (applied to} \\ j-1 \text{ instead of } j), \\ \text{since } i>\dim V)}} - a_{i-1,j} + a_{i-1,j-1} = a_{i-1,j} - a_{i-1,j-1} - a_{i-1,j} + a_{i-1,j-1} = 0.$$

Similarly, $b_{i,j} = 0$ if $j > \dim V$. Hence, we see that $b_{i,j} = 0$ if we have $i > \dim V$ or $j > \dim V$ (or both). Thus, $b_{i,j} = 0$ for all but finitely many pairs (i,j) (because all but finitely many pairs (i,j) satisfy $i > \dim V$ or $j > \dim V$). This proves (13.97.13).

[762] *Proof of (13.97.15):* Let $(i,j) \in \mathbb{N} \times \{1,2,3,\ldots\}$. Then,

$$\begin{aligned}
b_{1,j} + b_{2,j} + \cdots + b_{i,j} &= \sum_{k=1}^{i} \underbrace{b_{k,j}}_{\substack{=a_{k,j} - a_{k,j-1} - a_{k-1,j} + a_{k-1,j-1} \\ \text{(by the definition of } b_{k,j})}} = \sum_{k=1}^{i} \underbrace{\left(a_{k,j} - a_{k,j-1} - a_{k-1,j} + a_{k-1,j-1}\right)}_{=\left(a_{k,j} - a_{k,j-1}\right) - \left(a_{k-1,j} - a_{k-1,j-1}\right)} \\
&= \sum_{k=1}^{i} \left(\left(a_{k,j} - a_{k,j-1}\right) - \left(a_{k-1,j} - a_{k-1,j-1}\right)\right) = \left(a_{i,j} - a_{i,j-1}\right) - \underbrace{\left(a_{0,j} - a_{0,j-1}\right)}_{\substack{=\left(\mu^t\right)_j \\ \text{(by (13.97.11))}}} \\
&\qquad (\text{by the telescope principle}) \\
&= \left(a_{i,j} - a_{i,j-1}\right) - \left(\mu^t\right)_j,
\end{aligned}$$

and thus $\left(\mu^t\right)_j + \left(b_{1,j} + b_{2,j} + \cdots + b_{i,j}\right) = a_{i,j} - a_{i,j-1}$, qed.

[763] The partition that we call γ in Assertion $\mathcal{A}'$ is the partition that was called λ in Assertion $\mathcal{A}$.

[764] From here, it is easy to see that

$$\dim\left(\left(f^{j}\right)^{-1}(0)\right)-\dim\left(\left(f^{j-1}\right)^{-1}(0)\right)=\left(\lambda^{t}\right)_{j} \qquad \text{for every } j\in\{1,2,3,\ldots\}. \tag{13.97.19}$$

[765] Now, for every $j\in\{1,2,3,\ldots\}$, we have

$$(\text{the number of all entries in the } j\text{-th row of } T)=\gamma_j-\left(\mu^{t}\right)_{j}$$

(since T has shape γ/μ^t), so that

$$\begin{aligned}
\gamma_j-\left(\mu^{t}\right)_{j} &= (\text{the number of all entries in the } j\text{-th row of } T)\\
&=\sum_{i=1}^{\infty}\underbrace{(\text{the number of all entries } i \text{ in the } j\text{-th row of } T)}_{\substack{=b_{i,j}\\ \text{(by (13.97.16))}}}=\sum_{i=1}^{\infty}\underbrace{b_{i,j}}_{=a_{i,j}-a_{i,j-1}-a_{i-1,j}+a_{i-1,j-1}}\\
&=\sum_{i=1}^{\infty}\underbrace{\left(a_{i,j}-a_{i,j-1}-a_{i-1,j}+a_{i-1,j-1}\right)}_{=\left(a_{i,j}-a_{i-1,j}\right)-\left(a_{i,j-1}-a_{i-1,j-1}\right)}=\sum_{i=1}^{\infty}\left(\left(a_{i,j}-a_{i-1,j}\right)-\left(a_{i,j-1}-a_{i-1,j-1}\right)\right)\\
&=\underbrace{\sum_{i=1}^{\infty}\left(a_{i,j}-a_{i-1,j}\right)}_{\substack{=\dim\left(\left(f^{j}\right)^{-1}(0)\right)-\dim\left(U\cap\left(f^{j}\right)^{-1}(0)\right)\\ \text{(by Lemma 13.97.1(f), applied to}\\ W=V,\ A=U,\ B=0 \text{ and } w_{i,j}=a_{i,j})}}-\underbrace{\sum_{i=1}^{\infty}\left(a_{i,j-1}-a_{i-1,j-1}\right)}_{\substack{=\dim\left(\left(f^{j-1}\right)^{-1}(0)\right)-\dim\left(U\cap\left(f^{j-1}\right)^{-1}(0)\right)\\ \text{(by Lemma 13.97.1(f), applied to}\\ V,\ U,\ 0,\ a_{i,j} \text{ and } j-1 \text{ instead of } W,\ A,\ B,\ w_{i,j} \text{ and } j)}}\\
&=\left(\dim\left(\left(f^{j}\right)^{-1}(0)\right)-\dim\left(U\cap\left(f^{j}\right)^{-1}(0)\right)\right)-\left(\dim\left(\left(f^{j-1}\right)^{-1}(0)\right)-\dim\left(U\cap\left(f^{j-1}\right)^{-1}(0)\right)\right)\\
&=\underbrace{\left(\dim\left(\left(f^{j}\right)^{-1}(0)\right)-\dim\left(\left(f^{j-1}\right)^{-1}(0)\right)\right)}_{\substack{=\left(\lambda^{t}\right)_{j}\\ \text{(by (13.97.19))}}}-\left(\dim\left(U\cap\left(f^{j}\right)^{-1}(0)\right)-\dim\left(U\cap\left(f^{j-1}\right)^{-1}(0)\right)\right)\\
&=\left(\lambda^{t}\right)_{j}-\left(\dim\left(U\cap\left(f^{j}\right)^{-1}(0)\right)-\dim\left(U\cap\left(f^{j-1}\right)^{-1}(0)\right)\right).
\end{aligned}$$

[764] *Proof of (13.97.18):* Let $j\in\mathbb{N}$. We can represent the endomorphism f of V as an $n\times n$-matrix $F\in\mathbb{K}^{n\times n}$ for $n=\dim V$. This $n\times n$-matrix F is nilpotent (since f is nilpotent) and has Jordan type λ (since f has Jordan type λ). But Exercise 2.9.22(a) (applied to $N=F$) shows that F has Jordan type λ if and only if every $k\in\mathbb{N}$ satisfies $\dim\left(\ker\left(F^{k}\right)\right)=\left(\lambda^{t}\right)_{1}+\left(\lambda^{t}\right)_{2}+\cdots+\left(\lambda^{t}\right)_{k}$. Since we know that F has Jordan type λ, we thus conclude that every $k\in\mathbb{N}$ satisfies $\dim\left(\ker\left(F^{k}\right)\right)=\left(\lambda^{t}\right)_{1}+\left(\lambda^{t}\right)_{2}+\cdots+\left(\lambda^{t}\right)_{k}$. Applied to $k=j$, this yields $\dim\left(\ker\left(F^{j}\right)\right)=\left(\lambda^{t}\right)_{1}+\left(\lambda^{t}\right)_{2}+\cdots+\left(\lambda^{t}\right)_{j}$.

Since the matrix F represents the endomorphism f, we have

$$\dim\left(\ker\left(F^{j}\right)\right)=\dim\left(\underbrace{\ker\left(f^{j}\right)}_{=\left(f^{j}\right)^{-1}(0)}\right)=\dim\left(\left(f^{j}\right)^{-1}(0)\right).$$

Compared with $\dim\left(\ker\left(F^{j}\right)\right)=\left(\lambda^{t}\right)_{1}+\left(\lambda^{t}\right)_{2}+\cdots+\left(\lambda^{t}\right)_{j}$, this yields

$$\dim\left(\left(f^{j}\right)^{-1}(0)\right)=\left(\lambda^{t}\right)_{1}+\left(\lambda^{t}\right)_{2}+\cdots+\left(\lambda^{t}\right)_{j},$$

and thus (13.97.18) is proven.

[765] *Proof of (13.97.19):* Let $j\in\{1,2,3,\ldots\}$. Then,

$$\begin{aligned}
&\underbrace{\dim\left(\left(f^{j}\right)^{-1}(0)\right)}_{\substack{=\left(\lambda^{t}\right)_{1}+\left(\lambda^{t}\right)_{2}+\cdots+\left(\lambda^{t}\right)_{j}\\ \text{(by (13.97.18))}}}-\underbrace{\dim\left(\left(f^{j-1}\right)^{-1}(0)\right)}_{\substack{=\left(\lambda^{t}\right)_{1}+\left(\lambda^{t}\right)_{2}+\cdots+\left(\lambda^{t}\right)_{j-1}\\ \text{(by (13.97.18), applied}\\ \text{to } j-1 \text{ instead of } j)}}\\
&=\left(\left(\lambda^{t}\right)_{1}+\left(\lambda^{t}\right)_{2}+\cdots+\left(\lambda^{t}\right)_{j}\right)-\left(\left(\lambda^{t}\right)_{1}+\left(\lambda^{t}\right)_{2}+\cdots+\left(\lambda^{t}\right)_{j-1}\right)=\left(\lambda^{t}\right)_{j},
\end{aligned}$$

so that (13.97.19) is proven.

Adding this equality to the equality

$$\begin{aligned}\left(\mu^t\right)_j &= \underbrace{a_{0,j}}_{\substack{=\dim\left(\left(f^0\right)^{-1}(U)\cap\ker\left(f^j\right)\right)\\ \text{(by the definition of } a_{0,j}\text{)}}} - \underbrace{a_{0,j-1}}_{\substack{=\dim\left(\left(f^0\right)^{-1}(U)\cap\ker\left(f^{j-1}\right)\right)\\ \text{(by the definition of } a_{0,j-1}\text{)}}} \qquad \text{(by (13.97.11))}\\ &= \dim\left(\left(\underbrace{f^0}_{=\mathrm{id}}\right)^{-1}(U)\cap\ker\left(f^j\right)\right) - \dim\left(\left(\underbrace{f^0}_{=\mathrm{id}}\right)^{-1}(U)\cap\ker\left(f^{j-1}\right)\right)\\ &= \dim\left(\underbrace{(\mathrm{id})^{-1}(U)}_{=U}\cap\underbrace{\ker\left(f^j\right)}_{=\left(f^j\right)^{-1}(0)}\right) - \dim\left(\underbrace{(\mathrm{id})^{-1}(U)}_{=U}\cap\underbrace{\ker\left(f^{j-1}\right)}_{=\left(f^{j-1}\right)^{-1}(0)}\right)\\ &= \dim\left(U\cap\left(f^j\right)^{-1}(0)\right) - \dim\left(U\cap\left(f^{j-1}\right)^{-1}(0)\right),\end{aligned}$$

we obtain $\gamma_j = \left(\lambda^t\right)_j$ for every $j \in \{1,2,3,\ldots\}$. Thus, $\gamma = \lambda^t$. Hence, T is a column-strict tableau of shape λ^t/μ^t (since T is a column-strict tableau of shape γ/μ^t).

Next, let us prove that $\operatorname{cont} T = \nu^t$. Indeed, we first notice that

$$\dim\left(\left(f^i\right)^{-1}(U)\right) = (\dim U) + \left(\nu^t\right)_1 + \left(\nu^t\right)_2 + \cdots + \left(\nu^t\right)_i \qquad \text{for every } i \in \mathbb{N}. \tag{13.97.20}$$

[766] As a consequence of this, we have

$$\dim\left(\left(f^i\right)^{-1}(U)\right) - \dim\left(\left(f^{i-1}\right)^{-1}(U)\right) = \left(\nu^t\right)_i \qquad \text{for every } i \in \{1,2,3,\ldots\}. \tag{13.97.21}$$

[767]

[766] *Proof of (13.97.20):* Let $i \in \mathbb{N}$. Recall that the nilpotent endomorphism $\overline{f}$ of the quotient space V/U (induced by $f \in \operatorname{End} V$) has Jordan type ν. We can represent this nilpotent endomorphism $\overline{f}$ of V/U as an $\ell\times\ell$-matrix $H \in \mathbb{K}^{\ell\times\ell}$ for $\ell = \dim(V/U)$. This $\ell\times\ell$-matrix H is nilpotent (since $\overline{f}$ is nilpotent) and has Jordan type ν (since $\overline{f}$ has Jordan type ν). But Exercise 2.9.22(a) (applied to ℓ, H and ν instead of n, N and λ) shows that H has Jordan type ν if and only if every $k \in \mathbb{N}$ satisfies $\dim\left(\ker\left(H^k\right)\right) = \left(\nu^t\right)_1 + \left(\nu^t\right)_2 + \cdots + \left(\nu^t\right)_k$. Since we know that H has Jordan type ν, we thus conclude that every $k \in \mathbb{N}$ satisfies $\dim\left(\ker\left(H^k\right)\right) = \left(\nu^t\right)_1 + \left(\nu^t\right)_2 + \cdots + \left(\nu^t\right)_k$. Applied to $k = i$, this yields $\dim\left(\ker\left(H^i\right)\right) = \left(\nu^t\right)_1 + \left(\nu^t\right)_2 + \cdots + \left(\nu^t\right)_i$.

Since the matrix H represents the endomorphism $\overline{f}$, we have

$$\dim\left(\ker\left(\overline{f}^i\right)\right) = \dim\left(\ker\left(H^i\right)\right) = \left(\nu^t\right)_1 + \left(\nu^t\right)_2 + \cdots + \left(\nu^t\right)_i.$$

But $\left(f^i\right)^{-1}(U)$ is a $\mathbb{K}$-vector subspace of V containing U. Thus, $\left(f^i\right)^{-1}(U)/U$ is canonically a $\mathbb{K}$-vector subspace of V/U. Moreover, this subspace $\left(f^i\right)^{-1}(U)/U$ is precisely the kernel $\ker\left(\overline{f}^i\right)$ (this is straightforward to check). Hence,

$$\dim\left(\left(f^i\right)^{-1}(U)/U\right) = \dim\left(\ker\left(\overline{f}^i\right)\right) = \left(\nu^t\right)_1 + \left(\nu^t\right)_2 + \cdots + \left(\nu^t\right)_i.$$

Since

$$\dim\left(\left(f^i\right)^{-1}(U)/U\right) = \dim\left(\left(f^i\right)^{-1}(U)\right) - \dim U,$$

this rewrites as

$$\dim\left(\left(f^i\right)^{-1}(U)\right) - \dim U = \left(\nu^t\right)_1 + \left(\nu^t\right)_2 + \cdots + \left(\nu^t\right)_i.$$

Adding $\dim U$ to both sides of this equality yields (13.97.20).

[767] *Proof of (13.97.21):* Let $i \in \{1,2,3,\ldots\}$. Then,

$$\begin{aligned}&\underbrace{\dim\left(\left(f^i\right)^{-1}(U)\right)}_{\substack{=(\dim U)+\left(\nu^t\right)_1+\left(\nu^t\right)_2+\cdots+\left(\nu^t\right)_i\\ \text{(by (13.97.20))}}} - \underbrace{\dim\left(\left(f^{i-1}\right)^{-1}(U)\right)}_{\substack{=(\dim U)+\left(\nu^t\right)_1+\left(\nu^t\right)_2+\cdots+\left(\nu^t\right)_{i-1}\\ \text{(by (13.97.20), applied}\\ \text{to } i-1 \text{ instead of } i)}}\\ &= \left((\dim U) + \left(\nu^t\right)_1 + \left(\nu^t\right)_2 + \cdots + \left(\nu^t\right)_i\right) - \left((\dim U) + \left(\nu^t\right)_1 + \left(\nu^t\right)_2 + \cdots + \left(\nu^t\right)_{i-1}\right) = \left(\nu^t\right)_i.\end{aligned}$$

Thus, (13.97.21) is proven.

Now, for every $i \in \{1, 2, 3, \ldots\}$, we have

$$\begin{aligned}
(\operatorname{cont} T)_i &= \left|T^{-1}(i)\right| = (\text{the number of entries } i \text{ in } T) \\
&= \sum_{j=1}^{\infty} \underbrace{(\text{the number of all entries } i \text{ in the } j\text{-th row of } T)}_{\substack{=b_{i,j} \\ (\text{by (13.97.16)})}} = \sum_{j=1}^{\infty} \underbrace{b_{i,j}}_{=a_{i,j}-a_{i,j-1}-a_{i-1,j}+a_{i-1,j-1}} \\
&= \sum_{j=1}^{\infty} \underbrace{(a_{i,j}-a_{i,j-1}-a_{i-1,j}+a_{i-1,j-1})}_{=(a_{i,j}-a_{i,j-1})-(a_{i-1,j}-a_{i-1,j-1})} = \sum_{j=1}^{\infty} \left((a_{i,j}-a_{i,j-1})-(a_{i-1,j}-a_{i-1,j-1})\right) \\
&= \underbrace{\sum_{j=1}^{\infty} (a_{i,j}-a_{i,j-1})}_{\substack{=\dim\left(\left(f^i\right)^{-1}(U)\right)-\dim\left(\left(f^i\right)^{-1}(U)\cap 0\right) \\ (\text{by Lemma 13.97.1(g), applied to} \\ W=V,\ A=U,\ B=0 \text{ and } w_{i,j}=a_{i,j})}} - \underbrace{\sum_{j=1}^{\infty} (a_{i-1,j}-a_{i-1,j-1})}_{\substack{=\dim\left(\left(f^{i-1}\right)^{-1}(U)\right)-\dim\left(\left(f^{i-1}\right)^{-1}(U)\cap 0\right) \\ (\text{by Lemma 13.97.1(f), applied to} \\ V,\ U,\ 0,\ a_{i,j} \text{ and } i-1 \text{ instead of } W,\ A,\ B,\ w_{i,j} \text{ and } i)}} \\
&= \left(\dim\left(\left(f^i\right)^{-1}(U)\right) - \dim \underbrace{\left(\left(f^i\right)^{-1}(U)\cap 0\right)}_{=0}\right) - \left(\dim\left(\left(f^{i-1}\right)^{-1}(U)\right) - \dim \underbrace{\left(\left(f^{i-1}\right)^{-1}(U)\cap 0\right)}_{=0}\right) \\
&= \left(\dim\left(\left(f^i\right)^{-1}(U)\right) - \underbrace{\dim 0}_{=0}\right) - \left(\dim\left(\left(f^{i-1}\right)^{-1}(U)\right) - \underbrace{\dim 0}_{=0}\right) \\
&= \dim\left(\left(f^i\right)^{-1}(U)\right) - \dim\left(\left(f^{i-1}\right)^{-1}(U)\right) = \left(\nu^t\right)_i \qquad (\text{by (13.97.21)}).
\end{aligned}$$

Hence, $\operatorname{cont} T = \nu^t$.

We shall next see that for every positive integer j, the weak composition $\operatorname{cont}(T|_{\text{cols}\geq j})$ is a partition. Here, and in the following, we are using the notations of Exercise 2.9.18.

We first notice that every $i \in \{1, 2, 3, \ldots\}$ and $j \in \mathbb{N}$ satisfy

$$(\text{the number of entries } i \text{ in the first } j \text{ rows of } T) = a_{i,j} - a_{i-1,j}. \tag{13.97.22}$$

[768] Now, for every positive integers j and i, we have

$$\begin{aligned}
&(\text{the number of entries } i+1 \text{ in the first } j \text{ rows of } T) \\
&= a_{i+1,j} - a_{i,j} \qquad (\text{by (13.97.22), applied to } i+1 \text{ instead of } i) \\
&\leq a_{i,j-1} - a_{i-1,j-1} \qquad \left(\begin{array}{c}\text{by Lemma 13.97.1(c), applied to} \\ V,\ U,\ 0,\ a_{i,j} \text{ and } j-1 \text{ instead of } W,\ A,\ B,\ w_{i,j} \text{ and } j\end{array}\right) \\
&= (\text{the number of entries } i \text{ in the first } j-1 \text{ rows of } T)
\end{aligned}$$

(since (13.97.22) (applied to $j-1$ instead of j) yields (the number of entries i in the first $j-1$ rows of T) $= a_{i,j-1} - a_{i-1,j-1}$). In other words, for every positive integers j and i, the number of entries $i+1$ in the first j rows of T is $\leq$ to the number of entries i in the first $j-1$ rows of T. In other words, Assertion $\mathcal{D}$ of Exercise 2.9.18(b) (with λ and μ replaced by λ^t and μ^t) is satisfied. Hence, Assertion $\mathcal{C}$ of Exercise 2.9.18(b) (with λ and μ replaced by λ^t and μ^t) is satisfied as well (because Exercise 2.9.18(b) yields that these Assertions $\mathcal{C}$ and $\mathcal{D}$ are equivalent). In other words, for every positive integer j, the weak composition $\operatorname{cont}(T|_{\text{cols}\geq j})$ is a partition.

Now, let us forget that we defined ν and T. We thus have found a column-strict tableau T of shape λ^t/μ^t with $\operatorname{cont} T = \nu^t$ which has the property that for every positive integer j, the weak composition $\operatorname{cont}(T|_{\text{cols}\geq j})$ is a partition. But we know that the number of such tableaux is $c^{\lambda^t}_{\mu^t,\nu^t}$ (by Corollary 2.6.12, applied to λ^t, μ^t and ν^t instead of λ, μ and ν). Hence, this number $c^{\lambda^t}_{\mu^t,\nu^t}$ must be $\neq 0$ (because we have found such a tableau T). So we have proven that $c^{\lambda^t}_{\mu^t,\nu^t} \neq 0$. Now, Exercise 2.7.11(c) yields $c^{\lambda}_{\mu,\nu} = c^{\lambda^t}_{\mu^t,\nu^t} \neq 0$, so that Exercise 2.9.22(b) is solved.

13.98. **Solution to Exercise 2.9.24.** *Solution to Exercise 2.9.24.* Define a set $\mathfrak{D}$ by

$$\mathfrak{D} = \left\{ g \in \Lambda \mid g^{\perp} a = (\omega(g))^{\perp} a \right\}. \tag{13.98.1}$$

We shall show that $\mathfrak{D}$ is a $\mathbf{k}$-subalgebra of Λ.

Define a map $\kappa : \Lambda \to \Lambda$ by

$$\left(\kappa(g) = g^{\perp} a - (\omega(g))^{\perp} a \qquad \text{for every } g \in \Lambda \right).$$

[768] *Proof of (13.97.22):* Let $i \in \{1,2,3,\ldots\}$ and $j \in \mathbb{N}$. We have

$$\begin{aligned}
&(\text{the number of entries } i \text{ in the first } j \text{ rows of } T) \\
&= \sum_{k=1}^{j} \underbrace{(\text{the number of all entries } i \text{ in the } k\text{-th row of } T)}_{\substack{=b_{i,k} \\ (\text{since (13.97.16) (applied to } (i,k) \text{ instead of } (i,j)) \text{ yields} \\ b_{i,k}=(\text{the number of all entries } i \text{ in the } k\text{-th row of } T))}} \\
&= \sum_{k=1}^{j} \underbrace{b_{i,k}}_{\substack{=a_{i,k}-a_{i,k-1}-a_{i-1,k}+a_{i-1,k-1} \\ (\text{by the definition of } b_{i,k})}} = \sum_{k=1}^{j} \underbrace{\left(a_{i,k} - a_{i,k-1} - a_{i-1,k} + a_{i-1,k-1}\right)}_{=(a_{i,k}-a_{i-1,k})-(a_{i,k-1}-a_{i-1,k-1})} \\
&= \sum_{k=1}^{j} \left(\left(a_{i,k} - a_{i-1,k}\right) - \left(a_{i,k-1} - a_{i-1,k-1}\right) \right) \\
&= \left(a_{i,j} - a_{i-1,j}\right) - \left(\underbrace{a_{i,0}}_{\substack{=0 \\ (\text{by (13.97.9)})}} - \underbrace{a_{i-1,0}}_{\substack{=0 \\ (\text{by (13.97.9), applied} \\ \text{to } i-1 \text{ instead of } i)}} \right) \qquad (\text{by the telescope principle}) \\
&= \left(a_{i,j} - a_{i-1,j}\right) - (0-0) = a_{i,j} - a_{i-1,j}.
\end{aligned}$$

This proves (13.97.22).

This map κ is $\mathbf{k}$-linear[769]. Hence, its kernel $\ker\kappa$ is a $\mathbf{k}$-submodule of Λ. Since

$$\ker\kappa = \left\{ g \in \Lambda \mid \underbrace{\kappa(g)}_{\substack{=g^\perp a - (\omega(g))^\perp a \\ \text{(by the definition of } \kappa)}} = 0 \right\} = \left\{ g \in \Lambda \mid \underbrace{g^\perp a - (\omega(g))^\perp a = 0}_{\iff \left(g^\perp a = (\omega(g))^\perp a\right)} \right\}$$
$$= \left\{ g \in \Lambda \mid g^\perp a = (\omega(g))^\perp a \right\} = \mathfrak{D} \qquad \text{(by (13.98.1))},$$

this rewrites as follows: The set $\mathfrak{D}$ is a $\mathbf{k}$-submodule of Λ.

Furthermore, we have the following two observations:

- We have $1 \in \mathfrak{D}$ [770].
- We have $xy \in \mathfrak{D}$ for each $x \in \mathfrak{D}$ and $y \in \mathfrak{D}$ [771].

[769] *Proof.* For each $g \in \Lambda$, we have $g^\perp a = \sum (g, a_1) a_2$ (by the definition of $g^\perp$). Hence, the element $g^\perp a$ of Λ depends $\mathbf{k}$-linearly on g (because the Hall inner product $(\cdot,\cdot)$ is $\mathbf{k}$-bilinear). Thus, the element $(\omega(g))^\perp a$ also depends $\mathbf{k}$-linearly on g (since the map ω is $\mathbf{k}$-linear). Hence, the difference $g^\perp a - (\omega(g))^\perp a$ of these two elements also depends $\mathbf{k}$-linearly on g. In other words, $\kappa(g)$ depends $\mathbf{k}$-linearly on g (since $\kappa(g) = g^\perp a - (\omega(g))^\perp a$). In other words, the map κ is $\mathbf{k}$-linear.

[770] *Proof.* Recall that ω is a $\mathbf{k}$-algebra endomorphism of Λ. Hence, $\omega(1) = 1$. Thus, $\left(\underbrace{\omega(1)}_{=1}\right)^\perp a = 1^\perp a$, so that $1^\perp a = (\omega(1))^\perp a$. Hence, 1 is an element of Λ and satisfies $1^\perp a = (\omega(1))^\perp a$. Thus,

$$1 \in \left\{ g \in \Lambda \mid g^\perp a = (\omega(g))^\perp a \right\} = \mathfrak{D} \qquad \text{(by (13.98.1))},$$

qed.

[771] *Proof.* Let $x \in \mathfrak{D}$ and $y \in \mathfrak{D}$.

We have $x \in \mathfrak{D} = \left\{ g \in \Lambda \mid g^\perp a = (\omega(g))^\perp a \right\}$. In other words, x is an element of Λ and satisfies $x^\perp a = (\omega(x))^\perp a$.

We have $y \in \mathfrak{D} = \left\{ g \in \Lambda \mid g^\perp a = (\omega(g))^\perp a \right\}$. In other words, y is an element of Λ and satisfies $y^\perp a = (\omega(y))^\perp a$.

Proposition 2.8.2(ii) (applied to $f = \omega(y)$ and $g = \omega(x)$) yields

$$(\omega(y)\,\omega(x))^\perp a = (\omega(x))^\perp \left((\omega(y))^\perp a\right). \tag{13.98.2}$$

But $\omega(xy) = \omega(x)\,\omega(y)$ (since ω is an endomorphism of the $\mathbf{k}$-algebra Λ). Hence,

$$\Bigg(\underbrace{\omega(xy)}_{\substack{=\omega(x)\omega(y) \\ =\omega(y)\omega(x)}}\Bigg)^\perp a = (\omega(y)\,\omega(x))^\perp a = (\omega(x))^\perp \left((\omega(y))^\perp a\right) \tag{13.98.3}$$

(by (13.98.2)).

On the other hand, Proposition 2.8.2(ii) (applied to $f = x$ and $g = y$) yields

$$(xy)^\perp a = y^\perp \Bigg(\underbrace{x^\perp a}_{=(\omega(x))^\perp a}\Bigg) = y^\perp \left((\omega(x))^\perp a\right).$$

Comparing this with

$$(\omega(x)\,y)^\perp a = y^\perp \left((\omega(x))^\perp a\right) \qquad \text{(by Proposition 2.8.2(ii) (applied to } f = \omega(x) \text{ and } g = y)),$$

we obtain

$$(xy)^\perp a = \Bigg(\underbrace{\omega(x)\,y}_{=y\omega(x)}\Bigg)^\perp a = (y\omega(x))^\perp a = (\omega(x))^\perp \Bigg(\underbrace{y^\perp a}_{=(\omega(y))^\perp a}\Bigg)$$
$$\text{(by Proposition 2.8.2(ii) (applied to } f = y \text{ and } g = \omega(x)\,))$$
$$= (\omega(x))^\perp \left((\omega(y))^\perp a\right) = (\omega(xy))^\perp a \qquad \text{(by (13.98.3))}.$$

Thus, xy is an element of Λ and satisfies $(xy)^\perp a = (\omega(xy))^\perp a$. Hence,

$$xy \in \left\{ g \in \Lambda \mid g^\perp a = (\omega(g))^\perp a \right\} = \mathfrak{D} \qquad \text{(by (13.98.1))},$$

qed.

Combining these two observations, we conclude that the set $\mathfrak{D}$ is a $\mathbf{k}$-subalgebra of Λ (since we already know that $\mathfrak{D}$ is a $\mathbf{k}$-submodule of Λ). In view of (13.98.1), this rewrites as follows: The set $\left\{g \in \Lambda \ \mid \ g^{\perp} a = (\omega(g))^{\perp} a\right\}$ is a $\mathbf{k}$-subalgebra of Λ. This solves Exercise 2.9.24(a).

(b) We have

$$e_k^{\perp} a = h_k^{\perp} a \qquad \text{for each positive integer } k \tag{13.98.4}$$

(by assumption). Thus,

$$e_n \in \mathfrak{D} \text{ for each } n \in \{1, 2, 3, \ldots\} \tag{13.98.5}$$

[772].

But Proposition 2.4.1 shows that the family $(e_n)_{n \in \{1,2,3,\ldots\}}$ generates Λ as a $\mathbf{k}$-algebra. Thus, the smallest $\mathbf{k}$-subalgebra of Λ that contains all elements of the family $(e_n)_{n \in \{1,2,3,\ldots\}}$ is Λ itself. In other words, if $\mathfrak{B}$ is a $\mathbf{k}$-subalgebra of Λ satisfying

$$(e_n \in \mathfrak{B} \text{ for each } n \in \{1, 2, 3, \ldots\}),$$

then $\mathfrak{B} = \Lambda$. Applying this to $\mathfrak{B} = \mathfrak{D}$, we conclude that $\mathfrak{D} = \Lambda$ (because $\mathfrak{D}$ is a $\mathbf{k}$-subalgebra of Λ, and because it satisfies (13.98.5)).

Now, let $f \in \Lambda$. Then, $f \in \Lambda = \mathfrak{D} = \left\{g \in \Lambda \ \mid \ g^{\perp} a = (\omega(g))^{\perp} a\right\}$ (by (13.98.1)). In other words, f is an element of Λ and satisfies $f^{\perp} a = (\omega(f))^{\perp} a$.

Now, forget that we fixed f. We thus have proven that $f^{\perp} a = (\omega(f))^{\perp} a$ for each $f \in \Lambda$. Renaming f as g in this statement, we conclude the following: $g^{\perp} a = (\omega(g))^{\perp} a$ for each $g \in \Lambda$. This solves Exercise 2.9.24(b).

13.99. **Solution to Exercise 2.9.25.** *Solution to Exercise 2.9.25.* Let us begin by proving a few simple lemmas:

Lemma 13.99.1. *Let $n \in \mathbb{N}$. Let ρ be the partition $(n-1, n-2, \ldots, 1)$. Let $\mu \in \operatorname{Par}$. Then, ρ/μ is a horizontal strip if and only if ρ/μ is a vertical strip.*

Lemma 13.99.1 becomes visually obvious if one draws in one's mind the Ferrers diagram of the staircase partition ρ and attempts to cut off either a horizontal strip or a vertical strip from it (in either case, the only possibilities are to remove some of its corners). But let us give a rigorous proof:

Proof of Lemma 13.99.1. Write the partition μ in the form $\mu = (\mu_1, \mu_2, \mu_3, \ldots)$. Write the partition ρ in the form $\rho = (\rho_1, \rho_2, \rho_3, \ldots)$. Then,

$$(\rho_1, \rho_2, \rho_3, \ldots) = \rho = (n-1, n-2, \ldots, 1) = (n-1, n-2, \ldots, 1, 0, 0, 0, \ldots).$$

Hence,

$$\rho_i = \begin{cases} n - i, & \text{if } i < n; \\ 0, & \text{if } i \geq n \end{cases} \qquad \text{for each positive integer } i. \tag{13.99.1}$$

Now, we are going to prove the following two claims:

Claim 1: If ρ/μ is a horizontal strip, then ρ/μ is a vertical strip.

Claim 2: If ρ/μ is a vertical strip, then ρ/μ is a horizontal strip.

[772] *Proof of (13.98.5):* Let $n \in \{1, 2, 3, \ldots\}$. Thus, n is a positive integer. Hence, $\omega(e_n) = h_n$ (by the definition of ω). Thus, $\Bigg(\underbrace{\omega(e_n)}_{=h_n}\Bigg)^{\perp} a = h_n^{\perp} a$. But (13.98.4) (applied to $k = n$) yields $e_n^{\perp} a = h_n^{\perp} a$. Comparing this with $(\omega(e_n))^{\perp} a = h_n^{\perp} a$, we find $e_n^{\perp} a = (\omega(e_n))^{\perp} a$. Hence, e_n is an element of Λ and satisfies $e_n^{\perp} a = (\omega(e_n))^{\perp} a$. Therefore,

$$e_n \in \left\{g \in \Lambda \ \mid \ g^{\perp} a = (\omega(g))^{\perp} a\right\} = \mathfrak{D} \qquad \text{(by (13.98.1))},$$

qed.

Proof of Claim 1: Assume that ρ/μ is a horizontal strip. We must then prove that ρ/μ is a vertical strip.

We have $\mu \subseteq \rho$ (since ρ/μ is a horizontal strip). Exercise 2.7.5(a) (applied to $\lambda = \rho$ and $\lambda_i = \rho_i$) yields that ρ/μ is a horizontal strip if and only if every $i \in \{1, 2, 3, \ldots\}$ satisfies $\mu_i \geq \rho_{i+1}$. Thus,

$$\text{every } i \in \{1, 2, 3, \ldots\} \text{ satisfies } \mu_i \geq \rho_{i+1} \tag{13.99.2}$$

(since ρ/μ is a horizontal strip).

Now, every $i \in \{1, 2, 3, \ldots\}$ satisfies $\rho_i \leq \mu_i + 1$ [773]. But Exercise 2.7.5(b) (applied to $\lambda = \rho$ and $\lambda_i = \rho_i$) yields that ρ/μ is a vertical strip if and only if every $i \in \{1, 2, 3, \ldots\}$ satisfies $\rho_i \leq \mu_i + 1$. Thus, ρ/μ is a vertical strip (since every $i \in \{1, 2, 3, \ldots\}$ satisfies $\rho_i \leq \mu_i + 1$). This proves Claim 1.

Proof of Claim 2: Assume that ρ/μ is a vertical strip. We must then prove that ρ/μ is a horizontal strip.

We know that ρ/μ is a vertical strip. In particular, $\mu \subseteq \rho$.

Exercise 2.7.5(b) (applied to $\lambda = \rho$ and $\lambda_i = \rho_i$) yields that ρ/μ is a vertical strip if and only if every $i \in \{1, 2, 3, \ldots\}$ satisfies $\rho_i \leq \mu_i + 1$. Thus,

$$\text{every } i \in \{1, 2, 3, \ldots\} \text{ satisfies } \rho_i \leq \mu_i + 1 \tag{13.99.3}$$

(since ρ/μ is a vertical strip).

Now, every $i \in \{1, 2, 3, \ldots\}$ satisfies $\mu_i \geq \rho_{i+1}$ [774]. But Exercise 2.7.5(a) (applied to $\lambda = \rho$ and $\lambda_i = \rho_i$) yields that ρ/μ is a horizontal strip if and only if every $i \in \{1, 2, 3, \ldots\}$ satisfies $\mu_i \geq \rho_{i+1}$. Thus, ρ/μ is a horizontal strip (since every $i \in \{1, 2, 3, \ldots\}$ satisfies $\mu_i \geq \rho_{i+1}$). This proves Claim 2.

[773] *Proof.* Let $i \in \{1, 2, 3, \ldots\}$. We must show that $\rho_i \leq \mu_i + 1$.

If $i \geq n - 1$, then

$$\begin{aligned} \rho_i &= \begin{cases} n - i, & \text{if } i < n; \\ 0, & \text{if } i \geq n \end{cases} \qquad \text{(by (13.99.1))} \\ &\leq \begin{cases} 1, & \text{if } i < n; \\ 0, & \text{if } i \geq n \end{cases} \qquad \text{(since } n - i \leq 1 \text{ (because } i \geq n - 1)) \\ &\leq \begin{cases} 1, & \text{if } i < n; \\ 1, & \text{if } i \geq n \end{cases} \qquad \text{(since } 0 \leq 1) \\ &= 1 = \underbrace{0}_{\leq \mu_i} + 1 \leq \mu_i + 1. \end{aligned}$$

Hence, if $i \geq n - 1$, then $\rho_i \leq \mu_i + 1$ holds. Thus, for the rest of this proof of $\rho_i \leq \mu_i + 1$, we WLOG assume that we don't have $i \geq n - 1$.

We have $i < n - 1$ (since we don't have $i \geq n - 1$). But (13.99.1) yields $\rho_i = \begin{cases} n - i, & \text{if } i < n; \\ 0, & \text{if } i \geq n \end{cases} = n - i$ (since $i < n - 1 < n$).

But from $i < n - 1$, we obtain $i + 1 < n$. Now, (13.99.1) (applied to $i + 1$ instead of i) yields $\rho_{i+1} = \begin{cases} n - (i + 1), & \text{if } i + 1 < n; \\ 0, & \text{if } i + 1 \geq n \end{cases} = n - (i + 1)$ (since $i + 1 < n$).

But (13.99.2) yields $\mu_i \geq \rho_{i+1} = n - (i + 1) = n - i - 1$. Adding 1 to both sides of this inequality, we obtain $\mu_i + 1 \geq n - i = \rho_i$ (since $\rho_i = n - i$). Hence, $\rho_i \leq \mu_i + 1$. This completes our proof of $\rho_i \leq \mu_i + 1$.

[774] *Proof.* Let $i \in \{1, 2, 3, \ldots\}$. We must show that $\mu_i \geq \rho_{i+1}$.

Applying (13.99.1) to $i + 1$ instead of i, we obtain $\rho_{i+1} = \begin{cases} n - (i + 1), & \text{if } i + 1 < n; \\ 0, & \text{if } i + 1 \geq n \end{cases}$.

If $i + 1 \geq n$, then

$$\begin{aligned} \rho_{i+1} &= \begin{cases} n - (i + 1), & \text{if } i + 1 < n; \\ 0, & \text{if } i + 1 \geq n \end{cases} = 0 \qquad \text{(since } i + 1 \geq n) \\ &\leq \mu_i. \end{aligned}$$

In other words, if $i + 1 \geq n$, then $\mu_i \geq \rho_{i+1}$. Thus, for the rest of this proof of $\mu_i \geq \rho_{i+1}$, we WLOG assume that we don't have $i + 1 \geq n$.

We have $i + 1 < n$ (since we don't have $i + 1 \geq n$). Thus, $\rho_{i+1} = \begin{cases} n - (i + 1), & \text{if } i + 1 < n; \\ 0, & \text{if } i + 1 \geq n \end{cases} = n - (i + 1)$ (since $i + 1 < n$).

But (13.99.1) yields $\rho_i = \begin{cases} n - i, & \text{if } i < n; \\ 0, & \text{if } i \geq n \end{cases} = n - i$ (since $i < i + 1 < n$).

From (13.99.3), we obtain $\rho_i \leq \mu_i + 1$, so that $\mu_i + 1 \geq \rho_i = n - i$. Subtracting 1 from both sides of this inequality, we find $\mu_i \geq n - i - 1 = n - (i + 1) = \rho_{i+1}$ (since $\rho_{i+1} = n - (i + 1)$). This completes our proof of $\mu_i \geq \rho_{i+1}$.

Combining Claim 1 with Claim 2, we conclude that ρ/μ is a horizontal strip if and only if ρ/μ is a vertical strip. Lemma 13.99.1 is thus proven. □

Lemma 13.99.2. *Let $n \in \mathbb{N}$. Let ρ be the partition $(n-1, n-2, \ldots, 1)$. Let k be a positive integer. Then, $e_k^\perp s_\rho = h_k^\perp s_\rho$.*

Proof of Lemma 13.99.2. Fix $\nu \in \operatorname{Par}$. We have the following logical equivalence:

$$\begin{aligned} & (\rho/\nu \text{ is a horizontal } k\text{-strip}) \\ \iff & (\rho/\nu \text{ is a horizontal strip and has size } k) \\ & \qquad (\text{by the definition of a "horizontal } k\text{-strip"}) \\ \iff & (\rho/\nu \text{ is a horizontal strip and satisfies } |\rho/\nu| = k) . \end{aligned} \tag{13.99.4}$$

Similarly, we also have the following logical equivalence:

$$\begin{aligned} & (\rho/\nu \text{ is a vertical } k\text{-strip}) \\ \iff & (\rho/\nu \text{ is a vertical strip and satisfies } |\rho/\nu| = k) . \end{aligned} \tag{13.99.5}$$

But Lemma 13.99.1 (applied to $\mu = \nu$) shows that ρ/ν is a horizontal strip if and only if ρ/ν is a vertical strip. In other words, we have the following logical equivalence:

$$(\rho/\nu \text{ is a horizontal strip}) \iff (\rho/\nu \text{ is a vertical strip}) . \tag{13.99.6}$$

Now, we have the following chain of logical equivalences:

$$\begin{aligned} & (\rho/\nu \text{ is a horizontal } k\text{-strip}) \\ \iff & \left(\underbrace{\rho/\nu \text{ is a horizontal strip}}_{\substack{\iff (\rho/\nu \text{ is a vertical strip}) \\ (\text{by (13.99.6)})}} \text{ and satisfies } |\rho/\nu| = k \right) \qquad (\text{by (13.99.4)}) \\ \iff & (\rho/\nu \text{ is a vertical strip and satisfies } |\rho/\nu| = k) \\ \iff & (\rho/\nu \text{ is a vertical } k\text{-strip}) \qquad (\text{by (13.99.5)}) . \end{aligned} \tag{13.99.7}$$

Now, forget that we fixed ν. We thus have proven the equivalence (13.99.7) for each $\nu \in \operatorname{Par}$.

Lemma 13.77.2(a) (applied to $\gamma = \rho$) yields

$$h_k^\perp s_\rho = \underbrace{\sum_{\substack{\nu \in \operatorname{Par}; \\ \rho/\nu \text{ is a} \\ \text{horizontal } k\text{-strip}}}}_{\substack{= \sum_{\substack{\nu \in \operatorname{Par}; \\ \rho/\nu \text{ is a} \\ \text{vertical } k\text{-strip}}} \\ (\text{by the equivalence (13.99.7)})}} s_\nu = \sum_{\substack{\nu \in \operatorname{Par}; \\ \rho/\nu \text{ is a} \\ \text{vertical } k\text{-strip}}} s_\nu .$$

Comparing this with

$$e_k^\perp s_\rho = \sum_{\substack{\nu \in \operatorname{Par}; \\ \rho/\nu \text{ is a} \\ \text{vertical } k\text{-strip}}} s_\nu \qquad (\text{by Lemma 13.77.2(b) (applied to } \gamma = \rho)) ,$$

we obtain $e_k^\perp s_\rho = h_k^\perp s_\rho$. This proves Lemma 13.99.2. □

Lemma 13.99.3. *Let $\lambda \in \operatorname{Par}$. Then, $\omega(s_\lambda) = s_{\lambda^t}$.*

Proof of Lemma 13.99.3. We have $\varnothing \subseteq \lambda$. Hence, the first equation of (2.4.15) (applied to $\mu = \varnothing$) yields $\omega\left(s_{\lambda/\varnothing}\right) = s_{\lambda^t/\varnothing^t} = s_{\lambda^t/\varnothing}$ (since $\varnothing^t = \varnothing$).

But recall that $s_\lambda = s_{\lambda/\varnothing}$. The same argument (but applied to λ^t instead of λ) shows that $s_{\lambda^t} = s_{\lambda^t/\varnothing}$. Comparing this with $\omega\left(\underbrace{s_\lambda}_{=s_{\lambda/\varnothing}}\right) = \omega\left(s_{\lambda/\varnothing}\right) = s_{\lambda^t/\varnothing}$, we obtain $\omega\left(s_\lambda\right) = s_{\lambda^t}$. This proves Lemma 13.99.3.

□

We can now solve Exercise 2.9.25 easily:

Solution to Exercise 2.9.25. Lemma 13.99.2 shows that $e_k^\perp s_\rho = h_k^\perp s_\rho$ for each positive integer k. Hence, Exercise 2.9.24(b) (applied to $a = s_\rho$) yields that

$$g^\perp s_\rho = \left(\omega\left(g\right)\right)^\perp s_\rho \qquad \text{for each } g \in \Lambda. \tag{13.99.8}$$

Recall that

$$s_\mu^\perp\left(s_\lambda\right) = s_{\lambda/\mu} \qquad \text{for every } \lambda \in \operatorname{Par} \text{ and } \mu \in \operatorname{Par}. \tag{13.99.9}$$

Fix $\mu \in \operatorname{Par}$. Then, Lemma 13.99.3 (applied to μ instead of λ) shows that $\omega\left(s_\mu\right) = s_{\mu^t}$. Now, (13.99.8) (applied to $g = s_\mu$) yields

$$s_\mu^\perp s_\rho = \left(\underbrace{\omega\left(s_\mu\right)}_{=s_{\mu^t}}\right)^\perp s_\rho = \left(s_{\mu^t}\right)^\perp s_\rho = \left(s_{\mu^t}\right)^\perp\left(s_\rho\right) = s_{\rho/\mu^t}$$

(by (13.99.9) (applied to ρ and μ^t instead of λ and μ)). Comparing this with

$$s_\mu^\perp s_\rho = s_\mu^\perp\left(s_\rho\right) = s_{\rho/\mu} \qquad \text{(by (13.99.9) (applied to } \lambda = \rho\text{))},$$

this yields $s_{\rho/\mu} = s_{\rho/\mu^t}$. This solves Exercise 2.9.25.

13.100. **Solution to Exercise 3.1.6.** *Solution to Exercise 3.1.6.* Let A be as in Proposition 3.1.2, and assume that $\mathbf{k} = \mathbb{Q}$. We have $\mathfrak{p} \cap I^2 = 0$ (by Proposition 3.1.2 (b)). Thus, Lemma 3.1.4 yields that A is commutative. Since A is self-dual, this yields that A is cocommutative. Hence, Exercise 1.5.14 (d) shows that the $\mathbf{k}$-algebra A is generated by the $\mathbf{k}$-submodule $\mathfrak{p}$. In other words, A is the $\mathbf{k}$-subalgebra of A generated by the $\mathbf{k}$-submodule $\mathfrak{p}$. Since the $\mathbf{k}$-subalgebra of A generated by the $\mathbf{k}$-submodule $\mathfrak{p}$ is $\sum_{n\geq 0}\mathfrak{p}^n$ (this is because generally, if B is a $\mathbf{k}$-algebra, and U is a $\mathbf{k}$-submodule of B, then the $\mathbf{k}$-subalgebra of B generated by the $\mathbf{k}$-submodule U is $\sum_{n\geq 0} U^n$), this rewrites as follows: A is $\sum_{n\geq 0}\mathfrak{p}^n$. Hence,

$$\begin{aligned}
A &= \sum_{n\geq 0}\mathfrak{p}^n = \underbrace{\mathfrak{p}^0}_{=\mathbf{k}\cdot 1_A} + \underbrace{\mathfrak{p}^1}_{=\mathfrak{p}} + \sum_{n\geq 2}\underbrace{\mathfrak{p}^n}_{=\mathfrak{p}\cdot\mathfrak{p}\cdot\mathfrak{p}^{n-2}} = \mathbf{k}\cdot 1_A + \mathfrak{p} + \sum_{n\geq 2}\underbrace{\mathfrak{p}}_{\subset I}\cdot\underbrace{\mathfrak{p}}_{\subset I}\cdot\mathfrak{p}^{n-2} \\
&\subset \mathbf{k}\cdot 1_A + \mathfrak{p} + \sum_{n\geq 2} I \cdot \underbrace{I\cdot\mathfrak{p}^{n-2}}_{\substack{\subset I \\ \text{(since } I \text{ is an ideal of } A\text{)}}} \subset \mathbf{k}\cdot 1_A + \mathfrak{p} + \underbrace{\sum_{n\geq 2} I\cdot I}_{\subset I\cdot I = I^2} \subset \mathbf{k}\cdot 1_A + \mathfrak{p} + I^2.
\end{aligned}$$

Now, let $a \in I$. Then, $\epsilon\left(a\right) = 0$ (by the definition of I). But we know that $a \in I \subset A \subset \mathbf{k}\cdot 1_A + \mathfrak{p} + I^2$. Hence, there exist some $\lambda \in \mathbf{k}$ and some $a' \in \mathfrak{p} + I^2$ such that $a = \lambda\cdot 1_A + a'$. Since $a' \in \mathfrak{p} + I^2 \subset I$, we have $\epsilon\left(a'\right) = 0$. Now, $\epsilon\left(a\right) = 0$, so that

$$0 = \epsilon\left(\underbrace{a}_{=\lambda\cdot 1_A + a'}\right) = \epsilon\left(\lambda\cdot 1_A + a'\right) = \lambda\underbrace{\epsilon\left(1_A\right)}_{=1} + \underbrace{\epsilon\left(a'\right)}_{=0} = \lambda,$$

so that $\lambda = 0$ and thus $a = \underbrace{\lambda}_{=0}\cdot 1_A + a' = a' \in \mathfrak{p} + I^2$. We thus have shown that every $a \in I$ satisfies $a \in \mathfrak{p} + I^2$. Hence, $I \subset \mathfrak{p} + I^2$. Combined with $\mathfrak{p} + I^2 \subset I$, this yields $I = \mathfrak{p} + I^2$, and thus $I = \mathfrak{p} \oplus I^2$ (since $\mathfrak{p} \cap I^2 = 0$). This proves Proposition 3.1.2(c).

13.101. **Solution to Exercise 3.1.9.** *Solution to Exercise 3.1.9.* Corollary 2.5.17(a) yields that $(h_\lambda)_{\lambda \in \mathrm{Par}}$ and $(m_\lambda)_{\lambda \in \mathrm{Par}}$ are dual bases with respect to the Hall inner product on Λ. In other words,

$$(h_\lambda, m_\mu) = \delta_{\lambda,\mu} \qquad \text{for any partitions } \lambda \text{ and } \mu.$$

Thus, every nonempty partition λ satisfies $(h_\lambda, m_\varnothing) = \delta_{\lambda,\varnothing} = 0$. Hence, every nonempty partition λ satisfies

$$\left(h_\lambda, \underbrace{1}_{=m_\varnothing}\right) = (h_\lambda, m_\varnothing) = 0. \tag{13.101.1}$$

Recall the following fundamental fact from linear algebra: If $\mathbf{k}$ is a commutative ring, if A is a $\mathbf{k}$-module, if $(\cdot,\cdot) : A \times A \to \mathbf{k}$ is a symmetric $\mathbf{k}$-bilinear form on A, and if $(u_\lambda)_{\lambda \in L}$ and $(v_\lambda)_{\lambda \in L}$ are two $\mathbf{k}$-bases of A which are dual to each other with respect to the form $(\cdot,\cdot)$ (where L is some indexing set), then every $a \in A$ satisfies

$$a = \sum_{\lambda \in L} (u_\lambda, a)\, v_\lambda. \tag{13.101.2}$$

We can apply this fact to $A = \Lambda$, $L = \mathrm{Par}$, $(u_\lambda)_{\lambda \in L} = (h_\lambda)_{\lambda \in \mathrm{Par}}$ and $(v_\lambda)_{\lambda \in L} = (m_\lambda)_{\lambda \in \mathrm{Par}}$ (since the bases $(h_\lambda)_{\lambda \in \mathrm{Par}}$ and $(m_\lambda)_{\lambda \in \mathrm{Par}}$ of Λ are dual to each other with respect to the Hall inner product $(\cdot,\cdot)$). As a result, we obtain that every $a \in \Lambda$ satisfies

$$a = \sum_{\lambda \in \mathrm{Par}} (h_\lambda, a)\, m_\lambda. \tag{13.101.3}$$

Now, let us solve the exercise. We need to show that for every $a \in \Lambda$, the element a of Λ is primitive if and only if a lies in the $\mathbf{k}$-linear span of $p_1, p_2, p_3, \ldots$. The "if" direction of this statement is obvious[775]. Hence, we only need to prove the "only if" statement. In other words, we need to prove that if a is primitive, then a lies in the $\mathbf{k}$-linear span of $p_1, p_2, p_3, \ldots$. So let us assume that a is primitive. That is, $\Delta(a) = 1 \otimes a + a \otimes 1$.

For every $x \in \Lambda$ and $y \in \Lambda$, we have (using the Sweedler notation)

$$x^\perp(a) = \sum_{(a)} (x, a_1)\, a_2 = (x,1)\, a + (x,a)\, 1 \qquad \left(\text{since } \sum_{(a)} a_1 \otimes a_2 = \Delta(a) = 1 \otimes a + a \otimes 1\right)$$

and therefore

$$\begin{aligned} (xy, a) &= \Bigg(y, \underbrace{x^\perp(a)}_{=(x,1)a+(x,a)1}\Bigg) \qquad (\text{by Proposition 2.8.2(i)}) \\ &= (y, (x,1)\, a + (x,a)\, 1) = (x,1)(y,a) + (x,a)(y,1) = (x,1)(y,a) + (y,1)(x,a). \end{aligned} \tag{13.101.4}$$

Using this, we can easily obtain $(1,a) = 0$ [776].

Now, using (13.101.4), it is easy to see that

$$(h_\lambda, a) = 0 \qquad \text{for every partition } \lambda \text{ satisfying } \ell(\lambda) \geq 2. \tag{13.101.5}$$

[775] because Proposition 2.3.6(i) shows that each of $p_1, p_2, p_3, \ldots$ is primitive, and therefore every element of their $\mathbf{k}$-linear span is also primitive

[776] Indeed, applying (13.101.4) to $x = 1$ and $y = 1$, we obtain $(1,a) = \underbrace{(1,1)}_{=1}(1,a) + \underbrace{(1,1)}_{=1}(1,a) = (1,a) + (1,a) = 2(1,a)$, so that $(1,a) = 0$, qed.

[777] Now, (13.101.3) becomes

$$\begin{aligned} a &= \sum_{\lambda\in\operatorname{Par}} (h_\lambda, a)\, m_\lambda = \underbrace{\sum_{\substack{\lambda\in\operatorname{Par};\\ \ell(\lambda)=0}} (h_\lambda, a)\, m_\lambda}_{=(h_\varnothing,a)m_\varnothing} + \underbrace{\sum_{\substack{\lambda\in\operatorname{Par};\\ \ell(\lambda)=1}} (h_\lambda, a)\, m_\lambda}_{=\sum_{n\geq 1}(h_{(n)},a)m_{(n)}} + \sum_{\substack{\lambda\in\operatorname{Par};\\ \ell(\lambda)\geq 2}} \underbrace{(h_\lambda, a)}_{\substack{=0\\ \text{(by (13.101.5))}}} m_\lambda \\ &= \left(\underbrace{h_\varnothing}_{=1}, a\right) m_\varnothing + \sum_{n\geq 1} \left(h_{(n)}, a\right) \underbrace{m_{(n)}}_{=p_n} + \underbrace{\sum_{\substack{\lambda\in\operatorname{Par};\\ \ell(\lambda)\geq 2}} 0 m_\lambda}_{=0} = \underbrace{(1, a)}_{=0} m_\varnothing + \sum_{n\geq 1} \left(h_{(n)}, a\right) p_n \\ &= \sum_{n\geq 1} \left(h_{(n)}, a\right) p_n . \end{aligned}$$

Thus, a lies in the $\mathbf{k}$-linear span of $p_1, p_2, p_3, \ldots$. This completes the solution to Exercise 3.1.9.

13.102. **Solution to Exercise 4.1.1.** *Solution to Exercise 4.1.1.*

(a) This is straightforward: Let g_1 and g_2 be two elements of G belonging to the same conjugacy class. Thus, g_1 and g_2 are conjugate. In other words, there exists some $x \in G$ such that $g_1 = x g_2 x^{-1}$. Consider this x.

By the definition of $\operatorname{Ind}_H^G f$, we have

$$\begin{aligned} \left(\operatorname{Ind}_H^G f\right)(g_1) &= \frac{1}{|H|} \sum_{\substack{k\in G:\\ kg_1k^{-1}\in H}} f\left(kg_1k^{-1}\right) = \frac{1}{|H|} \sum_{\substack{k\in G:\\ kxg_2x^{-1}k^{-1}\in H}} f\left(kxg_2x^{-1}k^{-1}\right) \qquad \left(\text{since } g_1 = xg_2x^{-1}\right) \\ &= \frac{1}{|H|} \sum_{\substack{k\in G:\\ kxg_2(kx)^{-1}\in H}} f\left(kxg_2\left(kx\right)^{-1}\right) \qquad \left(\text{since } x^{-1}k^{-1} = (kx)^{-1}\right) \\ &= \frac{1}{|H|} \sum_{\substack{k\in G:\\ kg_2k^{-1}\in H}} f\left(kg_2k^{-1}\right) \end{aligned}$$

(here, we have substituted k for kx in the sum, because the map $G \to G,\ k \to kx$ is a bijection). Compared with

$$\left(\operatorname{Ind}_H^G f\right)(g_2) = \frac{1}{|H|} \sum_{\substack{k\in G:\\ kg_2k^{-1}\in H}} f\left(kg_2k^{-1}\right)$$

(this follows from the definition of $\operatorname{Ind}_H^G f$), this yields $\left(\operatorname{Ind}_H^G f\right)(g_1) = \left(\operatorname{Ind}_H^G f\right)(g_2)$.

Now, forget that we fixed g_1 and g_2. We thus have proven that if g_1 and g_2 are two elements of G belonging to the same conjugacy class, then $\left(\operatorname{Ind}_H^G f\right)(g_1) = \left(\operatorname{Ind}_H^G f\right)(g_2)$. In other words, the map $\operatorname{Ind}_H^G f$ is constant

[777] *Proof of (13.101.5):* Let λ be a partition satisfying $\ell(\lambda) \geq 2$. Write the partition λ in the form $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_\ell)$ with $\ell = \ell(\lambda)$. Then, λ_1, λ_2, ..., λ_ℓ are positive integers. Let ν be the partition $(\lambda_2, \lambda_3, \ldots, \lambda_\ell)$; then, ν is nonempty (since $\ell = \ell(\lambda) \geq 2$), so that $(h_\nu, 1) = 0$ (by (13.101.1), applied to ν instead of λ). Also, $\left(h_{(\lambda_1)}, 1\right) = 0$ (by (13.101.1), applied to (λ_1) instead of λ).

By the definition of h_λ, we have $h_\lambda = h_{\lambda_1} h_{\lambda_2} \cdots h_{\lambda_\ell} = \underbrace{h_{\lambda_1}}_{=h_{(\lambda_1)}} \underbrace{(h_{\lambda_2} h_{\lambda_3} \cdots h_{\lambda_\ell})}_{\substack{=h_\nu\\ (\text{since } \nu=(\lambda_2,\lambda_3,\ldots,\lambda_\ell))}} = h_{(\lambda_1)} h_\nu$. Thus,

$$\begin{aligned} \left(\underbrace{h_\lambda}_{=h_{(\lambda_1)}h_\nu}, a\right) &= \left(h_{(\lambda_1)} h_\nu, a\right) = \underbrace{\left(h_{(\lambda_1)}, 1\right)}_{=0} (h_\nu, a) + \underbrace{(h_\nu, 1)}_{=0} \left(h_{(\lambda_1)}, a\right) \qquad \left(\text{by (13.101.4), applied to } x = h_{(\lambda_1)} \text{ and } y = h_\nu\right) \\ &= 0 + 0 = 0. \end{aligned}$$

This proves (13.101.5).

on conjugacy classes. Hence, $\operatorname{Ind}_H^G f$ is a class function on G. In other words, $\operatorname{Ind}_H^G f \in R_{\mathbb{C}}(G)$. This solves part (a) of the exercise.

(b) We have $G = \bigsqcup_{j \in J} Hj$. Thus, every $k \in G$ can be uniquely written in the form hj for some $j \in J$ and $h \in H$. Hence,

$$\begin{aligned}
\sum_{\substack{k\in G:\\ kgk^{-1}\in H}} f\left(kgk^{-1}\right) &= \underbrace{\sum_{\substack{j\in J;\ h\in H:\\ (hj)g(hj)^{-1}\in H}}}_{\substack{=\sum_{\substack{j\in J;\ h\in H:\\ hjgj^{-1}h^{-1}\in H}} = \sum_{\substack{j\in J;\ h\in H:\\ jgj^{-1}\in H}} \\ \text{(because under the condition that } h\in H\text{,} \\ \text{the relation } hjgj^{-1}h^{-1}\in H \text{ is equivalent} \\ \text{to the relation } jgj^{-1}\in H\text{)}}} f\left(\underbrace{(hj)\,g\,(hj)^{-1}}_{=hjgj^{-1}h^{-1}}\right) \\
&= \underbrace{\sum_{\substack{j\in J;\ h\in H:\\ jgj^{-1}\in H}}}_{=\sum_{\substack{j\in J:\\ jgj^{-1}\in H}} \sum_{h\in H}} \underbrace{f\left(hjgj^{-1}h^{-1}\right)}_{\substack{=f\left(jgj^{-1}\right) \\ \text{(since } f \text{ is a class function on } H\text{,} \\ \text{and } hjgj^{-1}h^{-1} \text{ is } H\text{-conjugate to } jgj^{-1}\text{)}}} \\
&= \sum_{\substack{j\in J:\\ jgj^{-1}\in H}} \underbrace{\sum_{h\in H} f\left(jgj^{-1}\right)}_{=|H| f\left(jgj^{-1}\right)} = |H| \sum_{\substack{j\in J:\\ jgj^{-1}\in H}} f\left(jgj^{-1}\right).
\end{aligned} \tag{13.102.1}$$

Now, the definition of $\operatorname{Ind}_H^G f$ yields

$$\left(\operatorname{Ind}_H^G f\right)(g) = \frac{1}{|H|} \underbrace{\sum_{\substack{k\in G:\\ kgk^{-1}\in H}} f\left(kgk^{-1}\right)}_{=|H| \sum_{\substack{j\in J:\\ jgj^{-1}\in H}} f\left(jgj^{-1}\right)} = \sum_{\substack{j\in J:\\ jgj^{-1}\in H}} f\left(jgj^{-1}\right).$$

This solves part (b) of the exercise.

13.103. **Solution to Exercise 4.1.2.** *Solution to Exercise 4.1.2.* We have $\mathbb{C}G \otimes_{\mathbb{C}H} \mathbb{C}H \cong \mathbb{C}G$ as $(\mathbb{C}G, \mathbb{C}H)$-bimodules, and thus also as $(\mathbb{C}G, \mathbb{C}I)$-bimodules[778].

By the definition of induction, we have $\operatorname{Ind}_I^H U = \mathbb{C}H \otimes_{\mathbb{C}I} U$. But by the definition of induction, we also have

$$\begin{aligned}
\operatorname{Ind}_H^G \operatorname{Ind}_I^H U &= \mathbb{C}G \otimes_{\mathbb{C}H} \underbrace{\operatorname{Ind}_I^H U}_{=\mathbb{C}H \otimes_{\mathbb{C}I} U} = \mathbb{C}G \otimes_{\mathbb{C}H} \left(\mathbb{C}H \otimes_{\mathbb{C}I} U\right) \\
&\cong \underbrace{\left(\mathbb{C}G \otimes_{\mathbb{C}H} \mathbb{C}H\right)}_{\substack{\cong \mathbb{C}G \\ \text{(as } (\mathbb{C}G, \mathbb{C}I)\text{-bimodules)}}} \otimes_{\mathbb{C}I} U \qquad \text{(by the associativity of the tensor product)} \\
&\cong \mathbb{C}G \otimes_{\mathbb{C}I} U = \operatorname{Ind}_I^G U
\end{aligned}$$

(since $\operatorname{Ind}_I^G U = \mathbb{C}G \otimes_{\mathbb{C}I} U$ (by the definition of induction)). This solves Exercise 4.1.2.

---

[778] since the right $\mathbb{C}I$-module structures on $\mathbb{C}G \otimes_{\mathbb{C}H} \mathbb{C}H$ and on $\mathbb{C}G$ are obtained from the respective right $\mathbb{C}H$-module structures by restriction

13.104. **Solution to Exercise 4.1.3.** *Solution to Exercise 4.1.3.* The solution rests upon the following general fact:

Proposition 13.104.1. *Let* $\mathbf{k}$ *be a commutative ring. Let* A, B, C, A', B' *and* C' *be six* $\mathbf{k}$*-algebras. Let* P *be an* (A,B)*-bimodule*[779]. *Let* Q *be a* (B,C)*-bimodule. Let* P' *be an* (A',B')*-bimodule. Let* Q' *be a* (B',C')*-bimodule. Then,*

$$(P \otimes P') \otimes_{B \otimes B'} (Q \otimes Q') \cong (P \otimes_B Q) \otimes (P' \otimes_{B'} Q')$$

as $(A \otimes A', C \otimes C')$*-bimodules. Here, all* $\otimes$ *signs without subscript stand for* $\otimes_{\mathbf{k}}$.

This proposition is proven by straightforward (repeated) application of the universal property of the tensor product. (Of course, the $(A \otimes A', C \otimes C')$-bimodule isomorphism $(P \otimes P') \otimes_{B \otimes B'} (Q \otimes Q') \to (P \otimes_B Q) \otimes (P' \otimes_{B'} Q')$ sends every $(p \otimes p') \otimes_{B \otimes B'} (q \otimes q')$ to $(p \otimes_B q) \otimes (p' \otimes_B q')$.) We leave all details to the reader.

Now, let us come to the solution of Exercise 4.1.3.[780] We want to prove the isomorphism

$$\operatorname{Ind}_{H_1 \times H_2}^{G_1 \times G_2} (U_1 \otimes U_2) \cong \left(\operatorname{Ind}_{H_1}^{G_1} U_1 \right) \otimes \left(\operatorname{Ind}_{H_2}^{G_2} U_2 \right)$$

of $\mathbb{C}[G_1 \times G_2]$-modules. Recalling the definition of Ind, we rewrite this as

$$\mathbb{C}[G_1 \times G_2] \otimes_{\mathbb{C}[H_1 \times H_2]} (U_1 \otimes U_2) \cong (\mathbb{C}G_1 \otimes_{\mathbb{C}H_1} U_1) \otimes (\mathbb{C}G_2 \otimes_{\mathbb{C}H_2} U_2) . \qquad (13.104.1)$$

But Proposition 13.104.1 (applied to $\mathbf{k} = \mathbb{C}$, $A = \mathbb{C}G_1$, $B = \mathbb{C}H_1$, $C = \mathbb{C}$, $A' = \mathbb{C}G_2$, $B' = \mathbb{C}H_2$, $C' = \mathbb{C}$, $P = \mathbb{C}G_1$, $Q = U_1$, $P' = \mathbb{C}G_2$ and $Q' = U_2$) yields

$$(\mathbb{C}G_1 \otimes \mathbb{C}G_2) \otimes_{\mathbb{C}H_1 \otimes \mathbb{C}H_2} (U_1 \otimes U_2) \cong (\mathbb{C}G_1 \otimes_{\mathbb{C}H_1} U_1) \otimes (\mathbb{C}G_2 \otimes_{\mathbb{C}H_2} U_2)$$

as $(\mathbb{C}G_1 \otimes \mathbb{C}G_2, \mathbb{C} \otimes \mathbb{C})$-bimodules (thus, as left $\mathbb{C}G_1 \otimes \mathbb{C}G_2$-modules). In order to derive (13.104.1) from this isomorphism, we need to realize that:

- there exist algebra isomorphisms $\mathbb{C}[G_1 \times G_2] \to \mathbb{C}G_1 \otimes \mathbb{C}G_2$ and $\mathbb{C}[H_1 \times H_2] \to \mathbb{C}H_1 \otimes \mathbb{C}H_2$ which commute with the canonical inclusion maps $\mathbb{C}[H_1 \times H_2] \to \mathbb{C}[G_1 \times G_2]$ and $\mathbb{C}H_1 \otimes \mathbb{C}H_2 \to \mathbb{C}G_1 \otimes \mathbb{C}G_2$;
- when we identify $\mathbb{C}[H_1 \times H_2]$ with $\mathbb{C}H_1 \otimes \mathbb{C}H_2$ along the isomorphism $\mathbb{C}[H_1 \times H_2] \to \mathbb{C}H_1 \otimes \mathbb{C}H_2$, the $\mathbb{C}H_1 \otimes \mathbb{C}H_2$-module $U_1 \otimes U_2$ becomes exactly the $\mathbb{C}[H_1 \times H_2]$-module $U_1 \otimes U_2$;
- when we identify $\mathbb{C}[G_1 \times G_2]$ with $\mathbb{C}G_1 \otimes \mathbb{C}G_2$ along the isomorphism $\mathbb{C}[G_1 \times G_2] \to \mathbb{C}G_1 \otimes \mathbb{C}G_2$, the $\mathbb{C}G_1 \otimes \mathbb{C}G_2$-module $(\mathbb{C}G_1 \otimes_{\mathbb{C}H_1} U_1) \otimes (\mathbb{C}G_2 \otimes_{\mathbb{C}H_2} U_2)$ becomes exactly the $\mathbb{C}[G_1 \times G_2]$-module $(\mathbb{C}G_1 \otimes_{\mathbb{C}H_1} U_1) \otimes (\mathbb{C}G_2 \otimes_{\mathbb{C}H_2} U_2)$.

These facts are all trivial to verify (of course, the isomorphism $\mathbb{C}[G_1 \times G_2] \to \mathbb{C}G_1 \otimes \mathbb{C}G_2$ is given by sending every $t_{(g_1,g_2)}$ to $t_{g_1} \otimes t_{g_2}$, and similarly for the other isomorphism). Thus, (13.104.1) holds, and the exercise is solved.

13.105. **Solution to Exercise 4.1.4.** *Solution to Exercise 4.1.4.* In the following, we will write g for the element t_g of $\mathbb{C}G$ whenever g is an element of G. This is a relatively common abuse of notation, and it is harmless because the map $G \to \mathbb{C}G$, $g \mapsto t_g$ is an injective homomorphism of multiplicative monoids (so $t_{gh} = t_g t_h$ and $t_1 = 1$, which means that we won't run into ambiguities denoting t_g by g) and because every $\mathbb{C}G$-module M, every $m \in M$ and every $g \in G$ satisfy $gm = t_g m$.

Recall that $\operatorname{Ind}_H^G U$ is defined as the $\mathbb{C}G$-module $\mathbb{C}G \otimes_{\mathbb{C}H} U$, where $\mathbb{C}G$ is regarded as a $(\mathbb{C}G, \mathbb{C}H)$-bimodule.

The $\mathbb{C}$-vector space $\mathbb{C}G$ is endowed with both a $(\mathbb{C}G, \mathbb{C}H)$-bimodule structure (this is the structure used in the definition of $\operatorname{Ind}_H^G U$) and a $(\mathbb{C}H, \mathbb{C}G)$-bimodule structure (this is the structure used in the statement of Exercise 4.1.4). Thus, $\mathbb{C}G$ has a left $\mathbb{C}G$-module structure, a right $\mathbb{C}H$-module structure, a left $\mathbb{C}H$-module structure, and a right $\mathbb{C}G$-module structure. All four of these structures are simply given by multiplication inside $\mathbb{C}G$ (since $\mathbb{C}H$ is a $\mathbb{C}$-subalgebra of $\mathbb{C}G$). Therefore, notations like xy with x and y being two elements

[779] As usual, we understand the notion of a bimodule to be defined over $\mathbf{k}$; that is, the left A-module structure and the right B-module structure of an (A,B)-bimodule must restrict to one and the same $\mathbf{k}$-module structure.

[780] The following solution involves some handwaving: We are going to use certain isomorphisms to identify $\mathbb{C}[G_1 \times G_2]$ with $\mathbb{C}G_1 \otimes \mathbb{C}G_2$ and to identify $\mathbb{C}[H_1 \times H_2]$ with $\mathbb{C}H_1 \otimes \mathbb{C}H_2$. See the solution of Exercise 4.1.15 for an example of how to avoid this kind of handwaving. (Actually, Exercise 4.1.14(b) shows that Exercise 4.1.15 is a generalization of Exercise 4.1.3.)

of $\mathbb{C}G$ or $\mathbb{C}H$ will never be ambiguous: While they might be interpreted in different ways, all of the possible interpretations will produce identical results.

We recall that $\operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, U)$ is the left $\mathbb{C}G$-module consisting of all left $\mathbb{C}H$-module homomorphisms from $\mathbb{C}G$ to U. This uses only the $(\mathbb{C}H, \mathbb{C}G)$-bimodule structure on $\mathbb{C}G$ that was used in the statement of Exercise 4.1.4 (but not the $(\mathbb{C}G, \mathbb{C}H)$-bimodule structure that was used in the definition of $\operatorname{Ind}_H^G U$).

Let J be a system of distinct representatives for the right H-cosets in G. Then, $G = \bigsqcup_{j \in J} Hj$.

We now define a map $\alpha : \operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, U) \to \mathbb{C}G \otimes_{\mathbb{C}H} U$ by setting

$$\alpha(f) = \sum_{j \in J} j^{-1} \otimes_{\mathbb{C}H} f(j) \qquad \text{for all } f \in \operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, U).$$

This α is a map $\operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, U) \to \operatorname{Ind}_H^G U$ (since $\operatorname{Ind}_H^G U = \mathbb{C}G \otimes_{\mathbb{C}H} U$). It is easy to see that this map α does not depend on the choice of J. (In fact, if j_1 and j_2 are two elements of G lying in the same right H-coset, and if $f \in \operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, U)$, then it is easy to see that $j_1^{-1} \otimes_{\mathbb{C}H} f(j_1) = j_2^{-1} \otimes_{\mathbb{C}H} f(j_2)$.) In other words, if J' is any system of distinct representatives for the right H-cosets in G (which may and may not be equal to J), then

$$\alpha(f) = \sum_{j \in J'} j^{-1} \otimes_{\mathbb{C}H} f(j) \qquad \text{for all } f \in \operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, U). \tag{13.105.1}$$

Notice that if $g \in G$ is arbitrary, then $Jg = \{jg \mid j \in J\}$ is also a system of distinct representatives for the right H-cosets in G.

We will show that α is a $\mathbb{C}G$-module isomorphism.

First, let us prove that α is a left $\mathbb{C}G$-module homomorphism. In fact, any $f \in \operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, U)$ and $g \in G$ satisfy

$$\begin{aligned}
\alpha(gf) &= \sum_{j \in J} \underbrace{j^{-1}}_{= g(jg)^{-1}} \otimes_{\mathbb{C}H} \underbrace{(gf)(j)}_{\substack{= f(jg) \\ \text{(by the definition of } gf)}} = \sum_{j \in J} g(jg)^{-1} \otimes_{\mathbb{C}H} f(jg) \\
&= \sum_{j \in Jg} gj^{-1} \otimes_{\mathbb{C}H} f(j) \qquad \text{(here, we substituted } j \text{ for } jg \text{ in the sum)} \\
&= g \cdot \underbrace{\left(\sum_{j \in Jg} j^{-1} \otimes_{\mathbb{C}H} f(j) \right)}_{\substack{= \alpha(f) \\ \text{(by (13.105.1), applied to } J' = Jg)}} = g \cdot \alpha(f).
\end{aligned}$$

The map α is thus a homomorphism of left G-sets. Since α is furthermore $\mathbb{C}$-linear, this yields that α is a left $\mathbb{C}G$-module homomorphism.

We now are going to construct an inverse for α. This will be more cumbersome.

For every $g \in G$ and every $p \in \mathbb{C}G$, we denote by $\epsilon_g(p)$ the g-coordinate of p with respect to the basis G of the $\mathbb{C}$-vector space $\mathbb{C}G$. By the definition of "coordinate", we have

$$q = \sum_{g \in G} \epsilon_g(q) g \qquad \text{for every } q \in \mathbb{C}G. \tag{13.105.2}$$

For every $g \in G$, we have defined a map $\epsilon_g : \mathbb{C}G \to \mathbb{C}$ (because we have defined an element $\epsilon_g(p)$ for every $p \in \mathbb{C}G$). This map ϵ_g is $\mathbb{C}$-linear. Here are some simple properties of this map:

- For every $g \in G$ and $h \in G$, we have

$$\epsilon_g(h) = \delta_{g,h}. \tag{13.105.3}$$

[781]

- We have

$$\epsilon_1(pq) = \epsilon_1(qp) \qquad \text{for all } p \in \mathbb{C}G \text{ and } q \in \mathbb{C}G. \tag{13.105.4}$$

[781] *Proof.* Let $g \in G$ and $h \in G$. Then, $\epsilon_g(h)$ is defined as the g-coordinate of h with respect to the basis G of the $\mathbb{C}$-vector space $\mathbb{C}G$. This g-coordinate is precisely 1 (since h is an element of this basis G). Thus, $\epsilon_g(h) = 1$, qed.

[782]

- Moreover,

$$\epsilon_1\left(g^{-1}q\right) = \epsilon_g\left(q\right) \qquad \text{for every } g \in G \text{ and } q \in \mathbb{C}G. \tag{13.105.5}$$

[783]

Now, fix $q \in \mathbb{C}G$ and $u \in U$. We let $f_{q,u}$ be the map $\mathbb{C}G \to U$ defined by

$$f_{q,u}\left(p\right) = \sum_{h\in H} \epsilon_1\left(hpq\right) h^{-1}u \qquad \text{for every } p \in \mathbb{C}G. \tag{13.105.6}$$

It is obvious that this map $f_{q,u}$ is $\mathbb{C}$-linear. We will show that $f_{q,u}$ is a left $\mathbb{C}H$-module homomorphism.

The map $f_{q,u}$ is a homomorphism of left H-sets[784]. Since $f_{q,u}$ is furthermore $\mathbb{C}$-linear, this yields that $f_{q,u}$ is a left $\mathbb{C}H$-module homomorphism. Hence, $f_{q,u} \in \operatorname{Hom}_{\mathbb{C}H}\left(\mathbb{C}G, U\right)$.

Now, forget that we fixed q and u. We thus have defined a map $f_{q,u} \in \operatorname{Hom}_{\mathbb{C}H}\left(\mathbb{C}G, U\right)$ for every $q \in \mathbb{C}G$ and $u \in U$. It is easy to see that this map $f_{q,u}$ depends $\mathbb{C}$-linearly on each of q and u. Now, define a map $\widetilde{\beta} : \mathbb{C}G \times U \to \operatorname{Hom}_{\mathbb{C}H}\left(\mathbb{C}G, U\right)$ by

$$\widetilde{\beta}\left(q, u\right) = f_{q,u} \qquad \text{for every } \left(q, u\right) \in \mathbb{C}G \times U.$$

Then, $\widetilde{\beta}$ is a $\mathbb{C}$-bilinear map (because $\widetilde{\beta}\left(q, u\right) = f_{q,u}$ depends $\mathbb{C}$-linearly on each of q and u). We are now going to prove that the map $\widetilde{\beta}$ is $\mathbb{C}H$-bilinear with respect to the right $\mathbb{C}H$-module structure on $\mathbb{C}G$ and the left $\mathbb{C}H$-module structure on U.

In fact, every $h' \in H$, $q \in \mathbb{C}G$ and $u \in U$ satisfy

$$\widetilde{\beta}\left(q, h'u\right) = \widetilde{\beta}\left(qh', u\right). \tag{13.105.7}$$

[782] *Proof.* Let $p \in \mathbb{C}G$ and $q \in \mathbb{C}G$. We need to prove the equality (13.105.4). This equality is $\mathbb{C}$-linear in each of p and q, and thus we can WLOG assume that both p and q belong to the basis G of the $\mathbb{C}$-vector space $\mathbb{C}G$. Assume this. Then, pq and qp belong to G as well. Hence, (13.105.3) yields $\epsilon_1\left(pq\right) = \delta_{1,pq}$ and $\epsilon_1\left(qp\right) = \delta_{1,qp}$. But p and q are elements of the group G. Hence, we have $1 = pq$ if and only if $1 = qp$ (because both of these statements are equivalent to $q = p^{-1}$). Therefore, $\delta_{1,pq} = \delta_{1,qp}$, so that $\epsilon_1\left(pq\right) = \delta_{1,pq} = \delta_{1,qp} = \epsilon_1\left(qp\right)$. This proves (13.105.4).

[783] *Proof.* Let $g \in G$ and $q \in \mathbb{C}G$. We need to prove the equality (13.105.5). This equality is $\mathbb{C}$-linear in q, and thus we can WLOG assume that q belongs to the basis G of the $\mathbb{C}$-vector space $\mathbb{C}G$. Assume this. Then, $g^{-1}q \in G$ as well. Hence, (13.105.3) yields $\epsilon_1\left(g^{-1}q\right) = \delta_{1,g^{-1}q}$ and $\epsilon_g\left(q\right) = \delta_{g,q}$. But g and q are elements of the group G. Hence, we have $1 = g^{-1}q$ if and only if $g = q$. Therefore, $\delta_{1,g^{-1}q} = \delta_{g,q}$, so that $\epsilon_1\left(g^{-1}q\right) = \delta_{1,g^{-1}q} = \delta_{g,q} = \epsilon_g\left(q\right)$. This proves (13.105.5).

[784] *Proof.* In fact, for every $h' \in H$ and every $p \in \mathbb{C}G$, we have

$$\begin{aligned}
f_{q,u}\left(h'p\right) &= \sum_{h\in H} \epsilon_1\left(hh'pq\right) h^{-1}u \qquad \text{(by the definition of } f_{q,u}\text{)} \\
&= \sum_{h\in H} \epsilon_1\left(h \underbrace{\left(h'\right)^{-1} h'}_{=1} pq\right) \underbrace{\left(h\left(h'\right)^{-1}\right)^{-1}}_{=\left(\left(h'\right)^{-1}\right)^{-1}h^{-1}=h'h^{-1}} u \\
&\qquad \left(\begin{array}{c} \text{here, we have substituted } h\left(h'\right)^{-1} \text{ for } h \text{ in the sum,} \\ \text{because the map } H \to H,\ h \mapsto h\left(h'\right)^{-1} \text{ is a bijection} \\ \text{(since } H \text{ is a group and since } h' \in H\text{)} \end{array}\right) \\
&= \sum_{h\in H} \epsilon_1\left(hpq\right) h'h^{-1}u = h' \cdot \underbrace{\sum_{h\in H} \epsilon_1\left(hpq\right) h^{-1}u}_{\substack{=f_{q,u}(p) \\ \text{(by (13.105.6))}}} = h' \cdot f_{q,u}\left(p\right).
\end{aligned}$$

In other words, $f_{q,u}$ is a homomorphism of left H-sets, qed.

[785] As a consequence of this, every $r \in \mathbb{C}H$, $q \in \mathbb{C}G$ and $u \in U$ satisfy

$$\widetilde{\beta}(q, ru) = \widetilde{\beta}(qr, u) \tag{13.105.9}$$

(by $\mathbb{C}$-bilinearity of $\widetilde{\beta}$). In other words, the map $\widetilde{\beta}$ is $\mathbb{C}H$-bilinear with respect to the right $\mathbb{C}H$-module structure on $\mathbb{C}G$ and the left $\mathbb{C}H$-module structure on U. Hence, by the universal property of the tensor product, we conclude that there exists a unique $\mathbb{C}$-linear map $\beta : \mathbb{C}G \otimes_{\mathbb{C}H} U \to \operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, U)$ such that every $(q, u) \in \mathbb{C}G \times U$ satisfies

$$\beta(q \otimes_{\mathbb{C}H} u) = \widetilde{\beta}(q, u). \tag{13.105.10}$$

Consider this map β. Clearly, every $q \in \mathbb{C}G$ and $u \in U$ satisfy

$$\beta(q \otimes_{\mathbb{C}H} u) = \widetilde{\beta}(q, u) = f_{q,u} \qquad \left(\text{by the definition of } \widetilde{\beta}\right).$$

Hence, every $q \in \mathbb{C}G$, $u \in U$ and $p \in \mathbb{C}G$ satisfy

$$\left(\underbrace{\beta(q \otimes_{\mathbb{C}H} u)}_{=f_{q,u}}\right)(p) = f_{q,u}(p) = \sum_{h \in H} \epsilon_1(hpq)\, h^{-1} u \tag{13.105.11}$$

(by (13.105.6)).

We shall now show that the maps α and β are mutually inverse. To do so, we will show that $\alpha \circ \beta = \mathrm{id}$ and $\beta \circ \alpha = \mathrm{id}$.

[785] *Proof of* (13.105.7)*:* Let $h' \in H$, $q \in \mathbb{C}G$ and $u \in U$. Then, the map $H \to H,\ h \mapsto h'h$ is a bijection (since H is a group). Now, let $p \in \mathbb{C}G$. The definition of $\widetilde{\beta}$ yields $\widetilde{\beta}(q, h'u) = f_{q,h'u}$. Hence,

$$\begin{aligned}
\left(\underbrace{\widetilde{\beta}(q, h'u)}_{=f_{q,h'u}}\right)(p) &= f_{q,h'u}(p) = \sum_{h \in H} \epsilon_1(hpq)\, h^{-1} h' u \qquad \left(\text{by the definition of } f_{q,h'u}\right) \\
&= \sum_{h \in H} \epsilon_1(h'hpq) \underbrace{(h'h)^{-1}}_{=h^{-1}(h')^{-1}} h' u \\
&\qquad \left(\begin{array}{c}\text{here, we substituted } h'h \text{ for } h \text{ in the sum,} \\ \text{since the map } H \to H,\ h \mapsto h'h \text{ is a bijection}\end{array}\right) \\
&= \sum_{h \in H} \underbrace{\epsilon_1(h'hpq)}_{\substack{=\epsilon_1(hpqh') \\ \text{(by (13.105.4), applied to} \\ h' \text{ and } hpq \text{ instead of } p \text{ and } q)}} h^{-1} \underbrace{(h')^{-1} h'}_{=1} u \\
&= \sum_{h \in H} \epsilon_1(hpqh')\, h^{-1} u. \qquad (13.105.8)
\end{aligned}$$

But the definition of $\widetilde{\beta}$ also yields $\widetilde{\beta}(qh', u) = f_{qh',u}$. Hence,

$$\begin{aligned}
\left(\widetilde{\beta}(qh', u)\right)(p) &= f_{qh',u}(p) = \sum_{h \in H} \epsilon_1(hpqh')\, h^{-1} u \qquad \left(\text{by the definition of } f_{qh',u}\right) \\
&= \left(\widetilde{\beta}(q, h'u)\right)(p) \qquad (\text{by (13.105.8)}).
\end{aligned}$$

Now, forget that we fixed p. We have thus proven that $\left(\widetilde{\beta}(qh', u)\right)(p) = \left(\widetilde{\beta}(q, h'u)\right)(p)$ for every $p \in \mathbb{C}G$. In other words, $\widetilde{\beta}(qh', u) = \widetilde{\beta}(q, h'u)$, so that $\widetilde{\beta}(q, h'u) = \widetilde{\beta}(qh', u)$. This proves (13.105.7).

Let us first show that $\alpha \circ \beta = \operatorname{id}$. In fact, every $q \in \mathbb{C}G$ and $u \in U$ satisfy

$$\begin{aligned}
(\alpha\circ\beta)(q\otimes_{\mathbb{C}H}u) &= \alpha(\beta(q\otimes_{\mathbb{C}H}u)) = \sum_{j\in J} j^{-1}\otimes_{\mathbb{C}H} \underbrace{(\beta(q\otimes_{\mathbb{C}H}u))(j)}_{\substack{=\sum_{h\in H}\epsilon_1(hjq)h^{-1}u \\ \text{(by (13.105.11), applied to } p=j)}} \\
&\qquad \text{(by the definition of } \alpha\text{)} \\
&= \sum_{j\in J} j^{-1}\otimes_{\mathbb{C}H}\left(\sum_{h\in H}\epsilon_1(hjq)\,h^{-1}u\right) = \sum_{j\in J}\sum_{h\in H}\epsilon_1(hjq) \underbrace{j^{-1}\otimes_{\mathbb{C}H}h^{-1}u}_{\substack{=j^{-1}h^{-1}\otimes_{\mathbb{C}H}u \\ \text{(since } h^{-1}\in H \text{ (since } h\in H\text{))}}} \\
&= \sum_{j\in J}\sum_{h\in H}\epsilon_1(hjq)\underbrace{j^{-1}h^{-1}}_{=(hj)^{-1}}\otimes_{\mathbb{C}H}u = \sum_{j\in J}\sum_{h\in H}\epsilon_1(hjq)(hj)^{-1}\otimes_{\mathbb{C}H}u \\
&= \underbrace{\sum_{j\in J}\sum_{g\in Hj}}_{\substack{=\sum_{g\in G} \\ \text{(since } \bigsqcup_{j\in J}Hj=G\text{)}}} \epsilon_1(gq)\,g^{-1}\otimes_{\mathbb{C}H}u \\
&\qquad \left(\begin{array}{c}\text{here, we substituted } g \text{ for } hj \text{ in the second sum, since the map} \\ H\to Hj,\ h\mapsto hj \text{ is a bijection (because } G \text{ is a group)}\end{array}\right) \\
&= \sum_{g\in G}\epsilon_1(gq)\,g^{-1}\otimes_{\mathbb{C}H}u = \sum_{g\in G}\underbrace{\epsilon_1(g^{-1}q)}_{\substack{=\epsilon_g(q) \\ \text{(by (13.105.5))}}}\underbrace{(g^{-1})^{-1}}_{=g}\otimes_{\mathbb{C}H}u \\
&\qquad \left(\begin{array}{c}\text{here, we substituted } g^{-1} \text{ for } g \text{ in the sum, since the map} \\ G\to G,\ g\mapsto g^{-1} \text{ is a bijection (since } G \text{ is a group)}\end{array}\right) \\
&= \sum_{g\in G}\epsilon_g(q)\,g\otimes_{\mathbb{C}H}u = \underbrace{\left(\sum_{g\in G}\epsilon_g(q)\,g\right)}_{\substack{=q \\ \text{(by (13.105.2))}}}\otimes_{\mathbb{C}H}u \\
&= q\otimes_{\mathbb{C}H}u = \operatorname{id}(q\otimes_{\mathbb{C}H}u).
\end{aligned}$$

Thus, the two maps $\alpha \circ \beta$ and id are equal on every pure tensor. Since these two maps are $\mathbb{C}$-linear, this yields that $\alpha \circ \beta = \operatorname{id}$.

Next, we are going to show that $\beta \circ \alpha = \operatorname{id}$.

Let $f \in \operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, U)$. Let $p \in \mathbb{C}G$. The map f is left $\mathbb{C}H$-linear (since $f \in \operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, U)$), hence $\mathbb{C}$-linear. We have $\alpha(f) = \sum_{j\in J} j^{-1} \otimes_{\mathbb{C}H} f(j)$ (by the definition of α). Applying the map β to this equality, we obtain

$$\beta(\alpha(f)) = \beta\left(\sum_{j\in J} j^{-1}\otimes_{\mathbb{C}H} f(j)\right) = \sum_{j\in J}\beta\left(j^{-1}\otimes_{\mathbb{C}H} f(j)\right)$$

(since β is a $\mathbb{C}$-linear map). Thus,

$$
\begin{aligned}
&\underbrace{\left(\beta\left(\alpha\left(f\right)\right)\right)}_{=\sum_{j\in J}\beta\left(j^{-1}\otimes_{\mathbb{C}H}f\left(j\right)\right)}\left(p\right) \\
&=\left(\sum_{j\in J}\beta\left(j^{-1}\otimes_{\mathbb{C}H}f\left(j\right)\right)\right)\left(p\right)=\sum_{j\in J}\underbrace{\left(\beta\left(j^{-1}\otimes_{\mathbb{C}H}f\left(j\right)\right)\right)\left(p\right)}_{\substack{=\sum_{h\in H}\epsilon_1\left(hpj^{-1}\right)h^{-1}f\left(j\right)\\ \text{(by (13.105.11), applied to}\\ q=j^{-1}\text{ and }u=f\left(j\right))}} \\
&=\sum_{j\in J}\sum_{h\in H}\epsilon_1\left(hpj^{-1}\right)h^{-1}f\left(j\right)=\sum_{j\in J}\sum_{h\in H}\underbrace{\epsilon_1\left(h^{-1}pj^{-1}\right)}_{\substack{=\epsilon_1\left(j^{-1}h^{-1}p\right)\\ \text{(by (13.105.4), applied to}\\ h^{-1}p\text{ and }j^{-1}\text{ instead of }p\text{ and }q)}}\underbrace{\left(h^{-1}\right)^{-1}}_{=h}f\left(j\right) \\
&\qquad\left(\begin{array}{c}\text{here, we substituted }h^{-1}\text{ for }h\text{ in the second sum,}\\ \text{since the map }H\to H,\ h\mapsto h^{-1}\text{ is a bijection}\\ \text{(because }H\text{ is a group)}\end{array}\right) \\
&=\sum_{j\in J}\sum_{h\in H}\epsilon_1\left(\underbrace{j^{-1}h^{-1}}_{=\left(hj\right)^{-1}}p\right)\underbrace{hf\left(j\right)}_{\substack{=f\left(hj\right)\\ \text{(since }f\text{ is left }\mathbb{C}H\text{-linear and since }h\in H\subset\mathbb{C}H)}} \\
&=\sum_{j\in J}\sum_{h\in H}\epsilon_1\left(\left(hj\right)^{-1}p\right)f\left(hj\right)=\underbrace{\sum_{j\in J}\sum_{g\in Hj}}_{\substack{=\sum_{g\in G}\\ \text{(since }\bigsqcup_{j\in J}Hj=G)}}\epsilon_1\left(g^{-1}p\right)f\left(g\right) \\
&\qquad\left(\begin{array}{c}\text{here, we substituted }g\text{ for }hj\text{ in the second sum, since the map}\\ H\to Hj,\ h\mapsto hj\text{ is a bijection (because }G\text{ is a group)}\end{array}\right) \\
&=\sum_{g\in G}\underbrace{\epsilon_1\left(g^{-1}p\right)}_{\substack{=\epsilon_g\left(p\right)\\ \text{(by (13.105.5), applied to }q=p)}}f\left(g\right)=\sum_{g\in G}\epsilon_g\left(p\right)f\left(g\right) \\
&=f\left(\underbrace{\sum_{g\in G}\epsilon_g\left(p\right)g}_{\substack{=p\\ \text{(by (13.105.2))}}}\right)\qquad\left(\text{since }f\text{ is }\mathbb{C}\text{-linear}\right) \\
&=f\left(p\right).
\end{aligned}
$$

Now, forget that we fixed p. We thus have proven that $\left(\beta\left(\alpha\left(f\right)\right)\right)\left(p\right)=f\left(p\right)$ for every $p\in\mathbb{C}G$. In other words, $\beta\left(\alpha\left(f\right)\right)=f$. Hence, $\left(\beta\circ\alpha\right)\left(f\right)=\beta\left(\alpha\left(f\right)\right)=f=\operatorname{id}\left(f\right)$.

Since we have shown this for every f, we can thus conclude that $\beta\circ\alpha=\operatorname{id}$. Combined with $\alpha\circ\beta=\operatorname{id}$, this yields that the maps α and β are mutually inverse. Hence, the map α is invertible, and thus a left $\mathbb{C}G$-module isomorphism (as we already know that α is a left $\mathbb{C}G$-module homomorphism). Hence, $\operatorname{Hom}_{\mathbb{C}H}\left(\mathbb{C}G,U\right)\cong\mathbb{C}G\otimes_{\mathbb{C}H}U=\operatorname{Ind}_H^G U$ as left $\mathbb{C}G$-modules. This solves Exercise 4.1.4.

Remark: In our solution, we explicitly constructed a $\mathbb{C}G$-module isomorphism $\alpha:\operatorname{Hom}_{\mathbb{C}H}\left(\mathbb{C}G,U\right)\to\operatorname{Ind}_H^G U$. This isomorphism is functorial with respect to U. It is also independent on the choice of J (this is not immediately clear from its definition, but it can be shown very easily, by observing that the tensor $j^{-1}\otimes_{\mathbb{C}H}f\left(j\right)$ for $j\in G$ depends only on the coset Hj and not on j itself). One might ask whether this isomorphism is functorial in G and H; but to make sense of this question, one has to define the category with respect to which this functoriality is to be understood. I don't know a good answer.

It is also worth noting that in our solution, $\mathbb{C}$ could be replaced by any commutative ring. We used neither that $\mathbb{C}$ is a field, nor that $\mathbb{C}$ has characteristic 0.

13.106. **Solution to Exercise 4.1.6.** *Solution to Exercise 4.1.6.* Consider the $\mathbb{C}G$-module $\operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, V)$ defined as in Exercise 4.1.4. Then, Exercise 4.1.4 (applied to V instead of U) yields that this module $\operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, V)$ is isomorphic to $\operatorname{Ind}_H^G V$. That is, $\operatorname{Ind}_H^G V \cong \operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, V)$. Hence, $\operatorname{Hom}_{\mathbb{C}G}\left(U, \operatorname{Ind}_H^G V\right) \cong \operatorname{Hom}_{\mathbb{C}G}(U, \operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, V))$. But (4.1.8) (applied to $R = \mathbb{C}G$, $S = \mathbb{C}H$, $A = \mathbb{C}G$, $B = U$ and $C = V$) yields $\operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G \otimes_{\mathbb{C}G} U, V) \cong \operatorname{Hom}_{\mathbb{C}G}(U, \operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, V))$. Altogether, we thus have

$$\operatorname{Hom}_{\mathbb{C}G}\left(U, \operatorname{Ind}_H^G V\right) \cong \operatorname{Hom}_{\mathbb{C}G}(U, \operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, V)) \cong \operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G \otimes_{\mathbb{C}G} U, V).$$

It remains to see what the $\mathbb{C}H$-module $\mathbb{C}G \otimes_{\mathbb{C}G} U$ is. By the basic properties of tensor products, the $\mathbb{C}H$-module $\mathbb{C}G \otimes_{\mathbb{C}G} U$ is isomorphic to the $\mathbb{C}H$-module obtained by restricting the action of $\mathbb{C}G$ on the $\mathbb{C}G$-module U to the subalgebra $\mathbb{C}H$. Since this latter $\mathbb{C}H$-module is precisely $\operatorname{Res}_H^G U$ (in fact, this is how $\operatorname{Res}_H^G U$ was defined!), this rewrites as follows: The $\mathbb{C}H$-module $\mathbb{C}G \otimes_{\mathbb{C}G} U$ is isomorphic to $\operatorname{Res}_H^G U$. In other words, $\mathbb{C}G \otimes_{\mathbb{C}G} U \cong \operatorname{Res}_H^G U$. Now,

$$\operatorname{Hom}_{\mathbb{C}G}\left(U, \operatorname{Ind}_H^G V\right) \cong \operatorname{Hom}_{\mathbb{C}H}\left(\underbrace{\mathbb{C}G \otimes_{\mathbb{C}G} U}_{\cong \operatorname{Res}_H^G U}, V\right) \cong \operatorname{Hom}_{\mathbb{C}H}\left(\operatorname{Res}_H^G U, V\right).$$

This solves the exercise.

13.107. **Solution to Exercise 4.1.9.** *Solution to Exercise 4.1.9.* In the following, a "Hom" symbol without a subscript means "$\operatorname{Hom}_{\mathbb{C}}$" (rather than "$\operatorname{Hom}_{\mathbb{C}G}$" or whatever other meaning this symbol could possibly have in the context).

If G is a group and if M and N are two $\mathbb{C}G$-modules, then $\operatorname{Hom}(M, N)$ becomes a $\mathbb{C}G$-module, with G acting as follows: If $g \in G$ and $f \in \operatorname{Hom}(M, N)$, then $t_g f$ is the $\mathbb{C}$-linear map $M \to N$ sending every $m \in M$ to $t_g f\left(t_{g^{-1}} m\right)$. This $\mathbb{C}G$-module structure is precisely the one we know from Remark 1.4.11.

Now, it is well-known (and straightforward to verify) that every two $\mathbb{C}G$-modules M and N satisfy

$$\operatorname{Hom}_{\mathbb{C}G}(M, N) = (\operatorname{Hom}(M, N))^G \tag{13.107.1}$$

(where we regard $\operatorname{Hom}_{\mathbb{C}G}(M, N)$ as a $\mathbb{C}$-vector subspace of $\operatorname{Hom}(M, N)$ because every $\mathbb{C}G$-linear map $M \to N$ is a $\mathbb{C}$-linear map $M \to N$).

Let us now come to the solution of the exercise.

(a) Let ψ denote the $\mathbb{C}$-linear map

$$\operatorname{Hom}(V_1, W_1) \otimes \operatorname{Hom}(V_2, W_2) \to \operatorname{Hom}(V_1 \otimes V_2, W_1 \otimes W_2)$$

sending each tensor $f \otimes g$ to the tensor product $f \otimes g$ of homomorphisms. This map ψ is completely independent of G_1 and G_2 (it is defined whenever V_1, V_2, W_1 and W_2 are four $\mathbb{C}$-vector spaces) and is a vector space isomorphism (this is a basic fact from linear algebra, relying only on the finite-dimensionality of V_1, V_2, W_1 and W_2). But we can regard $\operatorname{Hom}(V_1, W_1)$ as a $\mathbb{C}G_1$-module, $\operatorname{Hom}(V_2, W_2)$ as a $\mathbb{C}G_2$-module and $\operatorname{Hom}(V_1 \otimes V_2, W_1 \otimes W_2)$ as a $\mathbb{C}[G_1 \times G_2]$-module. Then, the map ψ is a homomorphism of $\mathbb{C}[G_1 \times G_2]$-modules[786]. Hence, this map ψ must be an isomorphism of $\mathbb{C}[G_1 \times G_2]$-modules (being a vector space isomorphism). As a consequence, it sends the $G_1 \times G_2$-fixed space of its domain to the $G_1 \times G_2$-fixed space of its target:

$$\psi\left((\operatorname{Hom}(V_1, W_1) \otimes \operatorname{Hom}(V_2, W_2))^{G_1 \times G_2}\right) = (\operatorname{Hom}(V_1 \otimes V_2, W_1 \otimes W_2))^{G_1 \times G_2}.$$

[786] This is easy to verify by checking that $t_{(h_1,h_2)}(\psi(f \otimes g)) = \psi\left(t_{h_1} f \otimes t_{h_2} g\right)$ for all $(h_1, h_2) \in G_1 \times G_2$, $f \in \operatorname{Hom}(V_1, W_1)$ and $g \in \operatorname{Hom}(V_2, W_2)$.

Since

$$\begin{aligned}
&\left(\operatorname{Hom}\left(V_1,W_1\right)\otimes\operatorname{Hom}\left(V_2,W_2\right)\right)^{G_1\times G_2}\\
&= \underbrace{\left(\operatorname{Hom}\left(V_1,W_1\right)\right)^{G_1}}_{\substack{=\operatorname{Hom}_{\mathbb{C}G_1}(V_1,W_1)\\ \text{(by (13.107.1),}\\ \text{applied to } G=G_1,\ M=V_1 \text{ and } N=W_1)}} \otimes \underbrace{\left(\operatorname{Hom}\left(V_2,W_2\right)\right)^{G_2}}_{\substack{=\operatorname{Hom}_{\mathbb{C}G_2}(V_2,W_2)\\ \text{(by (13.107.1),}\\ \text{applied to } G=G_2,\ M=V_2 \text{ and } N=W_2)}}\\
&\qquad\left(\begin{array}{c}\text{by (4.1.15), applied to } K_1=G_1,\ K_2=G_2,\\ U_1=\operatorname{Hom}\left(V_1,W_1\right) \text{ and } U_2=\operatorname{Hom}\left(V_2,W_2\right)\end{array}\right)\\
&= \operatorname{Hom}_{\mathbb{C}G_1}\left(V_1,W_1\right)\otimes\operatorname{Hom}_{\mathbb{C}G_2}\left(V_2,W_2\right),
\end{aligned}$$

this rewrites as

$$\psi\left(\operatorname{Hom}_{\mathbb{C}G_1}\left(V_1,W_1\right)\otimes\operatorname{Hom}_{\mathbb{C}G_2}\left(V_2,W_2\right)\right)=\left(\operatorname{Hom}\left(V_1\otimes V_2,W_1\otimes W_2\right)\right)^{G_1\times G_2}.$$

Hence, the isomorphism ψ restricts to an isomorphism from $\operatorname{Hom}_{\mathbb{C}G_1}\left(V_1,W_1\right)\otimes\operatorname{Hom}_{\mathbb{C}G_2}\left(V_2,W_2\right)$ to $\left(\operatorname{Hom}\left(V_1\otimes V_2,W_1\otimes W_2\right)\right)^{G_1\times G_2}$. This restriction is precisely the $\mathbb{C}$-linear map

$$\operatorname{Hom}_{\mathbb{C}G_1}\left(V_1,W_1\right)\otimes\operatorname{Hom}_{\mathbb{C}G_2}\left(V_2,W_2\right)\to\operatorname{Hom}_{\mathbb{C}[G_1\times G_2]}\left(V_1\otimes V_2,W_1\otimes W_2\right)$$

sending each tensor $f\otimes g$ to the tensor product $f\otimes g$ of homomorphisms (i.e., the map alleged to be an isomorphism in the statement of the exercise). So we know now that this map is an isomorphism. This solves Exercise 4.1.9(a).

(b) Let G_1 and G_2 be two groups. Let V_i and W_i be finite-dimensional $\mathbb{C}G_i$-modules for every $i\in\{1,2\}$. Exercise 4.1.9(a) provides a vector space isomorphism from $\operatorname{Hom}_{\mathbb{C}G_1}\left(V_1,W_1\right)\otimes\operatorname{Hom}_{\mathbb{C}G_2}\left(V_2,W_2\right)$ to $\operatorname{Hom}_{\mathbb{C}[G_1\times G_2]}\left(V_1\otimes V_2,W_1\otimes W_2\right)$. As a consequence,

$$\begin{aligned}
&\dim_{\mathbb{C}}\left(\operatorname{Hom}_{\mathbb{C}G_1}\left(V_1,W_1\right)\otimes\operatorname{Hom}_{\mathbb{C}G_2}\left(V_2,W_2\right)\right)\\
&=\dim_{\mathbb{C}}\left(\operatorname{Hom}_{\mathbb{C}[G_1\times G_2]}\left(V_1\otimes V_2,W_1\otimes W_2\right)\right). \qquad (13.107.2)
\end{aligned}$$

Applying (4.1.1) to G_1, V_1 and W_1 instead of G, V_1 and V_2, we obtain

$$\left(\chi_{V_1},\chi_{W_1}\right)_{G_1}=\dim_{\mathbb{C}}\left(\operatorname{Hom}_{\mathbb{C}G_1}\left(V_1,W_1\right)\right). \qquad (13.107.3)$$

Applying (4.1.1) to G_2, V_2 and W_2 instead of G, V_1 and V_2, we obtain

$$\left(\chi_{V_2},\chi_{W_2}\right)_{G_2}=\dim_{\mathbb{C}}\left(\operatorname{Hom}_{\mathbb{C}G_2}\left(V_2,W_2\right)\right). \qquad (13.107.4)$$

Applying (4.1.1) to $G_1\times G_2$, $V_1\otimes V_2$ and $W_1\otimes W_2$ instead of G, V_1 and V_2, we obtain

$$\begin{aligned}
\left(\chi_{V_1\otimes V_2},\chi_{W_1\otimes W_2}\right)_{G_1\times G_2}&=\dim_{\mathbb{C}}\left(\operatorname{Hom}_{\mathbb{C}[G_1\times G_2]}\left(V_1\otimes V_2,W_1\otimes W_2\right)\right)\\
&=\dim_{\mathbb{C}}\left(\operatorname{Hom}_{\mathbb{C}G_1}\left(V_1,W_1\right)\otimes\operatorname{Hom}_{\mathbb{C}G_2}\left(V_2,W_2\right)\right) \qquad \text{(by (13.107.2))}\\
&=\underbrace{\dim_{\mathbb{C}}\left(\operatorname{Hom}_{\mathbb{C}G_1}\left(V_1,W_1\right)\right)}_{\substack{=\left(\chi_{V_1},\chi_{W_1}\right)_{G_1}\\ \text{(by (13.107.3))}}}\cdot\underbrace{\dim_{\mathbb{C}}\left(\operatorname{Hom}_{\mathbb{C}G_2}\left(V_2,W_2\right)\right)}_{\substack{=\left(\chi_{V_2},\chi_{W_2}\right)_{G_2}\\ \text{(by (13.107.4))}}}\\
&=\left(\chi_{V_1},\chi_{W_1}\right)_{G_1}\left(\chi_{V_2},\chi_{W_2}\right)_{G_2}.
\end{aligned}$$

This proves (4.1.2). Thus, Exercise 4.1.9(b) is solved.

13.108. **Solution to Exercise 4.1.10.** *Solution to Exercise 4.1.10.* If A and B are two algebras, P is a (B,A)-bimodule and Q is a left B-module, then $\operatorname{Hom}_B(P,Q)$ is a left A-module. Consequently, $\operatorname{Hom}_{\mathbb{C}H}(CG,U)$ is a left $\mathbb{C}G$-module, and $\operatorname{Hom}_{\mathbb{C}[H/(H\cap K)]}\left(\mathbb{C}\left[G/K\right],U^{H\cap K}\right)$ is a left $\mathbb{C}\left[G/K\right]$-module. Exercise 4.1.4 yields that the $\mathbb{C}G$-module $\operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G,U)$ is isomorphic to $\operatorname{Ind}_H^G U$. In other words,

$$\operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G,U)\cong\operatorname{Ind}_H^G U \qquad \text{as } \mathbb{C}G\text{-modules}. \qquad (13.108.1)$$

Also, Exercise 4.1.4 (applied to G/K, $H/(H\cap K)$ and $U^{H\cap K}$ instead of G, H and U) yields that the $\mathbb{C}\left[G/K\right]$-module $\operatorname{Hom}_{\mathbb{C}[H/(H\cap K)]}\left(\mathbb{C}\left[G/K\right],U^{H\cap K}\right)$ is isomorphic to $\operatorname{Ind}_{H/(H\cap K)}^{G/K}\left(U^{H\cap K}\right)$. In other words,

$$\operatorname{Hom}_{\mathbb{C}[H/(H\cap K)]}\left(\mathbb{C}\left[G/K\right],U^{H\cap K}\right)\cong\operatorname{Ind}_{H/(H\cap K)}^{G/K}\left(U^{H\cap K}\right) \qquad \text{as } \mathbb{C}\left[G/K\right]\text{-modules}. \qquad (13.108.2)$$

But we also have

$$(\operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, U))^K \cong \operatorname{Hom}_{\mathbb{C}[H/(H\cap K)]}\left(\mathbb{C}[G/K], U^{H\cap K}\right) \qquad \text{as } \mathbb{C}[G/K]\text{-modules.} \tag{13.108.3}$$

Here is just a brief sketch of the proof of (13.108.3): Let π be the canonical projection $G \to G/K$, and let $\mathbb{C}[\pi]$ be the $\mathbb{C}$-linear map $\mathbb{C}G \to \mathbb{C}[G/K]$ obtained by $\mathbb{C}$-linearly extending π. Clearly, $\mathbb{C}[\pi]$ is a surjective $\mathbb{C}G$-linear $\mathbb{C}$-algebra homomorphism, and we have

$$\begin{aligned}
\ker(\mathbb{C}[\pi]) &= \langle g - g' \mid g \in G,\ g' \in G,\ gK = g'K\rangle && (13.108.4)\\
&= \langle g - gk \mid g \in G,\ k \in K\rangle && (13.108.5)\\
&= \langle g - g' \mid g \in G,\ g' \in G,\ Kg = Kg'\rangle && (13.108.6)\\
&= \langle g - kg \mid g \in G,\ k \in K\rangle && (13.108.7)
\end{aligned}$$

(where the $\langle\cdot\rangle$ brackets stand for "$\mathbb{C}$-span").

Let f be an element of $(\operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, U))^K$. Then, $f \in (\operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, U))^K \subset \operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, U)$ is $\mathbb{C}H$-linear, and it can easily be shown that $\ker(\mathbb{C}[\pi]) \subset \ker f$ [787]. Hence, the map f factors through the surjective map $\mathbb{C}[\pi]$. The resulting map $\mathbb{C}[G/K] \to U$ is a $\mathbb{C}H$-module homomorphism (since $\mathbb{C}[\pi]$ and f were both $\mathbb{C}H$-linear), and it is easy to see that its image (i.e., the image of f) is contained in $U^{H\cap K}$ [788]. Hence, this map factors through the canonical inclusion $U^{H\cap K} \to U$, leaving behind a map $\mathbb{C}[G/K] \to U^{H\cap K}$ which we denote by $\Phi(f)$. This resulting map $\Phi(f)$ turns out to be $\mathbb{C}[H/(H\cap K)]$-linear[789], thus belongs to $\operatorname{Hom}_{\mathbb{C}[H/(H\cap K)]}\left(\mathbb{C}[G/K], U^{H\cap K}\right)$. Since this holds for every f, we thus obtain a $\mathbb{C}$-linear map

$$\begin{aligned}
(\operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, U))^K &\to \operatorname{Hom}_{\mathbb{C}[H/(H\cap K)]}\left(\mathbb{C}[G/K], U^{H\cap K}\right),\\
f &\mapsto \Phi(f).
\end{aligned}$$

This map is invertible[790] and $\mathbb{C}[G/K]$-linear[791], therefore an isomorphism of $\mathbb{C}[G/K]$-modules. This proves (13.108.3). All steps that were left to the reader are straightforward.

Now,

$$\begin{aligned}
\Bigg(\underbrace{\operatorname{Ind}_H^G U}_{\substack{\cong \operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G,U) \text{ as } \mathbb{C}G\text{-modules}\\ (\text{by } (13.108.1))}}\Bigg)^K &\cong (\operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, U))^K\\
&\cong \operatorname{Hom}_{\mathbb{C}[H/(H\cap K)]}\left(\mathbb{C}[G/K], U^{H\cap K}\right) \qquad (\text{by } (13.108.3))\\
&\cong \operatorname{Ind}_{H/(H\cap K)}^{G/K}\left(U^{H\cap K}\right) \qquad (\text{by } (13.108.2))
\end{aligned}$$

as $\mathbb{C}[G/K]$-modules. This solves Exercise 4.1.10.

[787] Indeed, $f \in (\operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, U))^K$. Hence, every $k \in K$ satisfies $kf = f$. Thus, every $g \in G$ and $k \in K$ satisfy $(kf)(g) = f(g)$. But since $(kf)(g) = f(gk)$ (by the definition of the action of G on $\operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, U)$), this becomes $f(gk) = f(g)$, so that $f(g - gk) = 0$. The map f therefore annihilates $g - gk$ for all $g \in G$ and $k \in K$. Due to (13.108.5), this yields $f(\ker(\mathbb{C}[\pi])) = 0$, so that $\ker(\mathbb{C}[\pi]) \subset \ker f$, qed.

[788] *Proof.* We want to show that the image of f is contained in $U^{H\cap K}$.

For this, it clearly suffices to prove that $f(g) \in U^{H\cap K}$ for every $g \in G$. So fix $g \in G$. Let $k \in H \cap K$. Then, $k \cdot f(g) = f(kg)$ (since f is $\mathbb{C}H$-linear and $k \in H$) and $f(g) = f(kg)$ (since $k \in K$, so that (13.108.7) yields $g - kg \in \ker(\mathbb{C}[\pi]) \subset \ker f$ and thus $f(g - kg) = 0$). Hence, $k \cdot f(g) = f(kg) = f(g)$. Since this has been proven for all $k \in H \cap K$, we thus have $f(g) \in U^{H\cap K}$, qed.

[789] This is straightforward to see (everything in sight is $\mathbb{C}H$-linear).

[790] In fact, defining the inverse is very easy (just send every map $g \in \operatorname{Hom}_{\mathbb{C}[H/(H\cap K)]}\left(\mathbb{C}[G/K], U^{H\cap K}\right)$ to the composition $\mathbb{C}G \xrightarrow{\mathbb{C}[\pi]} \mathbb{C}[G/K] \xrightarrow{g} U^{H\cap K} \xrightarrow{\text{inclusion}} U$). Checking that these maps are mutually inverse is also straightforward.

[791] Indeed, it is easier to check that its inverse is $\mathbb{C}[G/K]$-linear (this can be proven by straightforward computations).

13.109. **Solution to Exercise 4.1.11.** *Solution to Exercise 4.1.11.* In the following, we will write g for the element t_g of $\mathbb{C}G$ whenever g is an element of G. This is a relatively common abuse of notation, and it is harmless because the map $G \to \mathbb{C}G,\ g \mapsto t_g$ is an injective homomorphism of multiplicative monoids (so $t_{gh} = t_g t_h$ and $t_1 = 1$, which means that we won't run into ambiguities denoting t_g by g) and because every $\mathbb{C}G$-module M, every $m \in M$ and every $g \in G$ satisfy $gm = t_g m$. We will do the same abuse of notation for elements of H and of K.

Inflation does not change the underlying $\mathbb{C}$-vector space of a representation. Thus, $\operatorname{Infl}_{G/K}^{G} \operatorname{Ind}_{H/K}^{G/K} V = \operatorname{Ind}_{H/K}^{G/K} V$ as $\mathbb{C}$-vector spaces. For the same reason, $\operatorname{Infl}_{H/K}^{H} V = V$ as $\mathbb{C}$-vector spaces.

Let π_G be the canonical projection map $G \to G/K$. This gives rise to a surjective $\mathbb{C}$-algebra homomorphism $\mathbb{C}[\pi_G] : \mathbb{C}G \to \mathbb{C}[G/K]$ (which sends every $g \in G$ to $\pi_G(g)$). We now define a $\mathbb{C}$-linear map

$$\beta : \mathbb{C}G \otimes_{\mathbb{C}H} \operatorname{Infl}_{H/K}^{H} V \to \mathbb{C}[G/K] \otimes_{\mathbb{C}[H/K]} V, \\ s \otimes_{\mathbb{C}H} v \mapsto (\mathbb{C}[\pi_G])(s) \otimes_{\mathbb{C}[H/K]} v.$$

This is easily seen to be well-defined (using the universal property of the tensor product and the observation that every $t \in \mathbb{C}H$ satisfies $(\mathbb{C}[\pi_G])(t) \in \mathbb{C}[H/K]$).

We also want to define a map α in the opposite direction, but this will require some more work. First, for every $g \in G$, we define a $\mathbb{C}$-linear map

$$\mathbf{a}_g : V \to \mathbb{C}G \otimes_{\mathbb{C}H} \operatorname{Infl}_{H/K}^{H} V, \\ v \mapsto g \otimes_{\mathbb{C}H} v.$$

It is easily seen that if g_1 and g_2 are two elements of G satisfying $\pi_G(g_1) = \pi_G(g_2)$, then $\mathbf{a}_{g_1} = \mathbf{a}_{g_2}$ [792]. In other words, the map $\mathbf{a}_g$ depends only on $\pi_G(g)$ rather than on g itself. Thus, we can define a $\mathbb{C}$-linear map $\widetilde{\mathbf{a}}_p : V \to \mathbb{C}G \otimes_{\mathbb{C}H} \operatorname{Infl}_{H/K}^{H} V$ for every $p \in G/K$ by choosing any $g \in G$ satisfying $\pi_G(g) = p$, and then setting $\widetilde{\mathbf{a}}_p = \mathbf{a}_g$; the resulting map $\widetilde{\mathbf{a}}_p$ does not depend on the choice of g.

Hence, we have defined a $\mathbb{C}$-linear map $\widetilde{\mathbf{a}}_p : V \to \mathbb{C}G \otimes_{\mathbb{C}H} \operatorname{Infl}_{H/K}^{H} V$ for every $p \in G/K$. In other words, we have defined an element $\widetilde{\mathbf{a}}_p$ of $\operatorname{Hom}_{\mathbb{C}}\left(V, \mathbb{C}G \otimes_{\mathbb{C}H} \operatorname{Infl}_{H/K}^{H} V\right)$ for every $p \in G/K$. Hence, we can define a $\mathbb{C}$-linear map

$$\mathbf{A} : \mathbb{C}[G/K] \to \operatorname{Hom}_{\mathbb{C}}\left(V, \mathbb{C}G \otimes_{\mathbb{C}H} \operatorname{Infl}_{H/K}^{H} V\right), \\ p \mapsto \widetilde{\mathbf{a}}_p \qquad \text{for every } p \in G/K$$

(because in order to define a $\mathbb{C}$-linear map from the $\mathbb{C}$-vector space $\mathbb{C}[G/K]$, it is enough to assign its values on the basis G/K). Using this map $\mathbf{A}$, we can now define a $\mathbb{C}$-linear map

$$\alpha : \mathbb{C}[G/K] \otimes_{\mathbb{C}[H/K]} V \to \mathbb{C}G \otimes_{\mathbb{C}H} \operatorname{Infl}_{H/K}^{H} V, \\ s \otimes_{\mathbb{C}[H/K]} v \mapsto (\mathbf{A}(s))(v).$$

[792] Indeed, let g_1 and g_2 be two elements of G satisfying $\pi_G(g_1) = \pi_G(g_2)$. Then, $g_1 \in g_2 K$, so that there exists some $k \in K$ such that $g_1 = g_2 k$. Consider this k. Then, every $v \in V$ satisfies

$$\begin{aligned} \mathbf{a}_{g_1}(v) &= \underbrace{g_1}_{=g_2 k} \otimes_{\mathbb{C}H} v \qquad (\text{by the definition of } \mathbf{a}_{g_1}) \\ &= g_2 k \otimes_{\mathbb{C}H} v = g_2 \otimes_{\mathbb{C}H} \underbrace{kv}_{\substack{=v \\ (\text{since } k \in K \text{ acts trivially} \\ \text{on } \operatorname{Infl}_{H/K}^{H} V)}} \qquad (\text{since } k \text{ lies in } K \subset H, \text{ and thus can be moved past the } \otimes_{\mathbb{C}H} \text{ sign}) \\ &= g_2 \otimes_{\mathbb{C}H} v = \mathbf{a}_{g_2}(v) \qquad (\text{by the definition of } \mathbf{a}_{g_2}). \end{aligned}$$

Thus, $\mathbf{a}_{g_1} = \mathbf{a}_{g_2}$, qed.

It is again easy to check that this is well-defined[793]. It is straightforward to show that $\alpha \circ \beta = \operatorname{id}$ (indeed, this only needs to be proven on tensors of the form $g \otimes_{\mathbb{C}H} v$ for $g \in G$ and $v \in \operatorname{Infl}_{H/K}^H V$, because the $\mathbb{C}$-vector space $\mathbb{C}G \otimes_{\mathbb{C}H} \operatorname{Infl}_{H/K}^H V$ is spanned by such tensors; but on such tensors it is very easy to check) and $\beta \circ \alpha = \operatorname{id}$ (using a similar argument). Thus, the maps α and β are mutually inverse, and therefore β is invertible. But we can regard β as a map from $\operatorname{Ind}_H^G \operatorname{Infl}_{H/K}^H V$ to $\operatorname{Infl}_{G/K}^G \operatorname{Ind}_{H/K}^{G/K} V$ (since $\operatorname{Ind}_H^G \operatorname{Infl}_{H/K}^H V = \mathbb{C}G \otimes_{\mathbb{C}H} \operatorname{Infl}_{H/K}^H V$ and $\operatorname{Infl}_{G/K}^G \operatorname{Ind}_{H/K}^{G/K} V = \operatorname{Ind}_{H/K}^{G/K} V = \mathbb{C}[G/K] \otimes_{\mathbb{C}[H/K]} V$ as $\mathbb{C}$-vector spaces), and it is easy to verify that β becomes a $\mathbb{C}G$-module homomorphism when regarded this way. Thus, β is an invertible $\mathbb{C}G$-module homomorphism from $\operatorname{Ind}_H^G \operatorname{Infl}_{H/K}^H V$ to $\operatorname{Infl}_{G/K}^G \operatorname{Ind}_{H/K}^{G/K} V$, hence a $\mathbb{C}G$-module isomorphism. Thus, such an isomorphism exists, i.e., we have $\operatorname{Infl}_{G/K}^G \operatorname{Ind}_{H/K}^{G/K} V \cong \operatorname{Ind}_H^G \operatorname{Infl}_{H/K}^H V$ as $\mathbb{C}G$-modules. This solves Exercise 4.1.11.

13.110. **Solution to Exercise 4.1.12.** *Solution to Exercise 4.1.12.* (a) It is clearly enough to show that $g(v-kv) \in I_{V,K}$ for all $g \in G$, $k \in K$ and $v \in V$. But if $g \in G$, $k \in K$ and $v \in V$, then gk has the form $gk = k'g$ for some $k' \in K$ (since $K \triangleleft G$), and thus we have $g(v-kv) = gv - \underbrace{gk}_{=k'g} v = gv - k'gv \in I_{V,K}$ (by the definition of $I_{V,K}$, since $k' \in K$ and $gv \in V$). Hence, part (a) of the exercise is solved.

(b) For every $v \in V$, let $\overline{v}$ denote the projection of v onto the quotient space $V/I_{V,K} = V_K$.

We can define a map $\Phi : V^K \to V_K$ by sending every $v \in V^K$ to $\overline{v} \in V_K$. This map Φ is easily seen to be a $\mathbb{C}G$-module homomorphism $\operatorname{Infl}_{G/K}^G \left(V^K\right) \to V_K$. If we can show that Φ is also bijective, then it will follow that Φ is a $\mathbb{C}G$-module isomorphism $\operatorname{Infl}_{G/K}^G \left(V^K\right) \to V_K$, whence part (b) of the exercise will be solved. Hence, all that remains to be done is proving that Φ is bijective.

Let us construct an inverse map to Φ. First, let us notice that every $v \in V$ satisfies

$$\frac{1}{|K|} \sum_{j \in K} jv \in V^K.$$

[794] Thus, we can define a map $\psi : V \to V^K$ by sending every $v \in V$ to $\frac{1}{|K|} \sum_{j \in K} jv$. This map ψ is $\mathbb{C}$-linear and vanishes on $I_{V,K}$ [795]. Hence, the map ψ factors through the quotient $V/I_{V,K} = V_K$. Let us denote

[793] *Proof.* In order to prove this well-definedness, we have to check that $(\mathbf{A}(s))(v)$ depends $\mathbb{C}[H/K]$-bilinearly on (s,v). It is very easy to see that $(\mathbf{A}(s))(v)$ depends $\mathbb{C}$-bilinearly on (s,v); therefore, it only remains to prove that $(\mathbf{A}(st))(v) = (\mathbf{A}(s))(tv)$ for all $s \in \mathbb{C}[G/K]$, $t \in \mathbb{C}[H/K]$ and $v \in V$. So let $s \in \mathbb{C}[G/K]$, $t \in \mathbb{C}[H/K]$ and $v \in V$ be arbitrary. We want to prove that $(\mathbf{A}(st))(v) = (\mathbf{A}(s))(tv)$. Since both sides of this equality are $\mathbb{C}$-linear in s, we can WLOG assume that s belongs to the basis G/K of $\mathbb{C}[G/K]$. Assume this, and pick $g \in G$ such that $s = \pi_G(g)$. (This g exists since π_G is surjective.)

Similarly, we can WLOG assume that t belongs to the basis H/K of $\mathbb{C}[H/K]$. Assume this and pick $h \in H$ such that $t = \pi_G(h)$. (This t exists since $H/K = \pi_G(H)$.) Since $s = \pi_G(g)$ and $t = \pi_G(h)$, we have $st = \pi_G(g) \cdot \pi_G(h) = \pi_G(gh)$ (as π_G is a group homomorphism). Notice that $\underbrace{t}_{=\pi_G(h)} v = \pi_G(h) v = hv$ (because the action of H on $\operatorname{Infl}_{H/K}^H V$ factors through π_G).

Now, the definition of $\mathbf{A}(st)$ yields $\mathbf{A}(st) = \widetilde{\mathbf{a}}_{st} = \mathbf{a}_{gh}$ (by the definition of $\widetilde{\mathbf{a}}_{st}$, since $st = \pi_G(gh)$). Similarly, $\mathbf{A}(s) = \mathbf{a}_g$. Now, comparing $\underbrace{(\mathbf{A}(st))}_{=\mathbf{a}_{gh}}(v) = \mathbf{a}_{gh}(v) = gh \otimes_{\mathbb{C}H} v = g \otimes_{\mathbb{C}H} hv$ (since h belongs to $\mathbb{C}H$ and thus can be moved past the $\otimes_{\mathbb{C}H}$ sign) with $\underbrace{(\mathbf{A}(s))}_{=\mathbf{a}_g}\left(\underbrace{tv}_{=hv}\right) = \mathbf{a}_g(hv) = g \otimes_{\mathbb{C}H} hv$, we obtain $(\mathbf{A}(st))(v) = (\mathbf{A}(s))(tv)$, which is precisely what we needed to prove.

[794] *Proof.* Let $v \in V$. Let $j \in K$. Then, the map $K \to K,\ s \mapsto js$ is a bijection (since $j \in K$ and since K is a group). Now,

$$j \cdot \left(\frac{1}{|K|} \sum_{s \in K} sv \right) = \frac{1}{|K|} \sum_{s \in K} jsv = \frac{1}{|K|} \sum_{s \in K} sv$$

(here, we have substituted s for js in the sum, because the map $K \to K,\ s \mapsto js$ is a bijection). Now, forget that we fixed j. We thus have shown that $j \cdot \left(\frac{1}{|K|} \sum_{s \in K} sv \right) = \frac{1}{|K|} \sum_{s \in K} sv$ for every $j \in K$. Hence, $\frac{1}{|K|} \sum_{s \in K} sv \in V^K$, qed.

[795] *Proof.* We want to show that ψ vanishes on $I_{V,K}$. In order to do so, we only need to check that $\psi(v-kv) = 0$ for all $k \in K$ and $v \in V$ (since $I_{V,K}$ is spanned by all $v - kv$ with $k \in K$ and $v \in V$). But this follows from the fact that all $k \in K$

the resulting $\mathbb{C}$-linear map $V_K \to V^K$ by Ψ. We are now going to show that the maps Φ and Ψ are mutually inverse.

We have $\Phi \circ \Psi = \mathrm{id}$ [796] and $\Psi \circ \Phi = \mathrm{id}$ [797]. Hence, the maps Φ and Ψ are mutually inverse. It follows that Φ is bijective, and so the solution is complete.

13.111. **Solution to Exercise 4.1.14.** *Solution to Exercise 4.1.14.* In the following, we will use the following convention: Whenever K is a group, and k is an element of K, we shall write k for the element t_k of $\mathbb{C}K$. This is a relatively common abuse of notation, and it is harmless because the map $K \to \mathbb{C}K,\ k \mapsto t_k$ is an injective homomorphism of multiplicative monoids (so $t_{gh} = t_g t_h$ and $t_1 = 1$, which means that we won't run into ambiguities denoting t_k by k) and because every $\mathbb{C}K$-module M, every $m \in M$ and every $k \in K$ satisfy $km = t_k m$.

and $v \in V$ satisfy

$$\begin{aligned}
\psi(v - kv) &= \frac{1}{|K|} \sum_{j \in K} \underbrace{j(v - kv)}_{= jv - jkv} \qquad \text{(by the definition of } \psi\text{)} \\
&= \frac{1}{|K|} \sum_{j \in K} (jv - jkv) = \frac{1}{|K|} \sum_{j \in K} jv - \frac{1}{|K|} \sum_{j \in K} jkv = \frac{1}{|K|} \sum_{j \in K} jv - \frac{1}{|K|} \sum_{j \in K} jv \\
&\qquad \left(\begin{array}{c} \text{here, we have substituted } j \text{ for } jk \text{ in the second sum, since the map } K \to K,\ j \mapsto jk \\ \text{is a bijection (because } k \in K \text{ and because } K \text{ is a group)} \end{array} \right) \\
&= 0.
\end{aligned}$$

[796] *Proof.* Let $w \in V_K$. Then, there exists some $v \in V$ such that $w = \overline{v}$. Consider this v. Since Ψ was defined as a quotient of the map ψ, we have $\Psi(\overline{v}) = \psi(v)$. Now,

$$\begin{aligned}
(\Phi \circ \Psi)\left(\underbrace{w}_{=\overline{v}}\right) &= (\Phi \circ \Psi)(\overline{v}) = \Phi\left(\underbrace{\Psi(\overline{v})}_{=\psi(v)}\right) = \Phi\left(\underbrace{\psi(v)}_{\substack{= \frac{1}{|K|}\sum_{j \in K} jv \\ \text{(by the definition of } \psi\text{)}}}\right) = \Phi\left(\frac{1}{|K|} \sum_{j \in K} jv\right) \\
&= \overline{\frac{1}{|K|} \sum_{j \in K} jv} \qquad \text{(by the definition of } \Phi\text{)} \\
&= \frac{1}{|K|} \sum_{j \in K} \underbrace{\overline{jv}}_{\substack{= \overline{v} \\ \text{(since } v - jv \in I_{V,K} \\ \text{(because } j \in K\text{))}}} = \frac{1}{|K|} \underbrace{\sum_{j \in K} \overline{v}}_{= |K| \cdot \overline{v}} = \frac{1}{|K|} |K| \cdot \overline{v} = \overline{v} = w.
\end{aligned}$$

Thus we have shown that $(\Phi \circ \Psi)(w) = w$ for every $w \in V_K$. In other words, $\Phi \circ \Psi = \mathrm{id}$, qed.

[797] *Proof.* For every $v \in V^K$, we have

$$\begin{aligned}
(\Psi \circ \Phi)(v) &= \Psi\left(\underbrace{\Phi(v)}_{\substack{= \overline{v} \\ \text{(by the definition of } \Phi\text{)}}}\right) = \Psi(\overline{v}) = \psi(v) \\
&\qquad \text{(since } \Psi \text{ was defined as the quotient of the map } \psi\text{)} \\
&= \frac{1}{|K|} \sum_{j \in K} \underbrace{jv}_{\substack{= v \\ \text{(since } v \in V^K \text{ and } j \in K\text{)}}} \qquad \text{(by the definition of } \psi\text{)} \\
&= \frac{1}{|K|} \underbrace{\sum_{j \in K} v}_{= |K| \cdot v} = \frac{1}{|K|} |K| \cdot v = v.
\end{aligned}$$

Thus, $\Psi \circ \Phi = \mathrm{id}$, qed.

Let us first notice a trivial fact: If K is a finite group, and if $f : K \to \mathbb{C}$ is any function, then we have the following equivalence:

$$\begin{aligned} & \left(f \in R_{\mathbb{C}}\left(K\right)\right) \\ \Longleftrightarrow\ & \left(f \text{ is a class function on } K\right) \\ & \qquad \left(\begin{array}{c} \text{since } R_{\mathbb{C}}\left(K\right) \text{ is defined to be} \\ \text{the set of all class functions on } K \end{array}\right) \\ \Longleftrightarrow\ & \left(f \text{ is constant on } K\text{-conjugacy classes}\right) \\ & \qquad \left(\begin{array}{c} \text{since a class function on } K \text{ is defined to mean a function} \\ K \to \mathbb{C} \text{ which is constant on } K\text{-conjugacy classes} \end{array}\right) \\ \Longleftrightarrow\ & \left(\text{any two conjugate elements } k \text{ and } k' \text{ of } K \text{ satisfy } f\left(k\right) = f\left(k'\right)\right). \end{aligned} \tag{13.111.1}$$

(a) Let $f \in R_{\mathbb{C}}\left(H\right)$.

We can apply (13.111.1) to $K = H$. As a consequence, we obtain the following equivalence:

$$\begin{aligned} & \left(f \in R_{\mathbb{C}}\left(H\right)\right) \\ \Longleftrightarrow\ & \left(\text{any two conjugate elements } k \text{ and } k' \text{ of } H \text{ satisfy } f\left(k\right) = f\left(k'\right)\right). \end{aligned}$$

Hence,

$$\text{any two conjugate elements } k \text{ and } k' \text{ of } H \text{ satisfy } f\left(k\right) = f\left(k'\right) \tag{13.111.2}$$

(because we know that $f \in R_{\mathbb{C}}\left(H\right)$).

Now, let g and g' be two conjugate elements of G. Then, there exists a $p \in G$ such that $g' = pgp^{-1}$ (since g and g' are conjugate). Consider this p.

Applying the map $\operatorname{Ind}_\rho f$ to both sides of the equality $g' = pgp^{-1}$, we obtain

$$(\operatorname{Ind}_\rho f)(g') = (\operatorname{Ind}_\rho f)\left(pgp^{-1}\right) = \frac{1}{|H|} \underbrace{\sum_{\substack{(h,k)\in H\times G;\\ k\rho(h)k^{-1}=pgp^{-1}}}}_{=\sum_{h\in H}\sum_{\substack{k\in G;\\ k\rho(h)k^{-1}=pgp^{-1}}}} f(h)$$

$$\left(\text{by the definition of } \operatorname{Ind}_\rho f\right)$$

$$= \frac{1}{|H|}\sum_{h\in H} \underbrace{\sum_{\substack{k\in G;\\ k\rho(h)k^{-1}=pgp^{-1}}} f(h)}_{\substack{=\sum_{\substack{k\in G;\\ pk\rho(h)(pk)^{-1}=pgp^{-1}}} f(h)\\ \text{(here, we have substituted } pk \text{ for } k \text{ in the sum,}\\ \text{since the map } G\to G,\ k\mapsto pk \text{ is a bijection)}}}$$

$$= \frac{1}{|H|}\sum_{h\in H} \underbrace{\sum_{\substack{k\in G;\\ pk\rho(h)(pk)^{-1}=pgp^{-1}}}}_{\substack{=\sum_{\substack{k\in G;\\ pk\rho(h)k^{-1}p^{-1}=pgp^{-1}}}\\ \text{(since } (pk)^{-1}=k^{-1}p^{-1}\text{)}}} f(h) = \frac{1}{|H|}\sum_{h\in H} \underbrace{\sum_{\substack{k\in G;\\ pk\rho(h)k^{-1}p^{-1}=pgp^{-1}}}}_{\substack{=\sum_{\substack{k\in G;\\ pk\rho(h)k^{-1}=pg}}\\ \text{(since } pk\rho(h)k^{-1}p^{-1}=pgp^{-1} \text{ is}\\ \text{equivalent to } pk\rho(h)k^{-1}=pg\text{)}}} f(h)$$

$$= \frac{1}{|H|}\sum_{h\in H} \underbrace{\sum_{\substack{k\in G;\\ pk\rho(h)k^{-1}=pg}}}_{\substack{=\sum_{\substack{k\in G;\\ k\rho(h)k^{-1}=g}}\\ \text{(since } pk\rho(h)k^{-1}=pg \text{ is}\\ \text{equivalent to } k\rho(h)k^{-1}=g\text{)}}} f(h) = \frac{1}{|H|} \underbrace{\sum_{h\in H}\sum_{\substack{k\in G;\\ k\rho(h)k^{-1}=g}}}_{=\sum_{\substack{(h,k)\in H\times G;\\ k\rho(h)k^{-1}=g}}} f(h)$$

$$= \frac{1}{|H|}\sum_{\substack{(h,k)\in H\times G;\\ k\rho(h)k^{-1}=g}} f(h) = (\operatorname{Ind}_\rho f)(g)$$

(since the definition of $\operatorname{Ind}_\rho f$ yields $(\operatorname{Ind}_\rho f)(g) = \frac{1}{|H|}\sum_{\substack{(h,k)\in H\times G;\\ k\rho(h)k^{-1}=g}} f(h)$).

Let us now forget that we fixed g and g'. We thus have shown that any two conjugate elements g and g' of G satisfy $(\operatorname{Ind}_\rho f)(g) = (\operatorname{Ind}_\rho f)(g')$. Renaming g and g' as k and k' in this statement, we obtain the following: Any two conjugate elements k and k' of G satisfy $(\operatorname{Ind}_\rho f)(k) = (\operatorname{Ind}_\rho f)(k')$.

But (13.111.1) (applied to G and $\operatorname{Ind}_\rho f$ instead of K and f) yields the following equivalence:

$$\begin{aligned} &\left(\operatorname{Ind}_\rho f \in R_{\mathbb{C}}(G)\right)\\ \iff &\left(\text{any two conjugate elements } k \text{ and } k' \text{ of } G \text{ satisfy } (\operatorname{Ind}_\rho f)(k) = (\operatorname{Ind}_\rho f)(k')\right). \end{aligned}$$

Thus, $\operatorname{Ind}_\rho f \in R_{\mathbb{C}}(G)$ (because we know that any two conjugate elements k and k' of G satisfy $(\operatorname{Ind}_\rho f)(k) = (\operatorname{Ind}_\rho f)(k')$). This solves Exercise 4.1.14(a).

(b) Let us first introduce an elementary (but apocryphal) notion from linear algebra: the notion of *finite dual generating systems.*

Definition 13.111.1. Let $\mathbb{K}$ be a commutative ring. Let V be a $\mathbb{K}$-module. A *finite dual generating system* for V means a triple $\left(I, (a_i)_{i\in I}, (f_i)_{i\in I}\right)$, where

- I is a finite set;
- $(a_i)_{i\in I}$ is a family of elements of V;

- $(f_i)_{i\in I}$ is a family of elements of V^* (where V^* means $\operatorname{Hom}_{\mathbb{K}}(V,\mathbb{K})$)

such that every $v \in V$ satisfies $v = \sum\limits_{i\in I} f_i(v)\, a_i$.

In the following, we shall only use finite dual generating systems in the case when $\mathbb{K}$ is a field; nevertheless, they are more useful in the general case.

The first question one might ask about finite dual generating systems for V is when they exist. The answer is very simple when $\mathbb{K}$ is a field:

Proposition 13.111.2. *Let $\mathbb{K}$ be a field. Let V be a $\mathbb{K}$-vector space. Then, a finite dual generating system for V exists if and only if the vector space V is finite-dimensional.*

Proof of Proposition 13.111.2. Proposition 13.111.2 is an "if and only if" statement. Hence, in order to prove Proposition 13.111.2, it is sufficient to verify the following two claims:

> *Claim 1:* If a finite dual generating system for V exists, then the vector space V is finite-dimensional.
>
> *Claim 2:* If the vector space V is finite-dimensional, then a finite dual generating system for V exists.

Let us now prove these two claims.

Proof of Claim 1. Assume that a finite dual generating system for V exists. Let $\left(I, (a_i)_{i\in I}, (f_i)_{i\in I}\right)$ be such a finite dual generating system for V.

We know that I is a finite set, that $(a_i)_{i\in I}$ is a family of elements of V, and that every $v \in V$ satisfies $v = \sum\limits_{i\in I} f_i(v)\, a_i$. (Indeed, this is part of what it means for $\left(I, (a_i)_{i\in I}, (f_i)_{i\in I}\right)$ to be a finite dual generating system for V.)

Now, every $v \in V$ satisfies $v = \sum\limits_{i\in I} \underbrace{f_i(v)}_{\in\mathbb{K}} a_i \in \sum\limits_{i\in I} \mathbb{K}a_i$. Thus, $V \subset \sum\limits_{i\in I} \mathbb{K}a_i$. Combined with the (obvious) inclusion $\sum\limits_{i\in I} \mathbb{K}a_i \subset V$, this yields $V = \sum\limits_{i\in I} \mathbb{K}a_i$. But the vector space $\sum\limits_{i\in I} \mathbb{K}a_i$ is finite-dimensional (since I is a finite set). In other words, the vector space V is finite-dimensional (since $V = \sum\limits_{i\in I} \mathbb{K}a_i$). This proves Claim 1.

Proof of Claim 2. Assume that V is a finite-dimensional vector space. Then, V has a finite basis. Let $(e_i)_{i\in I}$ be such a basis. Thus, I is a finite set, and $(e_i)_{i\in I}$ is a basis of the $\mathbb{K}$-vector space V. Let $(e_i^*)_{i\in I}$ be the basis of V^* dual to the basis $(e_i)_{i\in I}$ of V. (This is well-defined, since V is finite-dimensional.)

We know that $(e_i^*)_{i\in I}$ is the basis of V^* dual to the basis $(e_i)_{i\in I}$ of V. Thus,

$$e_i^* \left(\sum_{j\in I} \lambda_j e_j \right) = \lambda_i \qquad \text{for all } i \in I \text{ and } (\lambda_j)_{j\in I} \in \mathbb{K}^I. \tag{13.111.3}$$

(Indeed, this is one of the ways to define a dual basis.)

Now, every $v \in V$ satisfies $v = \sum\limits_{i\in I} e_i^*(v)\, e_i$ [798].

So we know that $\left(I, (e_i)_{i\in I}, (e_i^*)_{i\in I}\right)$ is a triple such that

- I is a finite set;
- $(e_i)_{i\in I}$ is a family of elements of V;

[798] *Proof.* Let $v \in V$. Then, we can write v in the form $v = \sum\limits_{i\in I} \lambda_i e_i$ for some family $(\lambda_i)_{i\in I} \in \mathbb{K}^I$ (since $(e_i)_{i\in I}$ is a basis of V). Consider this family $(\lambda_i)_{i\in I}$. Now, $v = \sum\limits_{i\in I} \lambda_i e_i = \sum\limits_{j\in I} \lambda_j e_j$ (here, we have renamed the summation index i as j in the sum). For each $i \in I$, we now have

$$e_i^* \left(\underbrace{v}_{= \sum\limits_{j\in I} \lambda_j e_j} \right) = e_i^* \left(\sum_{j\in I} \lambda_j e_j \right) = \lambda_i \qquad \text{(by (13.111.3))}.$$

Thus, $\sum\limits_{i\in I} \underbrace{e_i^*(v)}_{=\lambda_i} e_i = \sum\limits_{i\in I} \lambda_i e_i$. Compared with $v = \sum\limits_{i\in I} \lambda_i e_i$, this yields $v = \sum\limits_{i\in I} e_i^*(v)\, e_i$, qed.

- $\left(e_i^*\right)_{i\in I}$ is a family of elements of V^* (where V^* means $\operatorname{Hom}_{\mathbb{K}}(V,\mathbb{K})$)

such that every $v \in V$ satisfies $v = \sum\limits_{i\in I} e_i^*(v)\, e_i$. In other words, $\left(I, (e_i)_{i\in I}, \left(e_i^*\right)_{i\in I}\right)$ is a finite dual generating system for V (by the definition of a "finite dual generating system"). Thus, a finite dual generating system for V exists (namely, $\left(I, (e_i)_{i\in I}, \left(e_i^*\right)_{i\in I}\right)$). This proves Claim 2.

Now, both Claim 1 and Claim 2 are proven. Hence, the proof of Proposition 13.111.2 is complete. □

[*Remark:* The notion of finite dual generating system for V is more versatile than the notion of a finite basis of V. One difference between these notions is that all bases of V have the same size, while the set I in a finite dual generating system $\left(I, (a_i)_{i\in I}, (f_i)_{i\in I}\right)$ of V can have any (finite) size $\geq \dim V$. Another difference manifests itself in the general setting when $\mathbb{K}$ is a commutative ring, not necessarily a field. In this generality, a finite basis of V exists if and only if V is a finite free $\mathbb{K}$-module (this is the definition of a finite free $\mathbb{K}$-module), whereas a finite dual generating system for V exists if and only if V is a finitely generated projective $\mathbb{K}$-module. Projective $\mathbb{K}$-modules are a more frequent occurrence in commutative algebra than free $\mathbb{K}$-modules, and in the absence of a finite basis, a finite dual generating system is the thing that comes closest to allowing "computing in a basis".]

One significant application of finite dual generating systems is computing traces of endomorphisms:

Proposition 13.111.3. *Let $\mathbb{K}$ be a field. Let V be a finite-dimensional $\mathbb{K}$-vector space. Let $\left(I, (a_i)_{i\in I}, (f_i)_{i\in I}\right)$ be a finite dual generating system for V. Let $T : V \to V$ be a $\mathbb{K}$-linear map. Then,*

$$\operatorname{trace} T = \sum_{i\in I} f_i(Ta_i).$$

Proposition 13.111.3 can be easily proven directly, but let us take a slight detour and derive it from the following more general fact:

Proposition 13.111.4. *Let $\mathbb{K}$ be a commutative ring. Let V be a $\mathbb{K}$-module. Let $\left(I, (a_i)_{i\in I}, (f_i)_{i\in I}\right)$ be a finite dual generating system for V. Let $\left(J, (b_j)_{j\in J}, (g_j)_{j\in J}\right)$ be a further finite dual generating system for V. Let $T : V \to V$ be a $\mathbb{K}$-linear map. Then,*

$$\sum_{i\in I} f_i(Ta_i) = \sum_{j\in J} g_j(Tb_j).$$

Proof of Proposition 13.111.4. We know that $\left(I, (a_i)_{i\in I}, (f_i)_{i\in I}\right)$ is a finite dual generating system for V. In other words, $\left(I, (a_i)_{i\in I}, (f_i)_{i\in I}\right)$ is a triple such that

- I is a finite set;
- $(a_i)_{i\in I}$ is a family of elements of V;
- $(f_i)_{i\in I}$ is a family of elements of V^* (where V^* means $\operatorname{Hom}_{\mathbb{K}}(V,\mathbb{K})$)

such that every $v \in V$ satisfies

$$v = \sum_{i\in I} f_i(v)\, a_i. \tag{13.111.4}$$

We know that $\left(J, (b_j)_{j\in J}, (g_j)_{j\in J}\right)$ is a finite dual generating system for V. In other words, $\left(J, (b_j)_{j\in J}, (g_j)_{j\in J}\right)$ is a triple such that

- J is a finite set;
- $(b_j)_{j\in J}$ is a family of elements of V;
- $(g_j)_{j\in J}$ is a family of elements of V^* (where V^* means $\operatorname{Hom}_{\mathbb{K}}(V,\mathbb{K})$)

such that every $v \in V$ satisfies

$$v = \sum_{j\in J} g_j(v)\, b_j. \tag{13.111.5}$$

Now,

$$\sum_{i\in I} f_i \left(\underbrace{Ta_i}_{\substack{=\sum\limits_{j\in J} g_j(Ta_i) b_j \\ \text{(by (13.111.5), applied} \\ \text{to } v=Ta_i)}} \right) = \sum_{i\in I} \underbrace{f_i\left(\sum_{j\in J} g_j\left(Ta_i\right) b_j\right)}_{\substack{=\sum\limits_{j\in J} g_j(Ta_i) f_i(b_j) \\ \text{(since the map } f_i \text{ is } \mathbb{K}\text{-linear)}}} = \underbrace{\sum_{i\in I}\sum_{j\in J}}_{=\sum_{j\in J}\sum_{i\in I}} \underbrace{g_j\left(Ta_i\right) f_i\left(b_j\right)}_{=f_i(b_j) g_j(Ta_i)}$$
$$= \sum_{j\in J}\sum_{i\in I} f_i\left(b_j\right) g_j\left(Ta_i\right).$$

Compared with

$$\sum_{j\in J} g_j \left(T \underbrace{b_j}_{\substack{=\sum\limits_{i\in I} f_i(b_j) a_i \\ \text{(by (13.111.4), applied} \\ \text{to } v=b_j)}} \right) = \sum_{j\in J} g_j \left(\underbrace{T \sum_{i\in I} f_i\left(b_j\right) a_i}_{\substack{=\sum_{i\in I} f_i(b_j) Ta_i \\ \text{(since the map } T \text{ is } \mathbb{K}\text{-linear)}}} \right) = \sum_{j\in J} \underbrace{g_j\left(\sum_{i\in I} f_i\left(b_j\right) Ta_i\right)}_{\substack{=\sum\limits_{i\in I} f_i(b_j) g_j(Ta_i) \\ \text{(since the map } g_j \text{ is } \mathbb{K}\text{-linear)}}}$$
$$= \sum_{j\in J}\sum_{i\in I} f_i\left(b_j\right) g_j\left(Ta_i\right),$$

this yields $\sum_{i\in I} f_i\left(Ta_i\right) = \sum_{j\in J} g_j\left(Tb_j\right)$. This proves Proposition 13.111.4. □

Proof of Proposition 13.111.3. The vector space V has a finite basis (since it is finite-dimensional). Let $\left(e_1, e_2, \ldots, e_n\right)$ be such a basis. Let $\left(m_{i,j}\right)_{1\leq i,j\leq n} \in \mathbb{K}^{n\times n}$ be the matrix which represents the map $T : V \to V$ with respect to this basis $\left(e_1, e_2, \ldots, e_n\right)$ of V. Then,

$$Te_j = \sum_{i=1}^{n} m_{i,j} e_i \qquad \text{for every } j \in \{1, 2, \ldots, n\} \tag{13.111.6}$$

(due to the definition of "the matrix which represents the map $T : V \to V$ with respect to this basis $\left(e_1, e_2, \ldots, e_n\right)$ of V").

Let $\left(e_1^*, e_2^*, \ldots, e_n^*\right)$ be the basis of V^* dual to the basis $\left(e_1, e_2, \ldots, e_n\right)$ of V. (This is well-defined, since V is finite-dimensional.) Thus,

$$e_k^*\left(\sum_{i=1}^{n} \lambda_i e_i\right) = \lambda_k \qquad \text{for all } k \in \{1, 2, \ldots, n\} \text{ and } \left(\lambda_1, \lambda_2, \ldots, \lambda_n\right) \in \mathbb{K}^n. \tag{13.111.7}$$

(This follows immediately from the definition of a "dual basis".) Now, every $k \in \{1, 2, \ldots, n\}$ and $j \in \{1, 2, \ldots, n\}$ satisfy

$$e_k^*\left(Te_j\right) = m_{k,j}. \tag{13.111.8}$$

[799] Now, every $v \in V$ satisfies $v = \sum\limits_{i\in\{1,2,\ldots,n\}} e_i^*\left(v\right) e_i$ [800].

So we know that $\left(\{1, 2, \ldots, n\}, \left(e_i\right)_{i\in\{1,2,\ldots,n\}}, \left(e_i^*\right)_{i\in\{1,2,\ldots,n\}}\right)$ is a triple such that

[799] *Proof of (13.111.8):* Fix $k \in \{1, 2, \ldots, n\}$ and $j \in \{1, 2, \ldots, n\}$. Applying the map e_k^* to both sides of (13.111.6), we obtain

$$e_k^*\left(Te_j\right) = e_k^*\left(\sum_{i=1}^{n} m_{i,j} e_i\right) = m_{k,j} \qquad \text{(by (13.111.7), applied to } \lambda_i = m_{i,j}\text{)}.$$

This proves (13.111.8).

[800] *Proof.* Let $v \in V$. Then, we can write v in the form $v = \sum\limits_{i=1}^{n} \lambda_i e_i$ for some n-tuple $\left(\lambda_1, \lambda_2, \ldots, \lambda_n\right) \in \mathbb{K}^n$ (since $\left(e_1, e_2, \ldots, e_n\right)$ is a basis of V). Consider this n-tuple $\left(\lambda_1, \lambda_2, \ldots, \lambda_n\right)$. Now, $v = \sum\limits_{i=1}^{n} \lambda_i e_i$. Hence, for each $k \in \{1, 2, \ldots, n\}$, we

- $\{1,2,\ldots,n\}$ is a finite set;
- $(e_i)_{i\in\{1,2,\ldots,n\}}$ is a family of elements of V;
- $(e_i^*)_{i\in\{1,2,\ldots,n\}}$ is a family of elements of V^* (where V^* means $\operatorname{Hom}_{\mathbb{K}}(V,\mathbb{K})$)

such that every $v\in V$ satisfies $v=\sum\limits_{i\in\{1,2,\ldots,n\}} e_i^*(v)\,e_i$. In other words, $\left(\{1,2,\ldots,n\},(e_i)_{i\in\{1,2,\ldots,n\}},(e_i^*)_{i\in\{1,2,\ldots,n\}}\right)$ is a finite dual generating system for V. Therefore, Proposition 13.111.4 (applied to $\left(J,(b_j)_{j\in J},(g_j)_{j\in J}\right)=\left(\{1,2,\ldots,n\},(e_i)_{i\in\{1,2,\ldots,n\}},(e_i^*)_{i\in\{1,2,\ldots,n\}}\right)$) yields

$$\begin{aligned}\sum_{i\in I} f_i(Ta_i) &= \underbrace{\sum_{j\in\{1,2,\ldots,n\}}}_{=\sum_{j=1}^n} \underbrace{e_j^*(Te_j)}_{\substack{=m_{j,j}\\ \text{(by (13.111.8), applied}\\ \text{to } k=j)}} = \sum_{j=1}^n m_{j,j}\\ &= \sum_{i=1}^n m_{i,i} \qquad \text{(here, we renamed the summation index } j \text{ as } i)\,. \end{aligned}\tag{13.111.9}$$

But recall that if $\mathfrak{G}$ is any endomorphism of the $\mathbb{K}$-vector space V, then the trace of $\mathfrak{G}$ equals the trace of any matrix which represents $\mathfrak{G}$ with respect to a basis of V [801]. Applying this to $\mathfrak{G}=T$, we conclude that the trace of T equals the trace of any matrix which represents T with respect to a basis of V. In particular, the trace of T equals the trace of the matrix $(m_{i,j})_{1\leq i,j\leq n}$ (since $(m_{i,j})_{1\leq i,j\leq n}$ is a matrix which represents T with respect to the basis $(e_1,e_2,\ldots,e_n)$ of V). In other words, $\operatorname{trace} T=\operatorname{trace}\left((m_{i,j})_{1\leq i,j\leq n}\right)=\sum_{i=1}^n m_{i,i}$ (by the definition of $\operatorname{trace}\left((m_{i,j})_{1\leq i,j\leq n}\right)$). Compared with (13.111.9), this yields $\operatorname{trace} T=\sum_{i\in I} f_i(Ta_i)$. This proves Proposition 13.111.3. □

[*Remark:* Proposition 13.111.4 can be used to define the trace of an endomorphism of a finitely generated projective $\mathbb{K}$-module when $\mathbb{K}$ is a commutative ring. Indeed, if $\mathbb{K}$ is a commutative ring and if V is a finitely generated projective $\mathbb{K}$-module, and if $T:V\to V$ is a $\mathbb{K}$-linear map, then the trace $\operatorname{trace}(T)$ of T can be defined as $\sum_{i\in I} f_i(Ta_i)$, where $\left(I,(a_i)_{i\in I},(f_i)_{i\in I}\right)$ is a finite dual generating system for V. This notion of trace is well-defined[802] and generalizes the classical notion from linear algebra (which is defined only for finitely generated **free** $\mathbb{K}$-modules)[803]. But we will not concern ourselves with these generalizations, since our exercise deals only with representations of groups over a field.]

After all these preparations, we finally come to the actual solution of Exercise 4.1.14(b). Let U be any finite-dimensional $\mathbb{C}H$-module. We want to prove $\chi_{\operatorname{Ind}_\rho U}=\operatorname{Ind}_\rho \chi_U$.

For every $g\in G$, we define a $\mathbb{C}$-linear map $g^*:\mathbb{C}G\to\mathbb{C}$ by

$$\left(g^*(k)=\delta_{g,k} \qquad \text{for all } k\in G\right). \tag{13.111.10}$$

(This is well-defined, since $(k)_{k\in G}$ is a basis of the $\mathbb{C}$-vector space $\mathbb{C}G$.) Then,

$$\sum_{g\in G} g^*(\gamma)\cdot g=\gamma \qquad \text{for every } \gamma\in\mathbb{C}G. \tag{13.111.11}$$

have

$$e_k^*\left(\underbrace{v}_{=\sum\limits_{i=1}^n \lambda_i e_i}\right)=e_k^*\left(\sum_{i=1}^n \lambda_i e_i\right)=\lambda_k \qquad \text{(by (13.111.7))}\,.$$

Renaming k as i in this statement, we obtain the following: For each $i\in\{1,2,\ldots,n\}$, we have $e_i^*(v)=\lambda_i$. Thus, $\underbrace{\sum\limits_{i\in\{1,2,\ldots,n\}}}_{=\sum_{i=1}^n} \underbrace{e_i^*(v)}_{=\lambda_i} e_i=\sum\limits_{i=1}^n \lambda_i e_i$. Compared with $v=\sum\limits_{i=1}^n \lambda_i e_i$, this yields $v=\sum\limits_{i\in\{1,2,\ldots,n\}} e_i^*(v)\,e_i$, qed.

[801] Indeed, this is how the trace of $\mathfrak{G}$ is defined.

[802] because Proposition 13.111.4 shows that $\sum_{i\in I} f_i(Ta_i)$ does not depend on the choice of $\left(I,(a_i)_{i\in I},(f_i)_{i\in I}\right)$

[803] because Proposition 13.111.3 (or, more precisely, its straightforward generalization to free $\mathbb{K}$-modules over commutative rings) shows that these two notions give the same result when V is a free $\mathbb{K}$-module

[804] Also,

$$(gr)^{*}(\gamma r) = g^{*}(\gamma) \qquad \text{for all } g \in G,\ r \in G \text{ and } \gamma \in \mathbb{C}G \tag{13.111.12}$$

[805].

Proposition 13.111.2 (applied to $\mathbb{K} = \mathbb{C}$ and $V = U$) yields that a finite dual generating system for U exists if and only if the vector space U is finite-dimensional. Thus, a finite dual generating system for U exists (since the vector space U is finite-dimensional). Let us fix such a finite dual generating system for U, and denote it by $\left(J, (b_j)_{j\in J}, (g_j)_{j\in J}\right)$.

Hence, $\left(J, (b_j)_{j\in J}, (g_j)_{j\in J}\right)$ is a finite dual generating system for U. In other words, $\left(J, (b_j)_{j\in J}, (g_j)_{j\in J}\right)$ is a triple such that

- J is a finite set;
- $(b_j)_{j\in J}$ is a family of elements of U;
- $(g_j)_{j\in J}$ is a family of elements of U^{*} (where U^{*} means $\operatorname{Hom}_{\mathbb{C}}(U, \mathbb{C})$)

such that every $v \in U$ satisfies

$$v = \sum_{j\in J} g_j(v)\, b_j. \tag{13.111.13}$$

We now endow $\mathbb{C}G$ with the $(\mathbb{C}G, \mathbb{C}H)$-bimodule structure that was used to define $\operatorname{Ind}_\rho U$. We recall that $\operatorname{Ind}_\rho U = \mathbb{C}G \otimes_{\mathbb{C}H} U$ (by the definition of $\operatorname{Ind}_\rho U$).

Let us fix $g \in G$. We define a map $\widetilde{F}_g : \mathbb{C}G \times U \to U$ by setting

$$\left(\widetilde{F}_g(\gamma, u) = \frac{1}{|H|} \sum_{h\in H} (g\rho(h))^{*}(\gamma)\, hu \qquad \text{for all } (\gamma, u) \in \mathbb{C}G \times U\right).$$

The map $\widetilde{F}_g$ is $\mathbb{C}H$-bilinear with respect to the right $\mathbb{C}H$-module structure on $\mathbb{C}G$ and the left $\mathbb{C}H$-module structure on U [806]. According to the universal property of the tensor product, this yields that there exists

[804] *Proof of (13.111.11):* Let $\gamma \in \mathbb{C}G$. We need to prove the equality (13.111.11). We notice that this equality is $\mathbb{C}$-linear in γ. Hence, we can WLOG assume that $\gamma \in G$ (since G is a basis of the $\mathbb{C}$-vector space $\mathbb{C}G$). Assume this. Now,

$$\begin{aligned}\sum_{g\in G} g^{*}(\gamma)\cdot g &= \sum_{g\in G} \underbrace{g^{*}(\gamma)}_{\substack{=\delta_{g,\gamma} \\ \text{(by (13.111.10), applied} \\ \text{to } k=\gamma \text{ (since } \gamma\in G))}} \cdot g = \sum_{g\in G} \delta_{g,\gamma}\cdot g \\ &= \sum_{\substack{g\in G;\\ g=\gamma}} \underbrace{\delta_{g,\gamma}}_{\substack{=1\\ \text{(since } g=\gamma)}} \cdot g + \sum_{\substack{g\in G;\\ g\neq\gamma}} \underbrace{\delta_{g,\gamma}}_{\substack{=0\\ \text{(since } g\neq\gamma)}} \cdot g = \sum_{\substack{g\in G;\\ g=\gamma}} 1\cdot g + \underbrace{\sum_{\substack{g\in G;\\ g\neq\gamma}} 0\cdot g}_{=0} \\ &= \sum_{\substack{g\in G;\\ g=\gamma}} \underbrace{1\cdot g}_{=g} = \sum_{\substack{g\in G;\\ g=\gamma}} g = \gamma \qquad (\text{since } \gamma \in G).\end{aligned}$$

This proves (13.111.11).

[805] *Proof of (13.111.12):* Let $g \in G$, $r \in G$ and $\gamma \in \mathbb{C}G$. We need to prove the equality (13.111.12). Since this equality is $\mathbb{C}$-linear in γ (because the maps $(gr)^{*}$ and g^{*} are $\mathbb{C}$-linear), we can WLOG assume that $\gamma \in G$ (since G is a basis of the $\mathbb{C}$-vector space $\mathbb{C}G$). Assume this.

We have $\gamma \in G$, and thus we can apply (13.111.10) to $k = \gamma$. We thus obtain $g^{*}(\gamma) = \delta_{g,\gamma}$. Also, $\underbrace{\gamma}_{\in G}\underbrace{r}_{\in G} \in GG \subset G$. Hence, (13.111.10) (applied to gr and γr instead of g and k) yields $(gr)^{*}(\gamma r) = \delta_{gr,\gamma r}$.

Now, we have $g = \gamma$ if and only if $gr = \gamma r$ (because G is a group). Now,

$$\begin{aligned} g^{*}(\gamma) = \delta_{g,\gamma} &= \begin{cases} 1, & \text{if } g = \gamma; \\ 0, & \text{if } g \neq \gamma \end{cases} = \begin{cases} 1, & \text{if } gr = \gamma r; \\ 0, & \text{if } gr \neq \gamma r \end{cases} \qquad (\text{since } g = \gamma \text{ if and only if } gr = \gamma r) \\ &= \delta_{gr,\gamma r} = (gr)^{*}(\gamma r).\end{aligned}$$

This proves (13.111.12).

[806] *Proof.* We notice that $\widetilde{F}_g(\gamma, u) = \frac{1}{|H|} \sum_{h\in H} (g\rho(h))^{*}(\gamma)\, hu$ depends $\mathbb{C}$-linearly on each of γ and u (for obvious reasons). In other words, the map $\widetilde{F}_g$ is $\mathbb{C}$-bilinear.

Now, let us fix $\gamma \in \mathbb{C}G$, $u \in U$ and $\kappa \in \mathbb{C}H$. We are going to prove the equality $\widetilde{F}_g(\gamma\kappa, u) = \widetilde{F}_g(\gamma, \kappa u)$.

a unique $\mathbb{C}$-linear map $F_g : \mathbb{C}G \otimes_{\mathbb{C}H} U \to U$ such that

$$\left(F_g\left(\gamma \otimes_{\mathbb{C}H} u\right) = \widetilde{F}_g\left(\gamma, u\right) \qquad \text{for all } \left(\gamma, u\right) \in \mathbb{C}G \times U\right).$$

Consider this map F_g.

So we know that any $(\gamma, u) \in \mathbb{C}G \times U$ satisfies

$$F_g\left(\gamma \otimes_{\mathbb{C}H} u\right) = \widetilde{F}_g\left(\gamma, u\right) = \frac{1}{|H|} \sum_{h \in H} \left(g\rho\left(h\right)\right)^* \left(\gamma\right) hu. \tag{13.111.14}$$

Let us now forget that we fixed g. We thus have constructed a $\mathbb{C}$-linear map $F_g : \mathbb{C}G \otimes_{\mathbb{C}H} U \to U$ for each $g \in G$. We have shown that this map satisfies (13.111.14) for any $(\gamma, u) \in \mathbb{C}G \times U$.

The set $G \times J$ is finite (since the sets G and J are finite). Now, we define a family $(a_i)_{i \in G \times J}$ of elements of $\mathbb{C}G \otimes_{\mathbb{C}H} U$ by

$$\left(a_{(k,j)} = k \otimes_{\mathbb{C}H} b_j \qquad \text{for all } (k, j) \in G \times J\right).$$

Furthermore, we define a family $(f_i)_{i \in G \times J}$ of elements of $(\mathbb{C}G \otimes_{\mathbb{C}H} U)^*$ by

$$\left(f_{(k,j)} = g_j \circ F_k \qquad \text{for all } (k, j) \in G \times J\right).$$

(This is well-defined because, for any $(k, j) \in G \times J$, the composition $g_j \circ F_k$ of the $\mathbb{C}$-linear maps $F_k : \mathbb{C}G \otimes_{\mathbb{C}H} U \to U$ and $g_j : U \to \mathbb{C}$ is a $\mathbb{C}$-linear map $\mathbb{C}G \otimes_{\mathbb{C}H} U \to \mathbb{C}$.)

Our goal is now to prove that $\left(G \times J, (a_i)_{i \in G \times J}, (f_i)_{i \in G \times J}\right)$ is a finite dual generating system for $\mathbb{C}G \otimes_{\mathbb{C}H} U$.

Since this equality is $\mathbb{C}$-linear in κ (because the map $\widetilde{F}_g$ is $\mathbb{C}$-bilinear), we can WLOG assume that $\kappa \in H$ (since H is a basis of the $\mathbb{C}$-vector space $\mathbb{C}H$). Assume this.

We have $\kappa \in H$. Thus, the map $H \to H,\ h \mapsto h\kappa$ is a bijection (since H is a group). Hence, we can substitute $h\kappa$ for h in the sum $\sum_{h \in H} (g\rho(h))^* (\gamma\rho(\kappa))\, hu$. As a result, we obtain $\sum_{h \in H} (g\rho(h))^* (\gamma\rho(\kappa))\, hu = \sum_{h \in H} (g\rho(h\kappa))^* (\gamma\rho(\kappa))\,(h\kappa)\, u$.

Now, the definition of $\widetilde{F}_g(\gamma\kappa, u)$ yields

$$\begin{aligned}\widetilde{F}_g\left(\gamma\kappa, u\right) &= \frac{1}{|H|} \underbrace{\sum_{h \in H} \left(g\rho\left(h\right)\right)^* \left(\gamma\rho\left(\kappa\right)\right) hu}_{= \sum_{h \in H} (g\rho(h\kappa))^* (\gamma\rho(\kappa))(h\kappa)u} = \frac{1}{|H|} \sum_{h \in H} \left(g \underbrace{\rho\left(h\kappa\right)}_{\substack{= \rho(h)\rho(\kappa) \\ \text{(since } \rho \text{ is a group} \\ \text{homomorphism)}}} \right)^* \left(\gamma\rho\left(\kappa\right)\right) \left(h\kappa\right) u \\ &= \frac{1}{|H|} \sum_{h \in H} \underbrace{\left(g\rho\left(h\right)\rho\left(\kappa\right)\right)^* \left(\gamma\rho\left(\kappa\right)\right)}_{\substack{= (g\rho(h))^*(\gamma) \\ \text{(by (13.111.12), applied to} \\ g\rho(h) \text{ and } \rho(\kappa) \text{ instead of } g \text{ and } r)}} h\kappa u = \frac{1}{|H|} \sum_{h \in H} \left(g\rho\left(h\right)\right)^* \left(\gamma\right) h\kappa u.\end{aligned}$$

Compared with

$$\widetilde{F}_g\left(\gamma, \kappa u\right) = \frac{1}{|H|} \sum_{h \in H} \left(g\rho\left(h\right)\right)^* \left(\gamma\right) h\kappa u \qquad \left(\text{by the definition of } \widetilde{F}_g\left(\gamma, \kappa u\right)\right),$$

this yields $\widetilde{F}_g(\gamma\kappa, u) = \widetilde{F}_g(\gamma, \kappa u)$.

Now let us forget that we fixed γ, u and κ. Thus, we have shown that $\widetilde{F}_g(\gamma\kappa, u) = \widetilde{F}_g(\gamma, \kappa u)$ for all $\gamma \in \mathbb{C}G$, $u \in U$ and $\kappa \in \mathbb{C}H$. This yields that the map $\widetilde{F}_g$ is $\mathbb{C}H$-bilinear with respect to the right $\mathbb{C}H$-module structure on $\mathbb{C}G$ and the left $\mathbb{C}H$-module structure on U (since we already know that this map $\widetilde{F}_g$ is $\mathbb{C}$-bilinear). Qed.

Indeed, every $v \in \mathbb{C}G \otimes_{\mathbb{C}H} U$ satisfies $v = \sum_{i \in G \times J} f_i(v) a_i$ [807]. Thus, we know that $\left(G \times J, (a_i)_{i \in G \times J}, (f_i)_{i \in G \times J} \right)$ is a triple such that

- $G \times J$ is a finite set;
- $(a_i)_{i \in G \times J}$ is a family of elements of $\mathbb{C}G \otimes_{\mathbb{C}H} U$;
- $(f_i)_{i \in G \times J}$ is a family of elements of $(\mathbb{C}G \otimes_{\mathbb{C}H} U)^*$ (where $(\mathbb{C}G \otimes_{\mathbb{C}H} U)^*$ means $\operatorname{Hom}_{\mathbb{C}} (\mathbb{C}G \otimes_{\mathbb{C}H} U, \mathbb{C})$)

[807] *Proof.* Let $v \in \mathbb{C}G \otimes_{\mathbb{C}H} U$. We need to prove the equality $v = \sum_{i \in G \times J} f_i(v) a_i$. Since this equality is $\mathbb{C}$-linear in v, we can WLOG assume that v is a pure tensor (since the pure tensors in $\mathbb{C}G \otimes_{\mathbb{C}H} U$ span the whole $\mathbb{C}$-vector space $\mathbb{C}G \otimes_{\mathbb{C}H} U$). Assume this. Thus, v is a pure tensor. In other words, $v = \gamma \otimes_{\mathbb{C}H} u$ for some $\gamma \in \mathbb{C}G$ and $u \in U$. Consider these γ and u.

We notice that for any $w \in G$, the map $G \to G,\ k \mapsto kw$ is a bijection (since G is a group). Hence, for any $w \in G$, we have

$$\begin{aligned}
\sum_{k \in G} (kw)^* (\gamma)\, kw &= \sum_{k \in G} k^* (\gamma)\, k && \left(\begin{array}{c} \text{here, we substituted } k \text{ for } kw \text{ in the sum,} \\ \text{since the map } G \to G,\ k \mapsto kw \text{ is a bijection} \end{array} \right) \\
&= \sum_{g \in G} g^* (\gamma)\, g && (\text{here, we renamed the summation index } k \text{ as } g) \\
&= \sum_{g \in G} g^* (\gamma) \cdot g = \gamma && (\text{by } (13.111.11)) . \qquad (13.111.15)
\end{aligned}$$

Now, every $h \in H$ satisfies

$$\begin{aligned}
\sum_{k \in G} (k\rho(h))^* (\gamma) \underbrace{kh}_{\substack{= k \cdot (\mathbb{C}[\rho])(h) \\ \text{(by the definition of the} \\ \text{right } \mathbb{C}H\text{-module} \\ \text{structure on } \mathbb{C}G)}} &= \sum_{k \in G} (k\rho(h))^* (\gamma)\, k \cdot \underbrace{(\mathbb{C}[\rho])(h)}_{\substack{= \rho(h) \\ (\text{since } h \in H)}} = \sum_{k \in G} (k\rho(h))^* (\gamma)\, k\rho(h) \\
&= \gamma \qquad (\text{by } (13.111.15), \text{ applied to } w = \rho(h)) . \qquad (13.111.16)
\end{aligned}$$

Now,

$$\begin{aligned}
\sum_{i \in G \times J} f_i(v) a_i &= \underbrace{\sum_{(k,j) \in G \times J}}_{= \sum_{k \in G} \sum_{j \in J}} \underbrace{f_{(k,j)}}_{= g_j \circ F_k} (v) \underbrace{a_{(k,j)}}_{= k \otimes_{\mathbb{C}H} b_j} \qquad (\text{here, we substituted } (k,j) \text{ for the summation index } i) \\
&= \sum_{k \in G} \underbrace{\sum_{j \in J} (g_j \circ F_k)(v)\, k \otimes_{\mathbb{C}H} b_j}_{\substack{= k \otimes_{\mathbb{C}H} \left(\sum_{j \in J} (g_j \circ F_k)(v) b_j \right) \\ \text{(since the tensor product is } \mathbb{C}\text{-bilinear)}}} = \sum_{k \in G} k \otimes_{\mathbb{C}H} \left(\sum_{j \in J} \underbrace{(g_j \circ F_k)(v)}_{= g_j(F_k(v))} b_j \right) \\
&= \sum_{k \in G} k \otimes_{\mathbb{C}H} \underbrace{\left(\sum_{j \in J} g_j (F_k(v))\, b_j \right)}_{\substack{= F_k(v) \\ \text{(because } F_k(v) = \sum_{j \in J} g_j(F_k(v)) b_j \\ \text{(by (13.111.13), applied to } F_k(v) \text{ instead of } v))}} = \sum_{k \in G} k \otimes_{\mathbb{C}H} F_k \left(\underbrace{v}_{= \gamma \otimes_{\mathbb{C}H} u} \right) \\
&= \sum_{k \in G} k \otimes_{\mathbb{C}H} \underbrace{F_k (\gamma \otimes_{\mathbb{C}H} u)}_{\substack{= \frac{1}{|H|} \sum_{h \in H} (k\rho(h))^* (\gamma) hu \\ \text{(by (13.111.14), applied to } g = k)}} = \sum_{k \in G} k \otimes_{\mathbb{C}H} \left(\frac{1}{|H|} \sum_{h \in H} (k\rho(h))^* (\gamma)\, hu \right) \\
&= \frac{1}{|H|} \underbrace{\sum_{k \in G} \sum_{h \in H}}_{= \sum_{h \in H} \sum_{k \in G}} (k\rho(h))^* (\gamma)\, k \otimes_{\mathbb{C}H} hu \qquad (\text{since the tensor product is } \mathbb{C}\text{-bilinear}) \\
&= \frac{1}{|H|} \sum_{h \in H} \sum_{k \in G} (k\rho(h))^* (\gamma) \underbrace{k \otimes_{\mathbb{C}H} hu}_{\substack{= kh \otimes_{\mathbb{C}H} u \\ \text{(since } h \text{ can be moved past} \\ \text{the } \otimes_{\mathbb{C}H} \text{ sign (since } h \in H \subset \mathbb{C}H))}} = \frac{1}{|H|} \sum_{h \in H} \sum_{k \in G} (k\rho(h))^* (\gamma)\, kh \otimes_{\mathbb{C}H} u \\
&= \frac{1}{|H|} \sum_{h \in H} \underbrace{\left(\sum_{k \in G} (k\rho(h))^* (\gamma)\, kh \right)}_{\substack{= \gamma \\ \text{(by (13.111.16))}}} \otimes_{\mathbb{C}H} u = \frac{1}{|H|} \sum_{h \in H} \underbrace{\gamma \otimes_{\mathbb{C}H} u}_{= v} = \frac{1}{|H|} \underbrace{\sum_{h \in H} v}_{= |H| \cdot v} = \frac{1}{|H|} |H| \cdot v = v.
\end{aligned}$$

such that every $v \in \mathbb{C}G \otimes_{\mathbb{C}H} U$ satisfies $v = \sum_{i \in G \times J} f_i(v)\, a_i$. In other words, $\left(G \times J, (a_i)_{i \in G \times J}, (f_i)_{i \in G \times J}\right)$ is a finite dual generating system for $\mathbb{C}G \otimes_{\mathbb{C}H} U$ (because this is precisely how a "finite dual generating system" was defined). In other words, $\left(G \times J, (a_i)_{i \in G \times J}, (f_i)_{i \in G \times J}\right)$ is a finite dual generating system for $\operatorname{Ind}_\rho U$ (since $\operatorname{Ind}_\rho U = \mathbb{C}G \otimes_{\mathbb{C}H} U$).

Let us now notice that

$$\sum_{j \in J} g_j(hb_j) = \chi_U(h) \qquad \text{for every } h \in H \tag{13.111.17}$$

[808].

Now, let us fix $g \in G$. We want to compute $\chi_{\operatorname{Ind}_\rho U}(g)$. The definition of $\chi_{\operatorname{Ind}_\rho U}(g)$ yields $\chi_{\operatorname{Ind}_\rho U}(g) = \operatorname{trace}(g : \operatorname{Ind}_\rho U \to \operatorname{Ind}_\rho U)$. But Proposition 13.111.3 (applied to $\mathbb{K} = \mathbb{C}$, $V = \operatorname{Ind}_\rho U$, $I = G \times J$ and $T = (g : \operatorname{Ind}_\rho U \to \operatorname{Ind}_\rho U)$) yields

$$\begin{aligned}
&\operatorname{trace}(g : \operatorname{Ind}_\rho U \to \operatorname{Ind}_\rho U) \\
&= \sum_{i \in G \times J} f_i \left(\underbrace{(g : \operatorname{Ind}_\rho U \to \operatorname{Ind}_\rho U)\, a_i}_{= g a_i} \right) = \sum_{i \in G \times J} f_i(g a_i) \\
&= \sum_{(k,j) \in G \times J} \underbrace{f_{(k,j)}}_{= g_j \circ F_k} \left(g \underbrace{a_{(k,j)}}_{= k \otimes_{\mathbb{C}H} b_j} \right) \qquad \text{(here, we substituted } (k,j) \text{ for the summation index } i\text{)} \\
&= \sum_{(k,j) \in G \times J} (g_j \circ F_k) \underbrace{(g(k \otimes_{\mathbb{C}H} b_j))}_{= gk \otimes_{\mathbb{C}H} b_j} = \sum_{(k,j) \in G \times J} \underbrace{(g_j \circ F_k)(gk \otimes_{\mathbb{C}H} b_j)}_{= g_j(F_k(gk \otimes_{\mathbb{C}H} b_j))} \\
&= \sum_{(k,j) \in G \times J} g_j \left(\underbrace{F_k(gk \otimes_{\mathbb{C}H} b_j)}_{\substack{= \frac{1}{|H|} \sum_{h \in H} (k\rho(h))^* (gk) h b_j \\ \text{(by (13.111.14), applied to } k,\ gk \text{ and } b_j \\ \text{instead of } g,\ \gamma \text{ and } u)}} \right) \\
&= \underbrace{\sum_{(k,j) \in G \times J}}_{= \sum_{k \in G} \sum_{j \in J}} \underbrace{g_j \left(\frac{1}{|H|} \sum_{h \in H} (k\rho(h))^* (gk)\, h b_j \right)}_{\substack{= \frac{1}{|H|} \sum_{h \in H} (k\rho(h))^* (gk) g_j(hb_j) \\ \text{(since the map } g_j \text{ is } \mathbb{C}\text{-linear)}}} = \sum_{k \in G} \sum_{j \in J} \frac{1}{|H|} \sum_{h \in H} (k\rho(h))^* (gk)\, g_j(hb_j) \\
&= \sum_{k \in G} \frac{1}{|H|} \sum_{h \in H} (k\rho(h))^* (gk) \sum_{j \in J} g_j(hb_j).
\end{aligned} \tag{13.111.18}$$

Thus, $v = \sum_{i \in G \times J} f_i(v)\, a_i$ is proven, qed.

[808] *Proof of (13.111.17):* Let $h \in H$. The definition of $\chi_U(h)$ yields $\chi_U(h) = \operatorname{trace}(h : U \to U)$. But Proposition 13.111.3 (applied to $\mathbb{C}$, U, $\left(J, (b_j)_{j \in J}, (g_j)_{j \in J}\right)$ and $(h : U \to U)$ instead of $\mathbb{C}$, U, $\left(I, (a_i)_{i \in I}, (f_i)_{i \in I}\right)$ and T) yields

$$\operatorname{trace}(h : U \to U) = \sum_{j \in J} g_j \left(\underbrace{(h : U \to U)\, b_j}_{= h b_j} \right) = \sum_{j \in J} g_j(hb_j).$$

Thus, $\chi_U(h) = \operatorname{trace}(h : U \to U) = \sum_{j \in J} g_j(hb_j)$. This proves (13.111.17).

But every $k \in G$ satisfies

$$\begin{aligned}
&\sum_{h\in H} \underbrace{(k\rho(h))^*(gk)}_{\substack{=\delta_{k\rho(h),gk} \\ \text{(by (13.111.10), applied to } k\rho(h) \\ \text{and } gk \text{ instead of } g \text{ and } k\text{)}}} \underbrace{\sum_{j\in J} g_j(hb_j)}_{\substack{=\chi_U(h) \\ \text{(by (13.111.17))}}} \\
&= \sum_{h\in H} \delta_{k\rho(h),gk}\chi_U(h) \\
&= \underbrace{\sum_{\substack{h\in H;\\ k\rho(h)=gk}}}_{\substack{=\sum_{\substack{h\in H;\\ k\rho(h)k^{-1}=g}} \\ \text{(because for every } h\in H\text{,} \\ \text{the statement } (k\rho(h)=gk) \\ \text{is equivalent} \\ \text{to } (k\rho(h)k^{-1}=g)\text{)}}} \underbrace{\delta_{k\rho(h),gk}}_{\substack{=1 \\ \text{(since } k\rho(h)=gk\text{)}}} \chi_U(h) + \sum_{\substack{h\in H;\\ k\rho(h)\neq gk}} \underbrace{\delta_{k\rho(h),gk}}_{\substack{=0 \\ \text{(since } k\rho(h)\neq gk\text{)}}} \chi_U(h) \\
&= \sum_{\substack{h\in H;\\ k\rho(h)k^{-1}=g}} 1\chi_U(h) + \underbrace{\sum_{\substack{h\in H;\\ k\rho(h)\neq gk}} 0\chi_U(h)}_{=0} = \sum_{\substack{h\in H;\\ k\rho(h)k^{-1}=g}} 1\chi_U(h) \\
&= \sum_{\substack{h\in H;\\ k\rho(h)k^{-1}=g}} \chi_U(h).
\end{aligned} \tag{13.111.19}$$

Now, recall that

$$\begin{aligned}
\chi_{\operatorname{Ind}_\rho U}(g) &= \operatorname{trace}(g : \operatorname{Ind}_\rho U \to \operatorname{Ind}_\rho U) \\
&= \sum_{k\in G} \frac{1}{|H|} \underbrace{\sum_{h\in H} (k\rho(h))^*(gk) \sum_{j\in J} g_j(hb_j)}_{\substack{=\sum_{\substack{h\in H;\\ k\rho(h)k^{-1}=g}} \chi_U(h) \\ \text{(by (13.111.19))}}} \qquad \text{(by (13.111.18))} \\
&= \sum_{k\in G} \frac{1}{|H|} \sum_{\substack{h\in H;\\ k\rho(h)k^{-1}=g}} \chi_U(h) = \frac{1}{|H|} \underbrace{\sum_{\substack{h\in H;\\ k\rho(h)k^{-1}=g}} \sum_{k\in G}}_{=\sum_{\substack{(h,k)\in H\times G;\\ k\rho(h)k^{-1}=g}}} \chi_U(h) = \frac{1}{|H|} \sum_{\substack{(h,k)\in H\times G;\\ k\rho(h)k^{-1}=g}} \chi_U(h).
\end{aligned}$$

Compared with

$$(\operatorname{Ind}_\rho \chi_U)(g) = \frac{1}{|H|} \sum_{\substack{(h,k)\in H\times G;\\ k\rho(h)k^{-1}=g}} \chi_U(h) \qquad \text{(by the definition of } (\operatorname{Ind}_\rho \chi_U)(g)\text{)},$$

this yields $\chi_{\operatorname{Ind}_\rho U}(g) = (\operatorname{Ind}_\rho \chi_U)(g)$.

Now, let us forget that we fixed g. We thus have proven that $\chi_{\operatorname{Ind}_\rho U}(g) = (\operatorname{Ind}_\rho \chi_U)(g)$ for every $g \in G$. In other words, $\chi_{\operatorname{Ind}_\rho U} = \operatorname{Ind}_\rho \chi_U$. This solves Exercise 4.1.14(b).

(c) Assume that H is a subgroup of G, and that $\rho : H \to G$ is the inclusion map. We need to show that $\operatorname{Ind}_\rho f = \operatorname{Ind}_H^G f$ for every $f \in R_{\mathbb{C}}(H)$.

We notice that the map $G \to G,\ k \mapsto k^{-1}$ is a bijection (since G is a group).

Let $f \in R_{\mathbb{C}}(H)$. Let $g \in G$. Then, the definition of $(\operatorname{Ind}_\rho f)(g)$ yields

$$\begin{aligned}
(\operatorname{Ind}_\rho f)(g) &= \frac{1}{|H|} \underbrace{\sum_{\substack{(h,k)\in H\times G;\\ k\rho(h)k^{-1}=g}}}_{\substack{=\sum_{\substack{(h,k)\in H\times G;\\ khk^{-1}=g}}\\ \text{(since } \rho(h)=h \text{ for every } h\in H\\ \text{(since } \rho:H\to G \text{ is the inclusion map))}}} f(h) = \frac{1}{|H|} \underbrace{\sum_{\substack{(h,k)\in H\times G;\\ khk^{-1}=g}}}_{=\sum_{k\in G}\sum_{\substack{h\in H;\\ khk^{-1}=g}}} f(h) \\
&= \frac{1}{|H|} \sum_{k\in G} \underbrace{\sum_{\substack{h\in H;\\ khk^{-1}=g}}}_{\substack{=\sum_{\substack{h\in H;\\ h=k^{-1}gk}}\\ \text{(because for every } h\in H,\\ \text{the statement } \left(khk^{-1}=g\right) \text{ is}\\ \text{equivalent to } \left(h=k^{-1}gk\right))}} f(h) \\
&= \frac{1}{|H|} \underbrace{\sum_{k\in G} \sum_{\substack{h\in H;\\ h=k^{-1}gk}} f(h)}_{\substack{=\sum_{\substack{k\in G;\\ k^{-1}gk\in H}} \sum_{\substack{h\in H;\\ h=k^{-1}gk}} f(h) + \sum_{\substack{k\in G;\\ k^{-1}gk\notin H}} \sum_{\substack{h\in H;\\ h=k^{-1}gk}} f(h)\\ \text{(since every } k\in G \text{ satisfies either } k^{-1}gk\in H \text{ or } k^{-1}gk\notin H \text{ (but never both))}}} \\
&= \frac{1}{|H|} \left(\sum_{\substack{k\in G;\\ k^{-1}gk\in H}} \underbrace{\sum_{\substack{h\in H;\\ h=k^{-1}gk}} f(h)}_{\substack{=f\left(k^{-1}gk\right)\\ \text{(since } k^{-1}gk\in H)}} + \sum_{\substack{k\in G;\\ k^{-1}gk\notin H}} \underbrace{\sum_{\substack{h\in H;\\ h=k^{-1}gk}} f(h)}_{\substack{=\text{(empty sum)}\\ \text{(since there is no } h\in H \text{ satisfying } h=k^{-1}gk\\ \text{(because } k^{-1}gk\notin H))}} \right) \\
&= \frac{1}{|H|} \left(\sum_{\substack{k\in G;\\ k^{-1}gk\in H}} f\left(k^{-1}gk\right) + \sum_{\substack{k\in G;\\ k^{-1}gk\notin H}} \underbrace{\text{(empty sum)}}_{=0} \right) = \frac{1}{|H|} \left(\sum_{\substack{k\in G;\\ k^{-1}gk\in H}} f\left(k^{-1}gk\right) + \underbrace{\sum_{\substack{k\in G;\\ k^{-1}gk\notin H}} 0}_{=0} \right) \\
&= \frac{1}{|H|} \sum_{\substack{k\in G;\\ k^{-1}gk\in H}} f\left(k^{-1}gk\right) = \frac{1}{|H|} \underbrace{\sum_{\substack{k\in G;\\ \left(k^{-1}\right)^{-1}gk^{-1}\in H}}}_{\substack{=\sum_{\substack{k\in G;\\ kgk^{-1}\in H}}\\ \text{(since } \left(k^{-1}\right)^{-1}=k \text{ for every } k\in G)}} f\Bigg(\underbrace{\left(k^{-1}\right)^{-1}}_{=k} gk^{-1} \Bigg) \\
&\qquad \left(\begin{array}{c} \text{here, we substituted } k^{-1} \text{ for } k \text{ in the sum, since} \\ \text{the map } G\to G,\ k\mapsto k^{-1} \text{ is a bijection} \end{array} \right) \\
&= \frac{1}{|H|} \sum_{\substack{k\in G;\\ kgk^{-1}\in H}} f\left(kgk^{-1}\right) = \left(\operatorname{Ind}_H^G f\right)(g)
\end{aligned}$$

(since (4.1.4) yields $\left(\operatorname{Ind}_H^G f\right)(g) = \dfrac{1}{|H|} \sum\limits_{\substack{k \in G; \\ kgk^{-1} \in H}} f\left(kgk^{-1}\right)$).

Let us now forget that we fixed g. We thus have proven that $\left(\operatorname{Ind}_\rho f\right)(g) = \left(\operatorname{Ind}_H^G f\right)(g)$ for every $g \in G$. In other words, $\operatorname{Ind}_\rho f = \operatorname{Ind}_H^G f$. This solves Exercise 4.1.14(c).

(d) Assume that H is a subgroup of G, and that $\rho : H \to G$ is the inclusion map. We need to show that $\operatorname{Ind}_\rho U = \operatorname{Ind}_H^G U$ for every $\mathbb{C}H$-module U.

We notice that $\rho : H \to G$ is the inclusion map. Hence, $\mathbb{C}[\rho] : \mathbb{C}H \to \mathbb{C}G$ is also the inclusion map (since we identify $\mathbb{C}H$ with a $\mathbb{C}$-subalgebra of $\mathbb{C}G$ along this map $\mathbb{C}[\rho]$). Thus,

$$(\mathbb{C}[\rho])\,\eta = \eta \qquad \text{for every } \eta \in \mathbb{C}H. \tag{13.111.20}$$

Let U be a $\mathbb{C}H$-module. Both $\operatorname{Ind}_H^G U$ and $\operatorname{Ind}_\rho U$ are defined as $\mathbb{C}G \otimes_{\mathbb{C}H} U$ for some $(\mathbb{C}G, \mathbb{C}H)$-bimodule structure on $\mathbb{C}G$; however, their definitions differ at how this $(\mathbb{C}G, \mathbb{C}H)$-bimodule structure is defined. We are now going to prove that these two $(\mathbb{C}G, \mathbb{C}H)$-bimodule structures on $\mathbb{C}G$ are identical.

The definition of $\operatorname{Ind}_\rho U$ shows that we have

$$\operatorname{Ind}_\rho U = \mathbb{C}G \otimes_{\mathbb{C}H} U, \tag{13.111.21}$$

where $\mathbb{C}G$ is regarded as a $(\mathbb{C}G, \mathbb{C}H)$-bimodule according to the following rule: The left $\mathbb{C}G$-module structure on $\mathbb{C}G$ is plain multiplication inside $\mathbb{C}G$; the right $\mathbb{C}H$-module structure on $\mathbb{C}G$ is induced by the $\mathbb{C}$-algebra homomorphism $\mathbb{C}[\rho] : \mathbb{C}H \to \mathbb{C}G$ (thus, it is explicitly given by $\gamma\eta = \gamma \cdot (\mathbb{C}[\rho])\,\eta$ for all $\gamma \in \mathbb{C}G$ and $\eta \in \mathbb{C}H$). We denote this $(\mathbb{C}G, \mathbb{C}H)$-bimodule structure on $\mathbb{C}G$ as the *first structure*.

On the other hand, the definition of $\operatorname{Ind}_H^G U$ shows that we have

$$\operatorname{Ind}_H^G U = \mathbb{C}G \otimes_{\mathbb{C}H} U, \tag{13.111.22}$$

where $\mathbb{C}G$ is regarded as a $(\mathbb{C}G, \mathbb{C}H)$-bimodule in the usual way (i.e., the left $\mathbb{C}G$-module structure on $\mathbb{C}G$ is plain multiplication inside $\mathbb{C}G$, and the right $\mathbb{C}H$-module structure on $\mathbb{C}G$ is also plain multiplication inside $\mathbb{C}G$ because $\mathbb{C}H \subset \mathbb{C}G$). We denote this $(\mathbb{C}G, \mathbb{C}H)$-bimodule structure on $\mathbb{C}G$ as the *second structure*.

The right hand sides of the equalities (13.111.21) and (13.111.22) appear identical, but so far we do not know if they actually mean the same thing, because the meanings of "$\mathbb{C}G$" possibly differ. Namely, we have two $(\mathbb{C}G, \mathbb{C}H)$-bimodule structures on the $\mathbb{C}$-vector space $\mathbb{C}G$: the first structure (used in (13.111.21)) and the second structure (used in (13.111.22)). These two structures clearly have the same left $\mathbb{C}G$-module structure. But they also have the same right $\mathbb{C}H$-module structure, because every $\gamma \in \mathbb{C}G$ and $\eta \in \mathbb{C}H$ satisfy

$$\begin{aligned}
&(\text{the result of the right action of } \eta \text{ on } \gamma \text{ according to the first structure}) \\
&= \gamma \cdot \underbrace{(\mathbb{C}[\rho])\,\eta}_{\substack{=\eta \\ (\text{by } (13.111.20))}} = \gamma \cdot \eta \\
&= (\text{the result of the right action of } \eta \text{ on } \gamma \text{ according to the second structure}).
\end{aligned}$$

Hence, the first structure and the second structure are identical. Thus, the right hand sides of the equalities (13.111.21) and (13.111.22) really mean the same thing. Thus, comparing the equalities (13.111.21) and (13.111.22), we obtain $\operatorname{Ind}_\rho U = \operatorname{Ind}_H^G U$ as left $\mathbb{C}G$-modules. This solves Exercise 4.1.14(d).

(e) Assume that $G = H/K$ for some normal subgroup K of H. Let $\rho : H \to G$ be the projection map. We want to prove that $\operatorname{Ind}_\rho f = f^K$ for every $f \in R_{\mathbb{C}}(H)$.

Let $f \in R_{\mathbb{C}}(H)$. Applying (13.111.1) to H instead of K, we obtain the following equivalence:

$$\begin{aligned}
&(f \in R_{\mathbb{C}}(H)) \\
\iff &(\text{any two conjugate elements } k \text{ and } k' \text{ of } H \text{ satisfy } f(k) = f(k')).
\end{aligned}$$

Hence,

$$(\text{any two conjugate elements } k \text{ and } k' \text{ of } H \text{ satisfy } f(k) = f(k')) \tag{13.111.23}$$

(since we know that $f \in R_{\mathbb{C}}(H)$). Thus,

$$f\left(y^{-1}zy\right) = f(z) \qquad \text{for all } z \in H \text{ and } y \in H \tag{13.111.24}$$

(because if $z \in H$ and $y \in H$, then $y^{-1}zy$ and z are two conjugate elements of H).

Let $g \in G$. Then, $g \in G = H/K$. Hence, there exists an $x \in H$ such that $g = xK$. Consider this x. The map $\rho : H \to G$ is the projection map from H to $G = H/K$, and thus sends every $h \in H$ to the coset $hK \in G$. In other words, $\rho(h) = hK$ for every $h \in H$. Applied to $h = x$, this yields $\rho(x) = xK = g$. Of course, $\ker \rho = K$ (since ρ is the projection map from H to H/K).

We have $|G| = |H/K| = [H : K] = |H| / |K|$.

We make another simple observation: If $h \in H$ and $y \in H$, then we have the following logical equivalence:

$$\left(\rho(h) = (\rho(y))^{-1} g \rho(y)\right) \iff \left(h \in y^{-1} x K y\right). \tag{13.111.25}$$

[809]

The definition of f^K yields $f^K(xK) = \dfrac{1}{|K|} \sum_{k \in K} f(xk)$. Hence,

$$f^K\left(\underbrace{g}_{=xK}\right) = f^K(xK) = \frac{1}{|K|} \sum_{k \in K} f(xk). \tag{13.111.26}$$

But the definition of $(\operatorname{Ind}_\rho f)(g)$ yields

$$\begin{aligned}
(\operatorname{Ind}_\rho f)(g) &= \frac{1}{|H|} \underbrace{\sum_{\substack{(h,k) \in H \times G; \\ k\rho(h)k^{-1} = g}}}_{\substack{= \sum_{\substack{(h,k) \in H \times G; \\ \rho(h) = k^{-1}gk}} \\ \text{(because for any } (h,k) \in H \times G, \\ \text{the statement } \left(k\rho(h)k^{-1} = g\right) \\ \text{is equivalent to } \left(\rho(h) = k^{-1}gk\right))}} f(h) = \frac{1}{|H|} \underbrace{\sum_{\substack{(h,k) \in H \times G; \\ \rho(h) = k^{-1}gk}}}_{= \sum_{k \in G} \sum_{\substack{h \in H; \\ \rho(h) = k^{-1}gk}}} f(h) \\
&= \frac{1}{|H|} \sum_{k \in G} \sum_{\substack{h \in H; \\ \rho(h) = k^{-1}gk}} f(h) = \frac{1}{|H|} \sum_{p \in G} \sum_{\substack{h \in H; \\ \rho(h) = p^{-1}gp}} f(h) \qquad (13.111.27)
\end{aligned}$$

[809] *Proof of (13.111.25):* Let $h \in H$ and $y \in H$. Recall that the map $\rho : H \to G$ is the projection map from H to $G = H/K$. Thus, $\rho(w) = wK$ for every $w \in H$.

We have the following chain of equivalences:

$$\begin{aligned}
&\left(\rho(h) = (\rho(y))^{-1} g \rho(y)\right) \\
&\iff (\rho(y)\rho(h) = g\rho(y)) \iff \left(\underbrace{\rho(y)\rho(h)(\rho(y))^{-1}}_{\substack{= \rho\left(yhy^{-1}\right) \\ \text{(since } \rho \text{ is a group} \\ \text{homomorphism)}}} = \underbrace{g}_{= \rho(x)}\right) \\
&\iff \left(\underbrace{\rho\left(yhy^{-1}\right)}_{\substack{= yhy^{-1}K \\ \text{(since } \rho(w) = wK \text{ for every } w \in H)}} = \underbrace{\rho(x)}_{\substack{= xK \\ \text{(since } \rho(w) = wK \text{ for every } w \in H)}}\right) \iff \left(yhy^{-1}K = xK\right) \\
&\iff \left(yhy^{-1} \in xK\right) \iff (yh \in xKy) \iff \left(h \in y^{-1}xKy\right).
\end{aligned}$$

This proves (13.111.25).

(here, we renamed the index k as p in the first sum). But every $p \in G$ satisfies $\sum\limits_{\substack{h\in H;\\ \rho(h)=p^{-1}gp}} f(h) = \sum\limits_{k\in K} f(xk)$ [810]. Hence, (13.111.27) becomes

$$\left(\operatorname{Ind}_\rho f\right)(g) = \frac{1}{|H|} \sum_{p\in G} \underbrace{\sum_{\substack{h\in H;\\ \rho(h)=p^{-1}gp}} f(h)}_{=\sum_{k\in K} f(xk)} = \frac{1}{|H|} \underbrace{\sum_{p\in G}\sum_{k\in K} f(xk)}_{=|G|\cdot\sum_{k\in K} f(xk)} = \frac{1}{|H|} \underbrace{|G|}_{=|H|/|K|} \cdot \sum_{k\in K} f(xk)$$

$$= \underbrace{\frac{1}{|H|}\cdot |H|/|K|}_{=\frac{1}{|K|}} \cdot \sum_{k\in K} f(xk) = \frac{1}{|K|} \sum_{k\in K} f(xk) = f^K(g)$$

(by (13.111.26)).

Let us now forget that we fixed g. We thus have shown that $\left(\operatorname{Ind}_\rho f\right)(g) = f^K(g)$ for every $g \in G$. In other words, $\operatorname{Ind}_\rho f = f^K$. This solves Exercise 4.1.14(e).

(f) Assume that $G = H/K$ for some normal subgroup K of H. Let $\rho : H \to G$ be the projection map. We want to prove that $\operatorname{Ind}_\rho U \cong U^K$ for every $\mathbb{C}H$-module U.

The map ρ is the projection map from H to $G = H/K$. Thus, the map ρ is surjective and has kernel $\ker\rho = K$. Furthermore,

$$\left|\rho^{-1}(g)\right| = |K| \qquad \text{for every } g \in G. \tag{13.111.29}$$

[811]

Let U be a $\mathbb{C}H$-module. Recall that U^K is a $\mathbb{C}[H/K]$-module, thus a $\mathbb{C}G$-module (since $H/K = G$).

Recall that $\operatorname{Ind}_\rho U$ is defined as the $\mathbb{C}G$-module $\mathbb{C}G \otimes_{\mathbb{C}H} U$, where $\mathbb{C}G$ is regarded as a $(\mathbb{C}G, \mathbb{C}H)$-bimodule according to the following rule: The left $\mathbb{C}G$-module structure on $\mathbb{C}G$ is plain multiplication inside $\mathbb{C}G$; the right $\mathbb{C}H$-module structure on $\mathbb{C}G$ is induced by the $\mathbb{C}$-algebra homomorphism $\mathbb{C}[\rho] : \mathbb{C}H \to \mathbb{C}G$ (thus, it is explicitly given by $\gamma\eta = \gamma \cdot (\mathbb{C}[\rho])\eta$ for all $\gamma \in \mathbb{C}G$ and $\eta \in \mathbb{C}H$). From now on, we regard $\mathbb{C}G$ as endowed with this $(\mathbb{C}G, \mathbb{C}H)$-bimodule structure.

[810] *Proof.* Let $p \in G$. Then, $p \in G = H/K$. Hence, there exists a $y \in H$ such that $p = yK$. Consider this y.

Recall that $\rho(h) = hK$ for every $h \in H$. Applying this to $h = y$, we obtain $\rho(y) = yK = p$.

We notice that the map $K \to y^{-1}xKy,\ k \mapsto y^{-1}xky$ is a bijection (since H is a group).

Now, for every $h \in H$, we have the following equivalence:

$$\begin{aligned} \left(\rho(h) = p^{-1}gp\right) &\iff \left(\rho(h) = (\rho(y))^{-1} g\rho(y)\right) && (\text{since } p = \rho(y)) \\ &\iff \left(h \in y^{-1}xKy\right) && (\text{according to } (13.111.25)). \end{aligned} \tag{13.111.28}$$

Now,

$$\underbrace{\sum_{\substack{h\in H;\\ \rho(h)=p^{-1}gp}}}_{\substack{=\sum_{\substack{h\in H;\\ h\in y^{-1}xKy}} \\ (\text{because for every } h\in H, \text{ the statement} \\ \left(\rho(h)=p^{-1}gp\right) \text{ is equivalent to } \left(h\in y^{-1}xKy\right) \\ (\text{because of } (13.111.28)))}} f(h) = \underbrace{\sum_{\substack{h\in H;\\ h\in y^{-1}xKy}}}_{\substack{=\sum_{h\in y^{-1}xKy} \\ (\text{since } y^{-1}xKy\subset H)}} f(h) = \sum_{h\in y^{-1}xKy} f(h) = \sum_{k\in K} f\left(y^{-1}xky\right)$$

(here, we substituted $y^{-1}xky$ for h in the sum, since the map $K \to y^{-1}xKy,\ k \mapsto y^{-1}xky$ is a bijection). Thus,

$$\sum_{\substack{h\in H;\\ \rho(h)=p^{-1}gp}} f(h) = \sum_{k\in K} \underbrace{f\left(y^{-1}xky\right)}_{\substack{=f(xk) \\ (\text{by } (13.111.24), \text{ applied to } z=xk)}} = \sum_{k\in K} f(xk),$$

qed.

[811] *Proof of (13.111.29):* Let $g \in G$. The map ρ is surjective. Hence, $\rho(H) = G$. Thus, $g \in G = \rho(H)$. Therefore, there exists some $x \in H$ such that $g = \rho(x)$. Let us fix such an x.

We know that H is a group. Hence, the map $K \to xK,\ k \mapsto xk$ is a bijection. Thus, the sets K and xK are in bijection. Therefore, $|xK| = |K|$.

We define a map $\alpha : U^K \to \mathbb{C}G \otimes_{\mathbb{C}H} U$ by setting

$$\alpha(u) = 1 \otimes_{\mathbb{C}H} u \qquad \text{for every } u \in U^K.$$

This α is a map $U^K \to \operatorname{Ind}_\rho U$ (since $\operatorname{Ind}_\rho U = \mathbb{C}G \otimes_{\mathbb{C}H} U$) and is $\mathbb{C}$-linear (since $\alpha(u) = 1 \otimes_{\mathbb{C}H} u$ depends $\mathbb{C}$-linearly on u).

We will show that α is a $\mathbb{C}G$-module isomorphism.

The $\mathbb{C}G$-module structure on U^K has the property that

$$\rho(h) \cdot v = hv \qquad \text{for any } h \in H \text{ and } v \in U^K. \tag{13.111.30}$$

[812]

Now, α is a $\mathbb{C}G$-module homomorphism[813]. We will eventually construct an inverse to α; but first we need to prepare.

For every $g \in G$ and every $p \in \mathbb{C}G$, we denote by $\epsilon_g(p)$ the g-coordinate of p with respect to the basis G of the $\mathbb{C}$-vector space $\mathbb{C}G$. By the definition of "coordinate", we have

$$q = \sum_{g \in G} \epsilon_g(q)\, g \qquad \text{for every } q \in \mathbb{C}G. \tag{13.111.32}$$

However, for every $y \in H$, we have the following logical equivalence:

$$\left(y \in \rho^{-1}(g)\right) \iff \left(\rho(y) = \underbrace{g}_{=\rho(x)}\right) \iff (\rho(y) = \rho(x)) \iff \left(\underbrace{\rho(y) \cdot (\rho(x))^{-1}}_{\substack{=\rho(yx^{-1}) \\ \text{(since } \rho \text{ is a group} \\ \text{homomorphism)}}} = 1\right)$$

$$\iff \left(\rho\left(yx^{-1}\right) = 1\right) \iff \left(yx^{-1} \in \underbrace{\ker \rho}_{=K}\right) \iff \left(yx^{-1} \in K\right) \iff (y \in xK).$$

Hence, $\rho^{-1}(g) = xK$, so that $\left|\rho^{-1}(g)\right| = |xK| = |K|$. This proves (13.111.29).

[812] *Proof of (13.111.30):* The map ρ is the projection map from H to H/K. Thus, the map ρ sends every $h \in H$ to the coset $hK \in H/K$. In other words,

$$\rho(h) = hK \qquad \text{for every } h \in H, \tag{13.111.31}$$

where hK means the coset $hK \in H/K$. But by the definition of the $\mathbb{C}[H/K]$-module structure on U^K, we have

$$(hK) \cdot v = hv \qquad \text{for any } h \in H \text{ and } v \in U^K,$$

where hK means the coset $hK \in H/K$. Thus, any $h \in H$ and $v \in U^K$ satisfy

$$\underbrace{\rho(h)}_{\substack{=hK \\ \text{(by (13.111.31))}}} \cdot v = (hK) \cdot v = hv,$$

where hK means the coset $hK \in H/K$. This proves (13.111.30).

[813] *Proof.* Let $u \in U^K$ and $g \in G$. We have $g \in G = \rho(H)$ (since the map $\rho : H \to G$ is surjective). Thus, $g = \rho(y)$ for some $y \in H$. Let us consider this y.

Applying (13.111.30) to $h = y$ and $v = u$, we obtain $\rho(y) \cdot u = yu$. Thus, $\underbrace{g}_{=\rho(y)} u = \rho(y) \cdot u = yu$.

Now, the definition of $\alpha(gu)$ yields

$$\alpha(gu) = 1 \otimes_{\mathbb{C}H} \underbrace{gu}_{=yu} = 1 \otimes_{\mathbb{C}H} yu = 1y \otimes_{\mathbb{C}H} u$$

(here, we moved the y past the tensor sign; this is allowed because $y \in H \subset \mathbb{C}H$).

By the definition of the left $\mathbb{C}G$-module structure on $\mathbb{C}G$, we have $g \cdot 1 = g1 = g$.

By the definition of the right $\mathbb{C}H$-module structure on $\mathbb{C}G$, we have $\gamma\eta = \gamma \cdot (\mathbb{C}[\rho])\eta$ for all $\gamma \in \mathbb{C}G$ and $\eta \in \mathbb{C}H$. Applying this to $\gamma = 1$ and $\eta = y$, we obtain $1y = 1 \cdot (\mathbb{C}[\rho])y$. But $y \in H$ and thus $(\mathbb{C}[\rho])y = \rho(y) = g$. Therefore, $1y = 1 \cdot \underbrace{(\mathbb{C}[\rho])y}_{=g} = 1 \cdot g = g = g \cdot 1$.

Now, $\alpha(gu) = \underbrace{1y}_{=g\cdot 1} \otimes_{\mathbb{C}H} u = g \cdot 1 \otimes_{\mathbb{C}H} u$. Compared with $g \cdot \underbrace{\alpha(u)}_{=1\otimes_{\mathbb{C}H} u} = g \cdot (1 \otimes_{\mathbb{C}H} u) = g \cdot 1 \otimes_{\mathbb{C}H} u$, this yields $\alpha(gu) = g \cdot \alpha(u)$.

Now, let us forget that we fixed u and g. We thus have shown that $\alpha(gu) = g \cdot \alpha(u)$ for all $u \in U^K$ and $g \in G$. Thus, the map α is a homomorphism of G-sets. Since the map α is also $\mathbb{C}$-linear, this yields that α is a $\mathbb{C}G$-module homomorphism. Qed.

For every $g \in G$, we have defined a map $\epsilon_g : \mathbb{C}G \to \mathbb{C}$ (because we have defined an element $\epsilon_g(p)$ for every $p \in \mathbb{C}G$). This map ϵ_g is $\mathbb{C}$-linear. We notice some basic properties of these maps:

- For every $g \in G$ and $h \in G$, we have

$$\epsilon_g(h) = \delta_{g,h}. \qquad (13.111.33)$$

[814]

- We have

$$\epsilon_1(pq) = \epsilon_1(qp) \qquad \text{for all } p \in \mathbb{C}G \text{ and } q \in \mathbb{C}G. \qquad (13.111.34)$$

[815]

- We have

$$\epsilon_1\left(g^{-1}q\right) = \epsilon_g(q) \qquad \text{for every } g \in G \text{ and } q \in \mathbb{C}G. \qquad (13.111.35)$$

[816]

Now, for every $(q,u) \in \mathbb{C}G \times U$, we have $\sum\limits_{h \in H} \epsilon_1(\rho(h)q)h^{-1}u \in U^K$ [817]. Hence, we can define a map $\widetilde{\beta} : \mathbb{C}G \times U \to U^K$ by setting

$$\widetilde{\beta}(q,u) = \sum_{h \in H} \epsilon_1(\rho(h)q)h^{-1}u \qquad \text{for every } (q,u) \in \mathbb{C}G \times U.$$

Consider this map $\widetilde{\beta}$. Then, $\widetilde{\beta}$ is a $\mathbb{C}$-bilinear map (because $\widetilde{\beta}(q,u) = \sum_{h \in H} \epsilon_1(\rho(h)q)h^{-1}u$ depends $\mathbb{C}$-linearly on each of q and u). We are now going to prove that the map $\widetilde{\beta}$ is $\mathbb{C}H$-bilinear with respect to the right $\mathbb{C}H$-module structure on $\mathbb{C}G$ and the left $\mathbb{C}H$-module structure on U.

In fact, every $h' \in H$, $q \in \mathbb{C}G$ and $u \in U$ satisfy

$$\widetilde{\beta}(q,h'u) = \widetilde{\beta}(qh',u). \qquad (13.111.36)$$

[814] This equality has been proven in our solution of Exercise 4.1.4. (Namely, it appeared there as (13.105.3).)

[815] This equality has been proven in our solution of Exercise 4.1.4. (Namely, it appeared there as (13.105.4).)

[816] This equality has been proven in our solution of Exercise 4.1.4. (Namely, it appeared there as (13.105.5).)

[817] *Proof.* Let $(q,u) \in \mathbb{C}G \times U$. Let $x = \sum\limits_{h \in H} \epsilon_1(\rho(h)q)h^{-1}u$. We shall show that $x \in U^K$.

Let $k \in K$. Then, the map $H \to H$, $h \mapsto hk$ is a bijection (since H is a group). Now, multiplying both sides of the equality $x = \sum\limits_{h \in H} \epsilon_1(\rho(h)q)h^{-1}u$ with k from the left, we obtain

$$\begin{aligned} kx &= k\sum_{h \in H} \epsilon_1(\rho(h)q)h^{-1}u = \sum_{h \in H} \epsilon_1(\rho(h)q)kh^{-1}u \\ &= \sum_{h \in H} \epsilon_1\left(\underbrace{\rho(hk)}_{\substack{=\rho(h)\rho(k) \\ \text{(since } \rho \text{ is a group} \\ \text{homomorphism)}}} q\right) k \underbrace{(hk)^{-1}}_{=k^{-1}h^{-1}} u \qquad \left(\begin{array}{c}\text{here, we substituted } hk \text{ for } h \text{ in the sum,} \\ \text{since the map } H \to H,\ h \mapsto hk \text{ is a bijection}\end{array}\right) \\ &= \sum_{h \in H} \epsilon_1\left(\rho(h) \underbrace{\rho(k)}_{\substack{=1 \\ \text{(since } k \in K = \ker\rho)}} q\right) \underbrace{kk^{-1}}_{=1} h^{-1}u = \sum_{h \in H} \epsilon_1(\rho(h)q)h^{-1}u = x. \end{aligned}$$

Let us now forget that we fixed k. We thus have shown that $kx = x$ for every $k \in K$. In other words, x is an element $y \in U$ satisfying $ky = y$ for every $k \in K$. Hence,

$$x \in \{y \in U \mid ky = y \text{ for every } k \in K\} = U^K.$$

Thus, $\sum\limits_{h \in H} \epsilon_1(\rho(h)q)h^{-1}u = x \in U^K$, qed.

[818] As a consequence of this, we can see that every $r \in \mathbb{C}H$, $q \in \mathbb{C}G$ and $u \in U$ satisfy

$$\widetilde{\beta}\left(q, ru\right) = \widetilde{\beta}\left(qr, u\right). \tag{13.111.38}$$

[819] In other words, the map $\widetilde{\beta}$ is $\mathbb{C}H$-bilinear with respect to the right $\mathbb{C}H$-module structure on $\mathbb{C}G$ and the left $\mathbb{C}H$-module structure on U (since we already know that $\widetilde{\beta}$ is $\mathbb{C}$-bilinear). Hence, by the universal property of the tensor product, we conclude that there exists a unique $\mathbb{C}$-linear map $\beta : \mathbb{C}G \otimes_{\mathbb{C}H} U \to U^K$ such that every $(q, u) \in \mathbb{C}G \times U$ satisfies

$$\beta\left(q \otimes_{\mathbb{C}H} u\right) = \widetilde{\beta}\left(q, u\right). \tag{13.111.39}$$

Consider this map β. Clearly, every $q \in \mathbb{C}G$ and $u \in U$ satisfy

$$\beta\left(q \otimes_{\mathbb{C}H} u\right) = \widetilde{\beta}\left(q, u\right) = \sum_{h \in H} \epsilon_1\left(\rho\left(h\right) q\right) h^{-1} u \qquad \left(\text{by the definition of } \widetilde{\beta}\right). \tag{13.111.40}$$

We shall now show that the maps α and $\dfrac{1}{|K|}\beta$ are mutually inverse. To do so, we will show that $\alpha \circ \beta = |K| \operatorname{id}$ and $\beta \circ \alpha = |K| \operatorname{id}$.

Let us first notice that

$$1 \otimes_{\mathbb{C}H} h^{-1} u = \left(\rho\left(h\right)\right)^{-1} \otimes_{\mathbb{C}H} u \qquad \text{for any } h \in H \text{ and } u \in U. \tag{13.111.41}$$

[820]

[818] *Proof of (13.111.36):* Let $h' \in H$, $q \in \mathbb{C}G$ and $u \in U$. Then, the map $H \to H,\ h \mapsto h'h$ is a bijection (since H is a group). The definition of $\widetilde{\beta}$ yields

$$\begin{aligned}
\widetilde{\beta}\left(q, h'u\right) &= \sum_{h \in H} \epsilon_1\left(\rho\left(h\right) q\right) h^{-1} h' u = \sum_{h \in H} \epsilon_1\left(\underbrace{\rho\left(h'h\right)}_{\substack{=\rho\left(h'\right)\rho\left(h\right) \\ \text{(since } \rho \text{ is a group} \\ \text{homomorphism)}}} q \right) \underbrace{\left(h'h\right)^{-1}}_{=h^{-1}\left(h'\right)^{-1}} h' u \\
&\qquad \left(\begin{array}{c}\text{here, we substituted } h'h \text{ for } h \text{ in the sum,} \\ \text{since the map } H \to H,\ h \mapsto h'h \text{ is a bijection}\end{array}\right) \\
&= \sum_{h \in H} \underbrace{\epsilon_1\left(\rho\left(h'\right)\rho\left(h\right) q\right)}_{\substack{=\epsilon_1\left(\rho\left(h\right) q \rho\left(h'\right)\right) \\ \text{(by (13.111.34), applied to} \\ \rho\left(h'\right) \text{ and } \rho\left(h\right) q \text{ instead of } p \text{ and } q)}} h^{-1} \underbrace{\left(h'\right)^{-1} h'}_{=1} u \\
&= \sum_{h \in H} \epsilon_1\left(\rho\left(h\right) q \rho\left(h'\right)\right) h^{-1} u. \qquad (13.111.37)
\end{aligned}$$

But the definition of the right $\mathbb{C}H$-module structure on $\mathbb{C}G$ yields $\gamma\eta = \gamma \cdot \left(\mathbb{C}\left[\rho\right]\right) \eta$ for all $\gamma \in \mathbb{C}G$ and $\eta \in \mathbb{C}H$. Applying this to $\gamma = q$ and $\eta = h'$, we obtain $qh' = q \cdot \left(\mathbb{C}\left[\rho\right]\right)\left(h'\right)$. But $h' \in H$ and thus $\left(\mathbb{C}\left[\rho\right]\right)\left(h'\right) = \rho\left(h'\right)$. Hence, $qh' = q \cdot \underbrace{\left(\mathbb{C}\left[\rho\right]\right)\left(h'\right)}_{=\rho\left(h'\right)} = q\rho\left(h'\right)$.

But the definition of $\widetilde{\beta}$ also yields

$$\begin{aligned}
\widetilde{\beta}\left(qh', u\right) &= \sum_{h \in H} \epsilon_1\left(\rho\left(h\right) \underbrace{qh'}_{=q\rho\left(h'\right)} \right) h^{-1} u = \sum_{h \in H} \epsilon_1\left(\rho\left(h\right) q \rho\left(h'\right)\right) h^{-1} u \\
&= \widetilde{\beta}\left(q, h'u\right) \qquad \left(\text{by (13.111.37)}\right).
\end{aligned}$$

In other words, $\widetilde{\beta}\left(q, h'u\right) = \widetilde{\beta}\left(qh', u\right)$. This proves (13.111.36).

[819] *Proof of (13.111.38):* Let $r \in \mathbb{C}H$, $q \in \mathbb{C}G$ and $u \in U$. We need to prove the equality (13.111.38). But this equality is $\mathbb{C}$-linear in r. Hence, we can WLOG assume that r belongs to the basis H of the $\mathbb{C}$-vector space $\mathbb{C}H$. Assume this. Then, (13.111.38) follows from (13.111.36) (applied to $h' = r$).

[820] *Proof of (13.111.41):* Let $h \in H$ and $u \in U$. We have $h \in H$. Since H is a group, this yields $h^{-1} \in H \subset \mathbb{C}H$. Hence, we can move the h^{-1} past the tensor sign in $1 \otimes_{\mathbb{C}H} h^{-1} u$. We thus obtain $1 \otimes_{\mathbb{C}H} h^{-1} u = 1h^{-1} \otimes_{\mathbb{C}H} u$.

But the definition of the right $\mathbb{C}H$-module structure on $\mathbb{C}G$ yields $\gamma\eta = \gamma \cdot \left(\mathbb{C}\left[\rho\right]\right) \eta$ for all $\gamma \in \mathbb{C}G$ and $\eta \in \mathbb{C}H$. Applying this to $\gamma = 1$ and $\eta = h^{-1}$, we obtain $1h^{-1} = 1 \cdot \left(\mathbb{C}\left[\rho\right]\right)\left(h^{-1}\right)$. But $h^{-1} \in H$ and thus $\left(\mathbb{C}\left[\rho\right]\right)\left(h^{-1}\right) = \rho\left(h^{-1}\right) = \left(\rho\left(h\right)\right)^{-1}$ (since ρ

Now, we are going to show that $\alpha \circ \beta = |K| \operatorname{id}$. In fact, let $q \in \mathbb{C}G$ and $u \in U$ be arbitrary. Then,

$$
\begin{aligned}
(\alpha\circ\beta)\left(q\otimes_{\mathbb{C}H} u\right) &= \alpha\left(\beta\left(q\otimes_{\mathbb{C}H}u\right)\right) = 1\otimes_{\mathbb{C}H} \underbrace{\beta\left(q\otimes_{\mathbb{C}H}u\right)}_{\substack{=\sum_{h\in H}\epsilon_1(\rho(h)q)h^{-1}u \\ \text{(by (13.111.40))}}} \\
&\qquad (\text{by the definition of } \alpha) \\
&= 1\otimes_{\mathbb{C}H}\left(\sum_{h\in H}\epsilon_1\left(\rho\left(h\right)q\right)h^{-1}u\right) = \underbrace{\sum_{h\in H}}_{=\sum_{g\in G}\sum_{\substack{h\in H;\\ \rho(h)=g}}} \epsilon_1\left(\rho\left(h\right)q\right) \underbrace{1\otimes_{\mathbb{C}H}h^{-1}u}_{\substack{=(\rho(h))^{-1}\otimes_{\mathbb{C}H}u\\ \text{(by (13.111.41))}}} \\
&= \sum_{g\in G}\underbrace{\sum_{\substack{h\in H;\\ \rho(h)=g}}}_{=\sum_{h\in\rho^{-1}(g)}} \epsilon_1\left(\underbrace{\rho\left(h\right)}_{=g}q\right)\left(\underbrace{\rho\left(h\right)}_{=g}\right)^{-1}\otimes_{\mathbb{C}H}u = \sum_{g\in G}\underbrace{\sum_{h\in\rho^{-1}(g)}\epsilon_1\left(gq\right)g^{-1}\otimes_{\mathbb{C}H}u}_{=|\rho^{-1}(g)|\epsilon_1(gq)g^{-1}\otimes_{\mathbb{C}H}u} \\
&= \sum_{g\in G}\underbrace{\left|\rho^{-1}\left(g\right)\right|}_{\substack{=|K|\\ \text{(by (13.111.29))}}}\epsilon_1\left(gq\right)g^{-1}\otimes_{\mathbb{C}H}u = \sum_{g\in G}|K|\,\epsilon_1\left(gq\right)g^{-1}\otimes_{\mathbb{C}H}u \\
&= |K|\left(\sum_{g\in G}\epsilon_1\left(gq\right)g^{-1}\right)\otimes_{\mathbb{C}H}u = |K|\left(\sum_{g\in G}\epsilon_1\left(g^{-1}q\right)\underbrace{\left(g^{-1}\right)^{-1}}_{=g}\right)\otimes_{\mathbb{C}H}u \\
&\qquad \left(\begin{array}{c}\text{here, we substituted } g^{-1} \text{ for } g \text{ in the sum, since the map}\\ G\to G,\ g\mapsto g^{-1} \text{ is a bijection (since } G \text{ is a group)}\end{array}\right) \\
&= |K|\left(\sum_{g\in G}\underbrace{\epsilon_1\left(g^{-1}q\right)}_{\substack{=\epsilon_g(q)\\ \text{(by (13.111.35))}}}g\right)\otimes_{\mathbb{C}H}u = |K|\underbrace{\left(\sum_{g\in G}\epsilon_g\left(q\right)g\right)}_{\substack{=q\\ \text{(by (13.111.32))}}}\otimes_{\mathbb{C}H}u \\
&= |K|\underbrace{q\otimes_{\mathbb{C}H}u}_{=\operatorname{id}(q\otimes_{\mathbb{C}H}u)} = |K|\operatorname{id}\left(q\otimes_{\mathbb{C}H}u\right).
\end{aligned}
$$

Now, let us forget that we fixed q and u. We thus have shown that $(\alpha\circ\beta)\left(q\otimes_{\mathbb{C}H}u\right) = |K|\operatorname{id}\left(q\otimes_{\mathbb{C}H}u\right)$ for all $q\in\mathbb{C}G$ and $u\in U$. In other words, the two maps $\alpha\circ\beta : \mathbb{C}G\otimes_{\mathbb{C}H}U\to\mathbb{C}G\otimes_{\mathbb{C}H}U$ and $|K|\operatorname{id} : \mathbb{C}G\otimes_{\mathbb{C}H}U\to\mathbb{C}G\otimes_{\mathbb{C}H}U$ are equal to each other on each pure tensor. Since these two maps are $\mathbb{C}$-linear, this yields that these two maps $\alpha\circ\beta : \mathbb{C}G\otimes_{\mathbb{C}H}U\to\mathbb{C}G\otimes_{\mathbb{C}H}U$ and $|K|\operatorname{id} : \mathbb{C}G\otimes_{\mathbb{C}H}U\to\mathbb{C}G\otimes_{\mathbb{C}H}U$ must be identical. In other words, $\alpha\circ\beta = |K|\operatorname{id}$.

Next, we are going to show that $\beta\circ\alpha = |K|\operatorname{id}$.

We first notice that

$$h^{-1}u = \left(\rho\left(h\right)\right)^{-1}u \qquad \text{for every } h\in H \text{ and } u\in U^K. \tag{13.111.42}$$

[821]

is a group homomorphism). Hence, $1h^{-1} = 1\cdot\underbrace{\left(\mathbb{C}\left[\rho\right]\right)\left(h^{-1}\right)}_{=(\rho(h))^{-1}} = 1\cdot\left(\rho\left(h\right)\right)^{-1} = \left(\rho\left(h\right)\right)^{-1}$. Now, $1\otimes_{\mathbb{C}H}h^{-1}u = \underbrace{1h^{-1}}_{=(\rho(h))^{-1}}\otimes_{\mathbb{C}H}u = \left(\rho\left(h\right)\right)^{-1}\otimes_{\mathbb{C}H}u$. This proves (13.111.41).

[821] *Proof of (13.111.42):* Let $h\in H$ and $u\in U^K$. Applying (13.111.30) to h^{-1} and u instead of h and v, we obtain $\rho\left(h^{-1}\right)\cdot u = h^{-1}u$. But ρ is a group homomorphism, and thus we have $\left(\rho\left(h\right)\right)^{-1} = \rho\left(h^{-1}\right)$. Hence, $\underbrace{\left(\rho\left(h\right)\right)^{-1}}_{=\rho(h^{-1})}u = \rho\left(h^{-1}\right)u = \rho\left(h^{-1}\right)\cdot u = h^{-1}u$. This proves (13.111.42).

Let $u \in U^K$. The definition of $\alpha(u)$ yields $\alpha(u) = 1 \otimes_{\mathbb{C}H} u$. Applying the map β to this equality, we obtain

$$\beta(\alpha(u)) = \beta(1 \otimes_{\mathbb{C}H} u) = \underbrace{\sum_{h \in H}}_{= \sum_{g \in G} \sum_{\substack{h \in H;\\ \rho(h) = g}}} \epsilon_1(\rho(h)1) \underbrace{h^{-1}u}_{\substack{= (\rho(h))^{-1}u \\ \text{(by (13.111.42))}}} \qquad \text{(by (13.111.40), applied to } q = 1\text{)}$$

$$= \sum_{g \in G} \underbrace{\sum_{\substack{h \in H;\\ \rho(h) = g}}}_{= \sum_{h \in \rho^{-1}(g)}} \epsilon_1\left(\underbrace{\rho(h)}_{=g} 1\right) \left(\underbrace{\rho(h)}_{=g}\right)^{-1} u = \sum_{g \in G} \underbrace{\sum_{h \in \rho^{-1}(g)} \epsilon_1(g1) g^{-1}u}_{= |\rho^{-1}(g)| \epsilon_1(g1) g^{-1}u}$$

$$= \sum_{g \in G} \underbrace{\left|\rho^{-1}(g)\right|}_{\substack{= |K| \\ \text{(by (13.111.29))}}} \epsilon_1(g1) g^{-1}u = \sum_{g \in G} |K| \epsilon_1(g1) g^{-1}u$$

$$= |K| \sum_{g \in G} \epsilon_1(g1) g^{-1}u = |K| \sum_{g \in G} \epsilon_1\left(g^{-1}1\right) \underbrace{\left(g^{-1}\right)^{-1}}_{=g} u$$

$$\left(\begin{array}{c} \text{here, we substituted } g^{-1} \text{ for } g \text{ in the sum, since the map} \\ G \to G,\ g \mapsto g^{-1} \text{ is a bijection (since } G \text{ is a group)} \end{array}\right)$$

$$= |K| \sum_{g \in G} \underbrace{\epsilon_1\left(g^{-1}1\right)}_{\substack{= \epsilon_g(1) \\ \text{(by (13.111.35), applied to } q=1\text{)}}} gu = |K| \sum_{g \in G} \epsilon_g(1) gu = |K| \underbrace{\left(\sum_{g \in G} \epsilon_g(1) g\right)}_{\substack{=1 \\ \text{(because } 1 = \sum_{g \in G} \epsilon_g(1) g \\ \text{(by (13.111.32), applied to } q=1\text{))}}} u$$

$$= |K| \underbrace{u}_{= \operatorname{id}(u)} = |K| \operatorname{id}(u).$$

In other words, $(\beta \circ \alpha)(u) = |K| \operatorname{id}(u)$ (since $(\beta \circ \alpha)(u) = \beta(\alpha(u))$).

Now, forget that we fixed u. We thus have shown that $(\beta \circ \alpha)(u) = |K| \operatorname{id}(u)$ for every $u \in U^K$. In other words, $\beta \circ \alpha = |K| \operatorname{id}$.

The equalities $\alpha \circ \left(\dfrac{1}{|K|}\beta\right) = \dfrac{1}{|K|} \underbrace{\alpha \circ \beta}_{= |K| \operatorname{id}} = \dfrac{1}{|K|} \cdot |K| \operatorname{id} = \operatorname{id}$ and $\left(\dfrac{1}{|K|}\beta\right) \circ \alpha = \dfrac{1}{|K|} \underbrace{\beta \circ \alpha}_{= |K| \operatorname{id}} = \dfrac{1}{|K|} \cdot |K| \operatorname{id} = \operatorname{id}$ show that the maps α and $\dfrac{1}{|K|}\beta$ are mutually inverse. Hence, the map α is invertible.

Now, we know that the map α is an invertible $\mathbb{C}G$-module homomorphism. Hence, α is a $\mathbb{C}G$-module isomorphism. Therefore, there exists a $\mathbb{C}G$-module isomorphism $U^K \to \mathbb{C}G \otimes_{\mathbb{C}H} U$ (namely, α). Therefore, $U^K \cong \mathbb{C}G \otimes_{\mathbb{C}H} U = \operatorname{Ind}_\rho U$ as $\mathbb{C}G$-modules. This solves Exercise 4.1.14(f).

[*Remark:* There is another solution of Exercise 4.1.14(f), which uses the result of Exercise 4.1.12(b). Yet another solution of Exercise 4.1.14(f) relies on Exercise 4.1.14(i), and will be given after the solution of the latter.]

(g) Let $\alpha \in R_{\mathbb{C}}(H)$ and $\beta \in R_{\mathbb{C}}(G)$. We notice that

$$\text{any two conjugate elements } k \text{ and } k' \text{ of } G \text{ satisfy } \beta(k) = \beta(k') \tag{13.111.43}$$

[822].

[822] *Proof of (13.111.43):* We can apply (13.111.1) to $K = G$ and $f = \beta$. As a consequence, we obtain the following equivalence:

$$(\beta \in R_{\mathbb{C}}(G))$$
$$\iff (\text{any two conjugate elements } k \text{ and } k' \text{ of } G \text{ satisfy } \beta(k) = \beta(k')).$$

Hence, any two conjugate elements k and k' of G satisfy $\beta(k) = \beta(k')$ (because we know that $\beta \in R_{\mathbb{C}}(G)$).

Let us first prove (4.1.17). The definition of $\langle \operatorname{Ind}_\rho \alpha, \beta \rangle_G$ yields

$$\begin{aligned}
\langle \operatorname{Ind}_\rho \alpha, \beta \rangle_G &= \frac{1}{|G|} \sum_{g \in G} \underbrace{\left(\operatorname{Ind}_\rho \alpha\right)(g)}_{\substack{= \frac{1}{|H|} \sum_{\substack{(h,k) \in H \times G; \\ k\rho(h)k^{-1} = g}} \alpha(h) \\ \text{(by the definition of } \operatorname{Ind}_\rho \alpha)}} \beta\left(g^{-1}\right) = \frac{1}{|G|} \sum_{g \in G} \left(\frac{1}{|H|} \sum_{\substack{(h,k) \in H \times G; \\ k\rho(h)k^{-1} = g}} \alpha(h) \right) \beta\left(g^{-1}\right) \\
&= \frac{1}{|G|} \frac{1}{|H|} \underbrace{\sum_{g \in G} \sum_{\substack{(h,k) \in H \times G; \\ k\rho(h)k^{-1} = g}}}_{= \sum_{(h,k) \in H \times G} \sum_{\substack{g \in G; \\ k\rho(h)k^{-1} = g}}} \alpha(h)\, \beta \left(\underbrace{g^{-1}}_{\substack{= \left(k\rho(h)k^{-1}\right)^{-1} \\ \text{(since } g = k\rho(h)k^{-1} \\ \text{(since } k\rho(h)k^{-1} = g))}} \right) \\
&= \frac{1}{|G|} \frac{1}{|H|} \sum_{(h,k) \in H \times G} \sum_{\substack{g \in G; \\ k\rho(h)k^{-1} = g}} \alpha(h)\, \beta\left(\left(k\rho(h)\, k^{-1} \right)^{-1} \right). \qquad (13.111.44)
\end{aligned}$$

However, for every $(h,k) \in H \times G$, we have

$$\sum_{\substack{g \in G; \\ k\rho(h)k^{-1} = g}} \alpha(h)\, \beta\left(\left(k\rho(h)\, k^{-1} \right)^{-1} \right) = \alpha(h) \left(\operatorname{Res}_\rho \beta \right) \left(h^{-1} \right) \qquad (13.111.45)$$

[823]. Thus, (13.111.44) becomes

$$\begin{aligned}\left\langle \operatorname{Ind}_\rho \alpha, \beta \right\rangle_G &= \frac{1}{|G|} \frac{1}{|H|} \underbrace{\sum_{(h,k) \in H \times G}}_{= \sum_{h \in H} \sum_{k \in G}} \underbrace{\sum_{\substack{g \in G; \\ k\rho(h)k^{-1} = g}} \alpha(h) \beta\left(\left(k\rho(h) k^{-1}\right)^{-1}\right)}_{= \alpha(h) (\operatorname{Res}_\rho \beta)(h^{-1})} \\ &= \frac{1}{|G|} \frac{1}{|H|} \sum_{h \in H} \underbrace{\sum_{k \in G} \alpha(h) (\operatorname{Res}_\rho \beta)\left(h^{-1}\right)}_{= |G| \cdot \alpha(h) (\operatorname{Res}_\rho \beta)(h^{-1})} = \frac{1}{|G|} \frac{1}{|H|} \sum_{h \in H} |G| \cdot \alpha(h) (\operatorname{Res}_\rho \beta)\left(h^{-1}\right) \\ &= \frac{1}{|H|} \sum_{h \in H} \alpha(h) (\operatorname{Res}_\rho \beta)\left(h^{-1}\right).\end{aligned}$$

Compared with

$$\begin{aligned}\left\langle \alpha, \operatorname{Res}_\rho \beta \right\rangle_H &= \frac{1}{|H|} \sum_{g \in H} \alpha(g) (\operatorname{Res}_\rho \beta)\left(g^{-1}\right) \qquad \left(\text{by the definition of } \left\langle \alpha, \operatorname{Res}_\rho \beta \right\rangle_H\right) \\ &= \frac{1}{|H|} \sum_{h \in H} \alpha(h) (\operatorname{Res}_\rho \beta)\left(h^{-1}\right) \\ &\qquad (\text{here, we renamed the summation index } g \text{ as } h),\end{aligned}$$

this yields $\left\langle \operatorname{Ind}_\rho \alpha, \beta \right\rangle_G = \left\langle \alpha, \operatorname{Res}_\rho \beta \right\rangle_H$. Thus, (4.1.17) is proven.

[823] *Proof of (13.111.45):* Fix a $(h,k) \in H \times G$. Thus, $h \in H$ and $k \in G$. Now, there exists only one $g \in G$ such that $k\rho(h)k^{-1} = g$ (namely, $g = k\rho(h)k^{-1}$). Hence, the sum $\sum_{\substack{g \in G; \\ k\rho(h)k^{-1} = g}} \alpha(h) \beta\left(\left(k\rho(h) k^{-1}\right)^{-1}\right)$ has only one addend. Hence, this sum simplifies as follows:

$$\begin{aligned}&\sum_{\substack{g \in G; \\ k\rho(h)k^{-1} = g}} \alpha(h) \beta\left(\left(k\rho(h) k^{-1}\right)^{-1}\right) \\ &= \alpha(h) \beta\left(\underbrace{\left(k\rho(h) k^{-1}\right)^{-1}}_{= (k^{-1})^{-1} (\rho(h))^{-1} k^{-1}}\right) = \alpha(h) \beta\left(\underbrace{\left(k^{-1}\right)^{-1}}_{= k} \underbrace{(\rho(h))^{-1}}_{\substack{= \rho(h^{-1}) \\ (\text{since } \rho \text{ is a group} \\ \text{homomorphism})}} k^{-1}\right) = \alpha(h) \beta\left(k\rho\left(h^{-1}\right) k^{-1}\right).\end{aligned} \tag{13.111.46}$$

But $\rho\left(h^{-1}\right)$ and $k\rho\left(h^{-1}\right) k^{-1}$ are two conjugate elements of G. Hence, (13.111.43) (applied to $\rho\left(h^{-1}\right)$ and $k\rho\left(h^{-1}\right) k^{-1}$ instead of k and k') yields $\beta\left(\rho\left(h^{-1}\right)\right) = \beta\left(k\rho\left(h^{-1}\right) k^{-1}\right)$. Thus,

$$\beta\left(\left(k\rho(h) k^{-1}\right)^{-1}\right) = \beta\left(\rho\left(h^{-1}\right)\right) = \underbrace{(\beta \circ \rho)}_{\substack{= \operatorname{Res}_\rho \beta \\ (\text{since } \operatorname{Res}_\rho \beta \text{ is defined as } \beta \circ \rho)}} \left(h^{-1}\right) = (\operatorname{Res}_\rho \beta)\left(h^{-1}\right).$$

Hence, (13.111.46) becomes

$$\sum_{\substack{g \in G; \\ k\rho(h)k^{-1} = g}} \alpha(h) \beta\left(\left(k\rho(h) k^{-1}\right)^{-1}\right) = \alpha(h) \underbrace{\beta\left(k\rho\left(h^{-1}\right) k^{-1}\right)}_{= (\operatorname{Res}_\rho \beta)(h^{-1})} = \alpha(h) (\operatorname{Res}_\rho \beta)\left(h^{-1}\right).$$

This proves (13.111.45).

Let us now prove (4.1.16). (This proof will be very much similar to the proof of (4.1.17) above, but in a few places even easier.) The definition of $\left(\operatorname{Ind}_\rho \alpha, \beta\right)_G$ yields

$$\begin{aligned}
\left(\operatorname{Ind}_\rho \alpha, \beta\right)_G &= \frac{1}{|G|} \sum_{g \in G} \underbrace{\left(\operatorname{Ind}_\rho \alpha\right)(g)}_{\substack{= \frac{1}{|H|} \sum_{\substack{(h,k) \in H \times G; \\ k\rho(h)k^{-1} = g}} \alpha(h) \\ \text{(by the definition of } \operatorname{Ind}_\rho \alpha)}} \overline{\beta(g)} = \frac{1}{|G|} \sum_{g \in G} \left(\frac{1}{|H|} \sum_{\substack{(h,k) \in H \times G; \\ k\rho(h)k^{-1} = g}} \alpha(h) \right) \overline{\beta(g)} \\
&= \frac{1}{|G|} \frac{1}{|H|} \underbrace{\sum_{g \in G} \sum_{\substack{(h,k) \in H \times G; \\ k\rho(h)k^{-1} = g}}}_{= \sum_{(h,k) \in H \times G} \sum_{\substack{g \in G; \\ k\rho(h)k^{-1} = g}}} \alpha(h) \overline{\beta\left(\underbrace{g}_{\substack{= k\rho(h)k^{-1} \\ \text{(since } k\rho(h)k^{-1} = g)}} \right)} \\
&= \frac{1}{|G|} \frac{1}{|H|} \sum_{(h,k) \in H \times G} \sum_{\substack{g \in G; \\ k\rho(h)k^{-1} = g}} \alpha(h) \overline{\beta\left(k\rho(h)k^{-1}\right)}. \qquad (13.111.47)
\end{aligned}$$

However, for every $(h,k) \in H \times G$, we have

$$\sum_{\substack{g \in G; \\ k\rho(h)k^{-1} = g}} \alpha(h) \overline{\beta\left(k\rho(h)k^{-1}\right)} = \alpha(h) \overline{\left(\operatorname{Res}_\rho \beta\right)(h)} \tag{13.111.48}$$

[824]. Thus, (13.111.47) becomes

$$\begin{aligned}(\operatorname{Ind}_\rho \alpha, \beta)_G &= \frac{1}{|G|}\frac{1}{|H|} \underbrace{\sum_{(h,k)\in H\times G}}_{=\sum_{h\in H}\sum_{k\in G}} \underbrace{\sum_{\substack{g\in G;\\ k\rho(h)k^{-1}=g}} \alpha(h)\overline{\beta(k\rho(h)k^{-1})}}_{=\alpha(h)\overline{(\operatorname{Res}_\rho \beta)(h)}} \\ &= \frac{1}{|G|}\frac{1}{|H|}\sum_{h\in H}\underbrace{\sum_{k\in G}\alpha(h)\overline{(\operatorname{Res}_\rho\beta)(h)}}_{=|G|\cdot\alpha(h)\overline{(\operatorname{Res}_\rho\beta)(h)}} = \frac{1}{|G|}\frac{1}{|H|}\sum_{h\in H}|G|\cdot\alpha(h)\overline{(\operatorname{Res}_\rho\beta)(h)} \\ &= \frac{1}{|H|}\sum_{h\in H}\alpha(h)\overline{(\operatorname{Res}_\rho\beta)(h)}.\end{aligned}$$

Compared with

$$\begin{aligned}(\alpha, \operatorname{Res}_\rho\beta)_H &= \frac{1}{|H|}\sum_{g\in H}\alpha(g)\overline{(\operatorname{Res}_\rho\beta)(g)} \qquad \left(\text{by the definition of } (\alpha,\operatorname{Res}_\rho\beta)_H\right) \\ &= \frac{1}{|H|}\sum_{h\in H}\alpha(h)\overline{(\operatorname{Res}_\rho\beta)(h)} \\ &\qquad (\text{here, we renamed the summation index } g \text{ as } h),\end{aligned}$$

this yields $(\operatorname{Ind}_\rho\alpha,\beta)_G = (\alpha,\operatorname{Res}_\rho\beta)_H$. Thus, (4.1.16) is proven. The solution of Exercise 4.1.14(g) is thus complete.

(h) Let U be a $\mathbb{C}H$-module, and let V be a $\mathbb{C}G$-module.

Recall that $\operatorname{Ind}_\rho U$ is defined as the $\mathbb{C}G$-module $\mathbb{C}G\otimes_{\mathbb{C}H}U$, where $\mathbb{C}G$ is regarded as a $(\mathbb{C}G,\mathbb{C}H)$-bimodule according to the following rule: The left $\mathbb{C}G$-module structure on $\mathbb{C}G$ is plain multiplication inside $\mathbb{C}G$; the right $\mathbb{C}H$-module structure on $\mathbb{C}G$ is induced by the $\mathbb{C}$-algebra homomorphism $\mathbb{C}[\rho]:\mathbb{C}H\to\mathbb{C}G$ (thus, it is explicitly given by $\gamma\eta=\gamma\cdot(\mathbb{C}[\rho])\eta$ for all $\gamma\in\mathbb{C}G$ and $\eta\in\mathbb{C}H$). From now on, we regard $\mathbb{C}G$ as endowed with this $(\mathbb{C}G,\mathbb{C}H)$-bimodule structure.

Now, (4.1.8) (applied to $R=\mathbb{C}H$, $S=\mathbb{C}G$, $A=\mathbb{C}G$, $B=U$ and $C=V$) yields

$$\operatorname{Hom}_{\mathbb{C}G}(\mathbb{C}G\otimes_{\mathbb{C}H}U,V)\cong\operatorname{Hom}_{\mathbb{C}H}(U,\operatorname{Hom}_{\mathbb{C}G}(\mathbb{C}G,V)). \tag{13.111.50}$$

We shall now prove that $\operatorname{Hom}_{\mathbb{C}G}(\mathbb{C}G,V)\cong\operatorname{Res}_\rho V$ as left $\mathbb{C}H$-modules.

Indeed, a fundamental fact in abstract algebra says the following: If $\mathfrak{A}$ is a $\mathbb{C}$-algebra, and if $\mathfrak{M}$ is a left $\mathfrak{A}$-module, then there exists a $\mathbb{C}$-vector space isomorphism $\Xi:\operatorname{Hom}_{\mathfrak{A}}(\mathfrak{A},\mathfrak{M})\to\mathfrak{M}$ which satisfies $(\Xi(f)=f(1)$ for every $f\in\operatorname{Hom}_{\mathfrak{A}}(\mathfrak{A},\mathfrak{M}))$. Applying this fact to $\mathfrak{A}=\mathbb{C}G$ and $\mathfrak{M}=V$, we conclude that

[824] *Proof of (13.111.48):* Fix a $(h,k)\in H\times G$. Thus, $h\in H$ and $k\in G$. Now, there exists only one $g\in G$ such that $k\rho(h)k^{-1}=g$ (namely, $g=k\rho(h)k^{-1}$). Hence, the sum $\sum_{\substack{g\in G;\\ k\rho(h)k^{-1}=g}}\alpha(h)\overline{\beta(k\rho(h)k^{-1})}$ has only one addend. Hence, this sum simplifies as follows:

$$\sum_{\substack{g\in G;\\ k\rho(h)k^{-1}=g}}\alpha(h)\overline{\beta(k\rho(h)k^{-1})}=\alpha(h)\overline{\beta(k\rho(h)k^{-1})}. \tag{13.111.49}$$

But $\rho(h)$ and $k\rho(h)k^{-1}$ are two conjugate elements of G. Hence, (13.111.43) (applied to $\rho(h)$ and $k\rho(h)k^{-1}$ instead of k and k') yields $\beta(\rho(h))=\beta\left(k\rho(h)k^{-1}\right)$. Thus,

$$\beta\left(k\rho(h)k^{-1}\right)=\beta(\rho(h))=\underbrace{(\beta\circ\rho)}_{\substack{=\operatorname{Res}_\rho\beta\\ (\text{since } \operatorname{Res}_\rho\beta \text{ is defined as } \beta\circ\rho)}}(h)=(\operatorname{Res}_\rho\beta)(h).$$

Hence, (13.111.49) becomes

$$\sum_{\substack{g\in G;\\ k\rho(h)k^{-1}=g}}\alpha(h)\overline{\beta(k\rho(h)k^{-1})}=\alpha(h)\underbrace{\overline{\beta(k\rho(h)k^{-1})}}_{\substack{=\overline{(\operatorname{Res}_\rho\beta)(h)}\\ (\text{since } \beta\left(k\rho(h)k^{-1}\right)=(\operatorname{Res}_\rho\beta)(h))}}=\alpha(h)\overline{(\operatorname{Res}_\rho\beta)(h)}.$$

This proves (13.111.48).

there exists a $\mathbb{C}$-vector space isomorphism $\Xi : \operatorname{Hom}_{\mathbb{C}G}(\mathbb{C}G, V) \to V$ which satisfies $(\Xi(f) = f(1)$ for every $f \in \operatorname{Hom}_{\mathbb{C}G}(\mathbb{C}G, V))$. Consider this Ξ.

The map Ξ is a homomorphism of H-sets from $\operatorname{Hom}_{\mathbb{C}G}(\mathbb{C}G, V)$ to $\operatorname{Res}_\rho V$ [825]. Therefore, Ξ is a $\mathbb{C}H$-module homomorphism from $\operatorname{Hom}_{\mathbb{C}G}(\mathbb{C}G, V)$ to $\operatorname{Res}_\rho V$ (since Ξ is a $\mathbb{C}$-linear map). Consequently, Ξ is a $\mathbb{C}H$-module isomorphism from $\operatorname{Hom}_{\mathbb{C}G}(\mathbb{C}G, V)$ to $\operatorname{Res}_\rho V$ (since Ξ is a $\mathbb{C}$-vector space isomorphism). Therefore, $\operatorname{Hom}_{\mathbb{C}G}(\mathbb{C}G, V) \cong \operatorname{Res}_\rho V$ as $\mathbb{C}H$-modules. Now,

$$\begin{aligned}
\operatorname{Hom}_{\mathbb{C}G}\left(\underbrace{\operatorname{Ind}_\rho U}_{=\mathbb{C}G\otimes_{\mathbb{C}H}U}, V\right) &= \operatorname{Hom}_{\mathbb{C}G}(\mathbb{C}G \otimes_{\mathbb{C}H} U, V) \\
&\cong \operatorname{Hom}_{\mathbb{C}H}\left(U, \underbrace{\operatorname{Hom}_{\mathbb{C}G}(\mathbb{C}G, V)}_{\cong \operatorname{Res}_\rho V \text{ as } \mathbb{C}H\text{-modules}}\right) \qquad (\text{by } (13.111.50)) \\
&\cong \operatorname{Hom}_{\mathbb{C}H}(U, \operatorname{Res}_\rho V).
\end{aligned}$$

This solves Exercise 4.1.14(h).

(i) The following solution will mostly be an imitation of the solution of Exercise 4.1.4.

Let U be any $\mathbb{C}H$-module. Recall that $\operatorname{Ind}_\rho U$ is defined as the $\mathbb{C}G$-module $\mathbb{C}G \otimes_{\mathbb{C}H} U$, where $\mathbb{C}G$ is regarded as a $(\mathbb{C}G, \mathbb{C}H)$-bimodule according to the following rule: The left $\mathbb{C}G$-module structure on $\mathbb{C}G$ is plain multiplication inside $\mathbb{C}G$; the right $\mathbb{C}H$-module structure on $\mathbb{C}G$ is induced by the $\mathbb{C}$-algebra homomorphism $\mathbb{C}[\rho] : \mathbb{C}H \to \mathbb{C}G$ (thus, it is explicitly given by $\gamma\eta = \gamma \cdot (\mathbb{C}[\rho])\eta$ for all $\gamma \in \mathbb{C}G$ and $\eta \in \mathbb{C}H$). From now on, we regard $\mathbb{C}G$ as endowed with this $(\mathbb{C}G, \mathbb{C}H)$-bimodule structure (besides the $(\mathbb{C}H, \mathbb{C}G)$-bimodule structure that was introduced in the statement of Exercise 4.1.14(i)).

The $\mathbb{C}$-vector space $\mathbb{C}G$ is thus endowed with a left $\mathbb{C}H$-module structure (which is part of the $(\mathbb{C}G, \mathbb{C}H)$-bimodule structure) and with a right $\mathbb{C}H$-module structure (which is part of the $(\mathbb{C}H, \mathbb{C}G)$-bimodule structure). These two structures, combined, form a $(\mathbb{C}H, \mathbb{C}H)$-bimodule structure[826]. This allows us to write expressions like xyz with $x \in \mathbb{C}H$, $y \in \mathbb{C}G$ and $z \in \mathbb{C}H$, without having to disambiguate whether they mean $(xy)z$ or $x(yz)$.

We recall that $\operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, U)$ is the left $\mathbb{C}G$-module consisting of all left $\mathbb{C}H$-module homomorphisms from $\mathbb{C}G$ to U. This uses only the $(\mathbb{C}H, \mathbb{C}G)$-bimodule structure on $\mathbb{C}G$ that was introduced in the statement of Exercise 4.1.14(i) (but not the $(\mathbb{C}G, \mathbb{C}H)$-bimodule structure on $\mathbb{C}G$ that was introduced in the definition of $\operatorname{Ind}_\rho U$).

[825] *Proof.* The map Ξ is clearly a map from $\operatorname{Hom}_{\mathbb{C}G}(\mathbb{C}G, V)$ to V, therefore a map from $\operatorname{Hom}_{\mathbb{C}G}(\mathbb{C}G, V)$ to $\operatorname{Res}_\rho V$ (since $\operatorname{Res}_\rho V = V$ as sets).

Recall that the $\mathbb{C}H$-module structure on $\operatorname{Res}_\rho V$ is given by

$$h \cdot v = \rho(h) \cdot v \qquad \text{for every } h \in H \text{ and } v \in V. \tag{13.111.51}$$

Now, let $f \in \operatorname{Hom}_{\mathbb{C}G}(\mathbb{C}G, V)$ and $h \in H$. We are going to prove that $\Xi(h \cdot f) = h \cdot \Xi(f)$, where $h \cdot \Xi(f)$ is computed in the $\mathbb{C}H$-module $\operatorname{Res}_\rho V$.

Indeed, the definition of the left $\mathbb{C}H$-module structure on $\operatorname{Hom}_{\mathbb{C}G}(\mathbb{C}G, V)$ yields $(\eta \cdot \alpha)(p) = \alpha(p\eta)$ for every $\eta \in \mathbb{C}H$, $\alpha \in \operatorname{Hom}_{\mathbb{C}G}(\mathbb{C}G, V)$ and $p \in \mathbb{C}G$. Applying this to $\eta = h$, $\alpha = f$ and $p = 1$, we obtain $(h \cdot f)(1) = f(1h)$ (since $h \in H \subset \mathbb{C}H$). But the definition of Ξ yields $\Xi(h \cdot f) = (h \cdot f)(1) = f(1h)$.

Recall that the right $\mathbb{C}H$-module structure on $\mathbb{C}G$ is given by the equality $\gamma\eta = \gamma \cdot (\mathbb{C}[\rho])\eta$ for all $\gamma \in \mathbb{C}G$ and $\eta \in \mathbb{C}H$. Applying this equality to $\gamma = 1$ and $\eta = h$, we obtain $1h = 1 \cdot (\mathbb{C}[\rho])(h)$ (since $h \in H \subset \mathbb{C}H$). But $h \in H$ and thus $(\mathbb{C}[\rho])(h) = \rho(h)$. Thus, $1h = 1 \cdot (\mathbb{C}[\rho])(h) = (\mathbb{C}[\rho])(h) = \rho(h) = \rho(h) \cdot 1$. Now, $\Xi(h \cdot f) = f\left(\underbrace{1h}_{=\rho(h)\cdot 1}\right) = f(\rho(h) \cdot 1) = \rho(h) \cdot f(1)$ (since f is a $\mathbb{C}G$-module homomorphism (because $f \in \operatorname{Hom}_{\mathbb{C}G}(\mathbb{C}G, V)$) and since $\rho(h) \in G \subset \mathbb{C}G$).

On the other hand, (13.111.51) (applied to $v = f(1)$) yields $h \cdot f(1) = \rho(h) \cdot f(1)$. Compared with $\Xi(h \cdot f) = \rho(h) \cdot f(1)$, this yields $\Xi(h \cdot f) = h \cdot f(1)$.

But the definition of Ξ yields $\Xi(f) = f(1)$. Thus, $f(1) = \Xi(f)$, so that $\Xi(h \cdot f) = h \cdot \underbrace{f(1)}_{=\Xi(f)} = h \cdot \Xi(f)$.

Now, let us forget that we fixed f and h. We thus have shown that $\Xi(h \cdot f) = h \cdot \Xi(f)$ for every $f \in \operatorname{Hom}_{\mathbb{C}G}(\mathbb{C}G, V)$ and $h \in H$. In other words, Ξ is a homomorphism of H-sets from $\operatorname{Hom}_{\mathbb{C}G}(\mathbb{C}G, V)$ to $\operatorname{Res}_\rho V$, qed.

[826] This is easy to check (we leave the details of this verification to the reader).

Recall that the right $\mathbb{C}H$-module structure on $\mathbb{C}G$ is given by

$$\gamma\eta = \gamma \cdot \left(\mathbb{C}\left[\rho\right]\right)\eta \qquad \text{for all } \gamma \in \mathbb{C}G \text{ and } \eta \in \mathbb{C}H. \tag{13.111.52}$$

For every $\gamma \in \mathbb{C}G$ and $\eta \in H$, we have

$$\begin{aligned} \gamma\eta &= \gamma \cdot \underbrace{\left(\mathbb{C}\left[\rho\right]\right)\eta}_{\substack{=\rho(\eta) \\ (\text{since } \eta\in H)}} \qquad (\text{by (13.111.52) (since } \eta \in H \subset \mathbb{C}H)) \\ &= \gamma \cdot \rho\left(\eta\right). \end{aligned} \tag{13.111.53}$$

On the other hand, the left $\mathbb{C}H$-module structure on $\mathbb{C}G$ is induced by the $\mathbb{C}$-algebra homomorphism $\mathbb{C}\left[\rho\right] : \mathbb{C}H \to \mathbb{C}G$. In other words, this structure is given by

$$\eta\gamma = \left(\mathbb{C}\left[\rho\right]\right)\eta \cdot \gamma \qquad \text{for all } \gamma \in \mathbb{C}G \text{ and } \eta \in \mathbb{C}H. \tag{13.111.54}$$

For every $\gamma \in \mathbb{C}G$ and $\eta \in H$, we have

$$\begin{aligned} \eta\gamma &= \underbrace{\left(\mathbb{C}\left[\rho\right]\right)\eta}_{\substack{=\rho(\eta) \\ (\text{since } \eta\in H)}} \cdot\gamma \qquad (\text{by (13.111.54) (since } \eta \in H \subset \mathbb{C}H)) \\ &= \rho\left(\eta\right)\cdot\gamma. \end{aligned} \tag{13.111.55}$$

The map ρ is a group homomorphism. Hence, $\rho\left(H\right)$ is a subgroup of G. Let us denote this subgroup by $\overline{H}$. Thus, $\overline{H} = \rho\left(H\right)$.

Let J be a system of distinct representatives for the right $\overline{H}$-cosets in G. Then, $G = \bigsqcup_{j\in J} \overline{H}j$.

For every $g \in G$, define a map $\mathfrak{R}_g : J \to J$ as follows: Let $i \in J$. Then, $ig \in G = \bigsqcup_{j\in J} \overline{H}j$. Thus, there exists a unique $j \in J$ such that $ig \in \overline{H}j$. Define $\mathfrak{R}_g\left(i\right)$ to be this j. Hence, we have defined $\mathfrak{R}_g\left(i\right)$ for every $i \in J$. Thus, we have defined a map $\mathfrak{R}_g : J \to J$.

For every $g \in G$ and $i \in J$, we have

$$ig \in \overline{H} \cdot \mathfrak{R}_g\left(i\right) \tag{13.111.56}$$

(because $\mathfrak{R}_g\left(i\right)$ is defined as the $j \in J$ satisfying $ig \in \overline{H}j$).

It is easy to see that for every $g \in G$, the map $\mathfrak{R}_g : J \to J$ is a bijection.[827] We have

$$j^{-1} \otimes_{\mathbb{C}H} f\left(jg\right) = g \cdot \left(\mathfrak{R}_g\left(j\right)\right)^{-1} \otimes_{\mathbb{C}H} f\left(\mathfrak{R}_g\left(j\right)\right) \qquad \text{in } \mathbb{C}G \otimes_{\mathbb{C}H} U \tag{13.111.58}$$

[827] *Proof.* Fix $g \in G$.

The sets $\overline{H}j$ for all $j \in J$ are disjoint (since $\bigsqcup_{j\in J} \overline{H}j$ is well-defined). In other words, if i and i' are two elements of J such that $\overline{H}i$ and $\overline{H}i'$ are not disjoint, then

$$i = i'. \tag{13.111.57}$$

Let us now prove that $\mathfrak{R}_g$ is injective.

Let i and i' be two elements of J. Assume that $\mathfrak{R}_g\left(i\right) = \mathfrak{R}_g\left(i'\right)$. We will show that $i = i'$.

From (13.111.56), we have $ig \in \overline{H} \cdot \mathfrak{R}_g\left(i\right)$. Hence, there exists some $h \in \overline{H}$ such that $ig = h \cdot \mathfrak{R}_g\left(i\right)$. Consider this h. We have

$$\overline{H} \underbrace{ig}_{=h\cdot\mathfrak{R}_g(i)} = \underbrace{\overline{H}h}_{\substack{=\overline{H} \\ (\text{since } h\in\overline{H} \text{ and since} \\ \overline{H} \text{ is a group})}} \cdot\mathfrak{R}_g\left(i\right) = \overline{H}\cdot\mathfrak{R}_g\left(i\right).$$

The same argument (but for i' instead of i) yields $\overline{H}i'g = \overline{H}\cdot\mathfrak{R}_g\left(i'\right)$. Hence, $\overline{H}ig = \overline{H}\cdot\underbrace{\mathfrak{R}_g\left(i\right)}_{=\mathfrak{R}_g(i')} = \overline{H}\cdot\mathfrak{R}_g\left(i'\right) = \overline{H}i'g$. Thus, $\underbrace{\overline{H}ig}_{=\overline{H}i'g} g^{-1} = \overline{H}i'gg^{-1} = \overline{H}i'$, so that $\overline{H}i' = \overline{H}igg^{-1} = \overline{H}i$. Thus, the sets $\overline{H}i$ and $\overline{H}i'$ are not disjoint (because $\left(\overline{H}i\right) \cap \underbrace{\left(\overline{H}i'\right)}_{=\overline{H}i} = \left(\overline{H}i\right) \cap \left(\overline{H}i\right) = \overline{H}i \neq \varnothing$). Therefore, $i = i'$ (by (13.111.57)).

Now, forget that we fixed i and i'. We thus have shown that if i and i' are two elements of J such that $\mathfrak{R}_g\left(i\right) = \mathfrak{R}_g\left(i'\right)$, then $i = i'$. In other words, the map $\mathfrak{R}_g$ is injective. But the set J is finite (since it is a subset of the finite set G). Hence, $\mathfrak{R}_g$ is a map from a finite set (namely, J) to itself. Since we know that $\mathfrak{R}_g$ is injective, this yields that $\mathfrak{R}_g$ is surjective (because any injective map from a finite set to itself must be surjective). Hence, $\mathfrak{R}_g$ is bijective (since $\mathfrak{R}_g$ is injective and surjective), that is, a bijection, qed.

for every $g \in G$, every $j \in J$ and every $f \in \operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, U)$. [828]

We now define a map $\alpha : \operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, U) \to \mathbb{C}G \otimes_{\mathbb{C}H} U$ by setting

$$\alpha(f) = \sum_{j \in J} j^{-1} \otimes_{\mathbb{C}H} f(j) \qquad \text{for all } f \in \operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, U).$$

This α is a map $\operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, U) \to \operatorname{Ind}_\rho U$ (since $\operatorname{Ind}_\rho U = \mathbb{C}G \otimes_{\mathbb{C}H} U$).

We will show that α is a $\mathbb{C}G$-module isomorphism.

First, let us prove that α is a left $\mathbb{C}G$-module homomorphism. In fact, any $f \in \operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, U)$ and $g \in G$ satisfy

$$\begin{aligned}
\alpha(gf) &= \sum_{j \in J} j^{-1} \otimes_{\mathbb{C}H} \underbrace{(gf)(j)}_{\substack{=f(jg) \\ \text{(by the definition of } gf)}} \qquad (\text{by the definition of } \alpha(gf)) \\
&= \sum_{j \in J} \underbrace{j^{-1} \otimes_{\mathbb{C}H} f(jg)}_{\substack{= g \cdot (\mathfrak{R}_g(j))^{-1} \otimes_{\mathbb{C}H} f(\mathfrak{R}_g(j)) \\ \text{(by (13.111.58))}}} = \sum_{j \in J} g \cdot (\mathfrak{R}_g(j))^{-1} \otimes_{\mathbb{C}H} f(\mathfrak{R}_g(j)) = \sum_{j \in J} g \cdot j^{-1} \otimes_{\mathbb{C}H} f(j) \\
&\qquad \left(\begin{array}{c} \text{here, we have substituted } j \text{ for } \mathfrak{R}_g(j) \text{ in the sum, since} \\ \text{the map } \mathfrak{R}_g : J \to J \text{ is a bijection} \end{array}\right) \\
&= g \cdot \underbrace{\sum_{j \in J} j^{-1} \otimes_{\mathbb{C}H} f(j)}_{=\alpha(f)} = g \cdot \alpha(f).
\end{aligned}$$

The map α is thus a homomorphism of left G-sets. Since α is furthermore $\mathbb{C}$-linear, this yields that α is a left $\mathbb{C}G$-module homomorphism.

We now are going to construct an inverse for α. This will be more cumbersome.

For every $g \in G$ and every $p \in \mathbb{C}G$, we denote by $\epsilon_g(p)$ the g-coordinate of p with respect to the basis G of the $\mathbb{C}$-vector space $\mathbb{C}G$. By the definition of "coordinate", we have

$$q = \sum_{g \in G} \epsilon_g(q) g \qquad \text{for every } q \in \mathbb{C}G. \tag{13.111.60}$$

For every $g \in G$, we have defined a map $\epsilon_g : \mathbb{C}G \to \mathbb{C}$ (because we have defined an element $\epsilon_g(p)$ for every $p \in \mathbb{C}G$). This map ϵ_g is $\mathbb{C}$-linear. We record some properties of these maps:

[828] *Proof of (13.111.58):* Let $g \in G$, $j \in J$ and $f \in \operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, U)$. Applying (13.111.56) to $i = j$, we obtain $jg \in \overline{H} \cdot \mathfrak{R}_g(j)$. In other words, there exists an $h \in \overline{H}$ such that $jg = h \cdot \mathfrak{R}_g(j)$. Consider this h. We have

$$j^{-1} \underbrace{jg}_{=h \cdot \mathfrak{R}_g(j)} \cdot (\mathfrak{R}_g(j))^{-1} = j^{-1} h \cdot \underbrace{\mathfrak{R}_g(j) \cdot (\mathfrak{R}_g(j))^{-1}}_{=1} = j^{-1} h,$$

hence

$$j^{-1} h = j^{-1} jg \cdot (\mathfrak{R}_g(j))^{-1} = g \cdot (\mathfrak{R}_g(j))^{-1}.$$

But $h \in \overline{H} = \rho(H)$. Hence, there exists a $h' \in H$ such that $h = \rho(h')$. Consider this h'.

Applying (13.111.55) to $\eta = h'$ and $\gamma = \mathfrak{R}_g(j)$, we obtain $h' \cdot \mathfrak{R}_g(j) = \underbrace{\rho(h')}_{=h} \cdot \mathfrak{R}_g(j) = h \cdot \mathfrak{R}_g(j) = jg$ (since $jg = h \cdot \mathfrak{R}_g(j)$).

But the map f is left $\mathbb{C}H$-linear (since $f \in \operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, U)$). Thus, $f(h' \cdot \mathfrak{R}_g(j)) = h' \cdot f(\mathfrak{R}_g(j))$ (since $h' \in H \subset \mathbb{C}H$). Since $h' \cdot \mathfrak{R}_g(j) = jg$, this rewrites as $f(jg) = h' \cdot f(\mathfrak{R}_g(j))$. Hence,

$$j^{-1} \otimes_{\mathbb{C}H} \underbrace{f(jg)}_{=h' \cdot f(\mathfrak{R}_g(j))} = j^{-1} \otimes_{\mathbb{C}H} h' \cdot f(\mathfrak{R}_g(j)) = j^{-1} h' \otimes_{\mathbb{C}H} f(\mathfrak{R}_g(j)) \tag{13.111.59}$$

(here, we have moved h' past the $\otimes_{\mathbb{C}H}$ sign, since $h' \in H \subset \mathbb{C}H$).

On the other hand, applying (13.111.53) to $\gamma = j^{-1}$ and $\eta = h'$, we obtain $j^{-1} h' = j^{-1} \cdot \underbrace{\rho(h')}_{=h} = j^{-1} h = g \cdot (\mathfrak{R}_g(j))^{-1}$.

Hence, (13.111.59) becomes

$$j^{-1} \otimes_{\mathbb{C}H} f(jg) = \underbrace{j^{-1} h'}_{=g \cdot (\mathfrak{R}_g(j))^{-1}} \otimes_{\mathbb{C}H} f(\mathfrak{R}_g(j)) = g \cdot (\mathfrak{R}_g(j))^{-1} \otimes_{\mathbb{C}H} f(\mathfrak{R}_g(j)).$$

This proves (13.111.58).

- For every $g \in G$ and $h \in G$, we have

$$\epsilon_g(h) = \delta_{g,h}. \tag{13.111.61}$$

[829]

- We have

$$\epsilon_1(pq) = \epsilon_1(qp) \qquad \text{for all } p \in \mathbb{C}G \text{ and } q \in \mathbb{C}G. \tag{13.111.62}$$

[830]

- We have

$$\epsilon_1\left(g^{-1}q\right) = \epsilon_g(q) \qquad \text{for every } g \in G \text{ and } q \in \mathbb{C}G. \tag{13.111.63}$$

[831]

Now, fix $q \in \mathbb{C}G$ and $u \in U$. We let $f_{q,u}$ be the map $\mathbb{C}G \to U$ defined by

$$f_{q,u}(p) = \sum_{h\in H} \epsilon_1(\rho(h)pq)h^{-1}u \qquad \text{for every } p \in \mathbb{C}G. \tag{13.111.64}$$

It is obvious that this map $f_{q,u}$ is $\mathbb{C}$-linear. We will show that $f_{q,u}$ is a left $\mathbb{C}H$-module homomorphism.

The map $f_{q,u}$ is a homomorphism of left H-sets[832]. Since $f_{q,u}$ is furthermore $\mathbb{C}$-linear, this yields that $f_{q,u}$ is a left $\mathbb{C}H$-module homomorphism. Hence, $f_{q,u} \in \operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, U)$.

Now, forget that we fixed q and u. We thus have defined a map $f_{q,u} \in \operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, U)$ for every $q \in \mathbb{C}G$ and $u \in U$. It is easy to see that this map $f_{q,u}$ depends $\mathbb{C}$-linearly on each of q and u. Now, define a map $\widetilde{\beta} : \mathbb{C}G \times U \to \operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, U)$ by

$$\widetilde{\beta}(q,u) = f_{q,u} \qquad \text{for every } (q,u) \in \mathbb{C}G \times U.$$

[829] This equality has been proven in our solution of Exercise 4.1.4. (Namely, it appeared there as (13.105.3).)

[830] This equality has been proven in our solution of Exercise 4.1.4. (Namely, it appeared there as (13.105.4).)

[831] This equality has been proven in our solution of Exercise 4.1.4. (Namely, it appeared there as (13.105.5).)

[832] *Proof.* In fact, for every $h' \in H$ and every $p \in \mathbb{C}G$, we have

$$f_{q,u}(h'p) = \sum_{h\in H} \epsilon_1\left(\rho(h) \underbrace{(h'p)}_{\substack{=\rho(h')\cdot p \\ (\text{by (13.111.55), applied} \\ \text{to } \gamma=p \text{ and } \eta=h')}} q\right) h^{-1}u \qquad (\text{by the definition of } f_{q,u})$$

$$= \sum_{h\in H} \epsilon_1(\rho(h)\rho(h')\cdot pq)h^{-1}u = \sum_{h\in H} \epsilon_1\left(\underbrace{\rho\left(h(h')^{-1}\right)}_{\substack{=\rho(h)(\rho(h'))^{-1} \\ (\text{since } \rho \text{ is a group} \\ \text{homomorphism})}} \rho(h')\cdot pq\right) \underbrace{\left(h(h')^{-1}\right)^{-1}}_{=((h')^{-1})^{-1}h^{-1}=h'h^{-1}} u$$

$$\left(\begin{array}{c}\text{here, we have substituted } h(h')^{-1} \text{ for } h \text{ in the sum,} \\ \text{because the map } H \to H,\ h \mapsto h(h')^{-1} \text{ is a bijection} \\ (\text{since } H \text{ is a group and since } h' \in H)\end{array}\right)$$

$$= \sum_{h\in H} \epsilon_1\left(\rho(h) \underbrace{(\rho(h'))^{-1}\rho(h')}_{=1} \cdot pq\right) h'h^{-1}u$$

$$= \sum_{h\in H} \epsilon_1(\rho(h)pq)h'h^{-1}u = h' \cdot \underbrace{\sum_{h\in H} \epsilon_1(\rho(h)pq)h^{-1}u}_{\substack{=f_{q,u}(p) \\ (\text{by (13.111.64)})}} = h' \cdot f_{q,u}(p).$$

In other words, $f_{q,u}$ is a homomorphism of left H-sets, qed.

Then, $\widetilde{\beta}$ is a $\mathbb{C}$-bilinear map (because $\widetilde{\beta}(q,u) = f_{q,u}$ depends $\mathbb{C}$-linearly on each of q and u). We are now going to prove that the map $\widetilde{\beta}$ is $\mathbb{C}H$-bilinear with respect to the right $\mathbb{C}H$-module structure on $\mathbb{C}G$ and the left $\mathbb{C}H$-module structure on U.

In fact, every $h' \in H$, $q \in \mathbb{C}G$ and $u \in U$ satisfy

$$\widetilde{\beta}(q, h'u) = \widetilde{\beta}(qh', u). \tag{13.111.65}$$

[833] As a consequence of this, we can see that every $r \in \mathbb{C}H$, $q \in \mathbb{C}G$ and $u \in U$ satisfy

$$\widetilde{\beta}(q, ru) = \widetilde{\beta}(qr, u). \tag{13.111.67}$$

[834] In other words, the map $\widetilde{\beta}$ is $\mathbb{C}H$-bilinear with respect to the right $\mathbb{C}H$-module structure on $\mathbb{C}G$ and the left $\mathbb{C}H$-module structure on U (since we already know that $\widetilde{\beta}$ is $\mathbb{C}$-bilinear). Hence, by the universal property of the tensor product, we conclude that there exists a unique $\mathbb{C}$-linear map $\beta : \mathbb{C}G \otimes_{\mathbb{C}H} U \to \operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, U)$ such that every $(q,u) \in \mathbb{C}G \times U$ satisfies

$$\beta(q \otimes_{\mathbb{C}H} u) = \widetilde{\beta}(q,u). \tag{13.111.68}$$

Consider this map β. Clearly, every $q \in \mathbb{C}G$ and $u \in U$ satisfy

$$\beta(q \otimes_{\mathbb{C}H} u) = \widetilde{\beta}(q,u) = f_{q,u} \qquad \left(\text{by the definition of } \widetilde{\beta}\right).$$

[833] *Proof of (13.111.65):* Let $h' \in H$, $q \in \mathbb{C}G$ and $u \in U$. Then, the map $H \to H,\ h \mapsto h'h$ is a bijection (since H is a group). Now, let $p \in \mathbb{C}G$. The definition of $\widetilde{\beta}$ yields $\widetilde{\beta}(q,h'u) = f_{q,h'u}$. Hence,

$$\begin{aligned}
\Bigl(\underbrace{\widetilde{\beta}(q,h'u)}_{=f_{q,h'u}}\Bigr)(p) &= f_{q,h'u}(p) = \sum_{h\in H} \epsilon_1(\rho(h)pq)\,h^{-1}h'u \qquad \left(\text{by the definition of } f_{q,h'u}\right) \\
&= \sum_{h\in H} \epsilon_1\Bigl(\underbrace{\rho(h'h)}_{\substack{=\rho(h')\rho(h) \\ \text{(since } \rho \text{ is a group} \\ \text{homomorphism)}}} pq\Bigr) \underbrace{(h'h)^{-1}}_{=h^{-1}(h')^{-1}} h'u \\
&\qquad \left(\begin{array}{c}\text{here, we substituted } h'h \text{ for } h \text{ in the sum,} \\ \text{since the map } H \to H,\ h \mapsto h'h \text{ is a bijection}\end{array}\right) \\
&= \sum_{h\in H} \underbrace{\epsilon_1(\rho(h')\rho(h)pq)}_{\substack{=\epsilon_1(\rho(h)pq\rho(h')) \\ \text{(by (13.111.62), applied to} \\ \rho(h') \text{ and } \rho(h)pq \text{ instead of } p \text{ and } q)}} h^{-1} \underbrace{(h')^{-1}h'}_{=1} u \\
&= \sum_{h\in H} \epsilon_1(\rho(h)pq\rho(h'))\,h^{-1}u. \qquad\qquad (13.111.66)
\end{aligned}$$

But the definition of $\widetilde{\beta}$ also yields $\widetilde{\beta}(qh',u) = f_{qh',u}$. Hence,

$$\begin{aligned}
\left(\widetilde{\beta}(qh',u)\right)(p) &= f_{qh',u}(p) = \sum_{h\in H} \epsilon_1\Bigl(\rho(h)p \underbrace{(qh')}_{\substack{=q\cdot\rho(h') \\ \text{(by (13.111.53), applied to} \\ \gamma=q \text{ and } \eta=h')}}\Bigr) h^{-1}u \qquad \left(\text{by the definition of } f_{qh',u}\right) \\
&= \sum_{h\in H} \epsilon_1(\rho(h)pq\cdot\rho(h'))\,h^{-1}u = \sum_{h\in H} \epsilon_1(\rho(h)pq\rho(h'))\,h^{-1}u = \left(\widetilde{\beta}(q,h'u)\right)(p) \qquad (\text{by (13.111.66)}).
\end{aligned}$$

Now, forget that we fixed p. We have thus proven that $\left(\widetilde{\beta}(qh',u)\right)(p) = \left(\widetilde{\beta}(q,h'u)\right)(p)$ for every $p \in \mathbb{C}G$. In other words, $\widetilde{\beta}(qh',u) = \widetilde{\beta}(q,h'u)$, so that $\widetilde{\beta}(q,h'u) = \widetilde{\beta}(qh',u)$. This proves (13.111.65).

[834] *Proof of (13.111.67):* Let $r \in \mathbb{C}H$, $q \in \mathbb{C}G$ and $u \in U$. We need to prove the equality (13.111.67). But this equality is $\mathbb{C}$-linear in r. Hence, we can WLOG assume that r belongs to the basis H of the $\mathbb{C}$-vector space $\mathbb{C}H$. Assume this. Then, (13.111.67) follows from (13.111.65) (applied to $h' = r$).

Hence, every $q \in \mathbb{C}G$, $u \in U$ and $p \in \mathbb{C}G$ satisfy

$$(13.111.69) \qquad \Bigg(\underbrace{\beta\left(q \otimes_{\mathbb{C}H} u\right)}_{=f_{q,u}}\Bigg)(p) = f_{q,u}(p) = \sum_{h \in H} \epsilon_1\left(\rho(h)\, pq\right) h^{-1} u$$

(by (13.111.64)).

Let $K = \ker \rho$. Thus, K is the kernel of a group homomorphism out of H (since ρ is a group homomorphism out of H), and therefore a normal subgroup of H.

We notice that

$$(13.111.70) \qquad \left|\rho^{-1}(g)\right| = |K| \qquad \text{for every } g \in \overline{H}.$$

[835]

We shall now show that the maps α and $\dfrac{1}{|K|}\beta$ are mutually inverse. To do so, we will show that $\alpha \circ \beta = |K| \operatorname{id}$ and $\beta \circ \alpha = |K| \operatorname{id}$.

[835] *Proof of (13.111.70):* Let $g \in \overline{H}$. Then, $g \in \overline{H} = \rho(H)$. Hence, $g = \rho(x)$ for some $x \in H$. Let us fix such a x.

We know that H is a group. Hence, the map $K \to xK$, $k \mapsto xk$ is a bijection. Thus, the sets K and xK are in bijection. Therefore, $|xK| = |K|$.

However, for every $y \in H$, we have the following logical equivalence:

$$\left(y \in \rho^{-1}(g)\right) \iff \Bigg(\rho(y) = \underbrace{g}_{=\rho(x)}\Bigg) \iff \left(\rho(y) = \rho(x)\right) \iff \Bigg(\underbrace{\rho(y) \cdot \left(\rho(x)\right)^{-1}}_{\substack{=\rho\left(yx^{-1}\right) \\ \text{(since } \rho \text{ is a group} \\ \text{homomorphism)}}} = 1\Bigg)$$
$$\iff \left(\rho\left(yx^{-1}\right) = 1\right) \iff \Bigg(yx^{-1} \in \underbrace{\ker \rho}_{=K}\Bigg) \iff \left(yx^{-1} \in K\right) \iff \left(y \in xK\right).$$

Hence, $\rho^{-1}(g) = xK$, so that $\left|\rho^{-1}(g)\right| = |xK| = |K|$. This proves (13.111.70).

Let us first show that $\alpha \circ \beta = \mathrm{id}$. In fact, let $q \in \mathbb{C}G$ and $u \in U$ be arbitrary. Then,

$$\begin{aligned}
\left(\alpha \circ \beta\right)\left(q \otimes_{\mathbb{C}H} u\right) &= \alpha\left(\beta\left(q \otimes_{\mathbb{C}H} u\right)\right) = \sum_{j \in J} j^{-1} \otimes_{\mathbb{C}H} \underbrace{\left(\beta\left(q \otimes_{\mathbb{C}H} u\right)\right)(j)}_{\substack{=\sum\limits_{h\in H} \epsilon_1(\rho(h)jq)h^{-1}u \\ \text{(by (13.111.69), applied to } p=j\text{)}}} \\
&\qquad \text{(by the definition of } \alpha\text{)} \\
&= \sum_{j \in J} j^{-1} \otimes_{\mathbb{C}H} \left(\sum_{h \in H} \epsilon_1\left(\rho\left(h\right) jq\right) h^{-1} u\right) = \sum_{j \in J} \sum_{h \in H} \epsilon_1\left(\rho\left(h\right) jq\right) \underbrace{j^{-1} \otimes_{\mathbb{C}H} h^{-1} u}_{\substack{=j^{-1}h^{-1} \otimes_{\mathbb{C}H} u \\ \text{(since } h^{-1} \in \mathbb{C}H \text{ (since } h \in H \\ \text{and thus } h^{-1} \in H \subset \mathbb{C}H\text{))}}} \\
&= \sum_{j \in J} \sum_{h \in H} \epsilon_1\left(\rho\left(h\right) jq\right) \underbrace{j^{-1} h^{-1}}_{\substack{=j^{-1} \cdot \rho\left(h^{-1}\right) \\ \text{(by (13.111.53), applied to } \gamma=j^{-1} \\ \text{and } \eta=h^{-1} \text{ (since } h^{-1} \in H \text{ (since } h \in H\text{)))}}} \otimes_{\mathbb{C}H} u \\
&= \sum_{j \in J} \underbrace{\sum_{h \in H}}_{\substack{=\sum_{g \in \overline{H}} \sum\limits_{\substack{h \in H; \\ \rho(h)=g}} \\ \text{(since every } h \in H \text{ satisfies} \\ \rho(h) \in \rho(H) = \overline{H}\text{)}}} \epsilon_1\left(\rho\left(h\right) jq\right) j^{-1} \cdot \underbrace{\rho\left(h^{-1}\right)}_{\substack{=(\rho(h))^{-1} \\ \text{(since } \rho \text{ is a group} \\ \text{homomorphism)}}} \otimes_{\mathbb{C}H} u \\
&= \sum_{j \in J} \sum_{g \in \overline{H}} \underbrace{\sum_{\substack{h \in H; \\ \rho(h)=g}}}_{=\sum_{h \in \rho^{-1}(g)}} \epsilon_1\left(\underbrace{\rho\left(h\right)}_{=g} jq\right) j^{-1} \cdot \left(\underbrace{\rho\left(h\right)}_{=g}\right)^{-1} \otimes_{\mathbb{C}H} u \\
&= \sum_{j \in J} \sum_{g \in \overline{H}} \underbrace{\sum_{h \in \rho^{-1}(g)} \epsilon_1\left(gjq\right) j^{-1} \cdot g^{-1} \otimes_{\mathbb{C}H} u}_{=\left|\rho^{-1}(g)\right| \cdot \epsilon_1(gjq) j^{-1} \cdot g^{-1} \otimes_{\mathbb{C}H} u} \\
&= \sum_{j \in J} \sum_{g \in \overline{H}} \underbrace{\left|\rho^{-1}\left(g\right)\right|}_{\substack{=|K| \\ \text{(by (13.111.70))}}} \cdot \epsilon_1\left(gjq\right) \underbrace{j^{-1} \cdot g^{-1}}_{=(gj)^{-1}} \otimes_{\mathbb{C}H} u
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{j\in J}\sum_{g\in\overline{H}} |K|\cdot\epsilon_1(gjq)(gj)^{-1}\otimes_{\mathbb{C}H} u = |K|\cdot\sum_{j\in J}\sum_{g\in\overline{H}}\epsilon_1(gjq)(gj)^{-1}\otimes_{\mathbb{C}H} u \\
&= |K|\cdot \underbrace{\sum_{j\in J}\sum_{g\in\overline{H}j}}_{\substack{=\sum_{g\in G}\\ (\text{since } \bigsqcup_{j\in J}\overline{H}j=G)}} \epsilon_1(gq)g^{-1}\otimes_{\mathbb{C}H} u \\
&\qquad \left(\begin{array}{c}\text{here, we substituted } g \text{ for } gj \text{ in the second sum, since the map}\\ \overline{H}\to\overline{H}j,\ g\mapsto gj \text{ is a bijection (because } G \text{ is a group)}\end{array}\right)\\
&= |K|\cdot\sum_{g\in G}\epsilon_1(gq)g^{-1}\otimes_{\mathbb{C}H} u = |K|\cdot\sum_{g\in G}\underbrace{\epsilon_1\left(g^{-1}q\right)}_{\substack{=\epsilon_g(q)\\ (\text{by (13.111.63)})}}\underbrace{\left(g^{-1}\right)^{-1}}_{=g}\otimes_{\mathbb{C}H} u \\
&\qquad \left(\begin{array}{c}\text{here, we substituted } g^{-1} \text{ for } g \text{ in the sum, since the map}\\ G\to G,\ g\mapsto g^{-1} \text{ is a bijection (since } G \text{ is a group)}\end{array}\right)\\
&= |K|\cdot\sum_{g\in G}\epsilon_g(q)g\otimes_{\mathbb{C}H} u = |K|\cdot\underbrace{\left(\sum_{g\in G}\epsilon_g(q)g\right)}_{\substack{=q\\ (\text{by (13.111.60)})}}\otimes_{\mathbb{C}H} u \\
&= |K|\cdot\underbrace{q\otimes_{\mathbb{C}H} u}_{=\mathrm{id}(q\otimes_{\mathbb{C}H} u)} = |K|\cdot\mathrm{id}(q\otimes_{\mathbb{C}H} u) = (|K|\,\mathrm{id})(q\otimes_{\mathbb{C}H} u).
\end{aligned}
$$

Now, let us forget that we fixed q and u. We thus have shown that $(\alpha\circ\beta)(q\otimes_{\mathbb{C}H} u) = (|K|\,\mathrm{id})(q\otimes_{\mathbb{C}H} u)$ for all $q\in\mathbb{C}G$ and $u\in U$. In other words, the two maps $\alpha\circ\beta : \mathbb{C}G\otimes_{\mathbb{C}H} U\to\mathbb{C}G\otimes_{\mathbb{C}H} U$ and $|K|\,\mathrm{id} : \mathbb{C}G\otimes_{\mathbb{C}H} U\to\mathbb{C}G\otimes_{\mathbb{C}H} U$ are equal to each other on each pure tensor. Since these two maps are $\mathbb{C}$-linear, this yields that these two maps $\alpha\circ\beta : \mathbb{C}G\otimes_{\mathbb{C}H} U\to\mathbb{C}G\otimes_{\mathbb{C}H} U$ and $|K|\,\mathrm{id} : \mathbb{C}G\otimes_{\mathbb{C}H} U\to\mathbb{C}G\otimes_{\mathbb{C}H} U$ must be identical. In other words, $\alpha\circ\beta = |K|\,\mathrm{id}$.

Next, we are going to show that $\beta\circ\alpha = |K|\,\mathrm{id}$.

Let $f\in\mathrm{Hom}_{\mathbb{C}H}(\mathbb{C}G,U)$. Let $p\in\mathbb{C}G$. The map f is left $\mathbb{C}H$-linear (since $f\in\mathrm{Hom}_{\mathbb{C}H}(\mathbb{C}G,U)$), hence $\mathbb{C}$-linear. We have $\alpha(f) = \sum_{j\in J} j^{-1}\otimes_{\mathbb{C}H} f(j)$ (by the definition of α). Applying the map β to this equality, we obtain

$$\beta(\alpha(f)) = \beta\left(\sum_{j\in J} j^{-1}\otimes_{\mathbb{C}H} f(j)\right) = \sum_{j\in J}\beta\left(j^{-1}\otimes_{\mathbb{C}H} f(j)\right)$$

(since β is a $\mathbb{C}$-linear map). Thus,

$$\begin{aligned}
&\underbrace{\left(\beta\left(\alpha\left(f\right)\right)\right)}_{=\sum_{j\in J}\beta\left(j^{-1}\otimes_{\mathbb{C}H}f(j)\right)}\left(p\right) \\
&=\left(\sum_{j\in J}\beta\left(j^{-1}\otimes_{\mathbb{C}H}f\left(j\right)\right)\right)\left(p\right)=\sum_{j\in J}\underbrace{\left(\beta\left(j^{-1}\otimes_{\mathbb{C}H}f\left(j\right)\right)\right)\left(p\right)}_{\substack{=\sum\limits_{h\in H}\epsilon_1\left(\rho(h)pj^{-1}\right)h^{-1}f(j)\\ \text{(by (13.111.69), applied to}\\ q=j^{-1}\text{ and }u=f(j))}} \\
&=\sum_{j\in J}\sum_{h\in H}\epsilon_1\left(\rho\left(h\right)pj^{-1}\right)h^{-1}f\left(j\right)=\sum_{j\in J}\sum_{h\in H}\underbrace{\epsilon_1\left(\rho\left(h^{-1}\right)pj^{-1}\right)}_{\substack{=\epsilon_1\left(j^{-1}\rho\left(h^{-1}\right)p\right)\\ \text{(by (13.111.62), applied to}\\ \rho\left(h^{-1}\right)p\text{ and }j^{-1}\text{ instead of }p\text{ and }q)}}\underbrace{\left(h^{-1}\right)^{-1}}_{=h}f\left(j\right) \\
&\qquad\left(\begin{array}{c}\text{here, we substituted }h^{-1}\text{ for }h\text{ in the second sum,}\\ \text{since the map }H\to H,\ h\mapsto h^{-1}\text{ is a bijection}\\ \text{(because }H\text{ is a group)}\end{array}\right) \\
&=\sum_{j\in J}\sum_{h\in H}\epsilon_1\left(j^{-1}\underbrace{\rho\left(h^{-1}\right)}_{\substack{=(\rho(h))^{-1}\\ \text{(since }\rho\text{ is a group}\\ \text{homomorphism)}}}p\right)\underbrace{hf\left(j\right)}_{\substack{=f(hj)\\ \text{(since }f\text{ is left }\mathbb{C}H\text{-linear and since }h\in H\subset\mathbb{C}H)}} \\
&=\sum_{j\in J}\sum_{h\in H}\epsilon_1\left(\underbrace{j^{-1}\left(\rho\left(h\right)\right)^{-1}}_{=(\rho(h)j)^{-1}}p\right)f\left(\underbrace{hj}_{\substack{=\rho(h)\cdot j\\ \text{(by (13.111.55), applied to}\\ \eta=h\text{ and }\gamma=j)}}\right) \\
&=\sum_{j\in J}\underbrace{\sum_{h\in H}}_{\substack{=\sum_{g\in\overline{H}}\sum\limits_{\substack{h\in H;\\ \rho(h)=g}}\\ \text{(since every }h\in H\text{ satisfies}\\ \rho(h)\in\rho(H)=\overline{H})}}\epsilon_1\left(\left(\rho\left(h\right)j\right)^{-1}p\right)f\left(\rho\left(h\right)j\right) \\
&=\sum_{j\in J}\sum_{g\in\overline{H}}\underbrace{\sum_{\substack{h\in H;\\ \rho(h)=g}}}_{=\sum_{h\in\rho^{-1}(g)}}\epsilon_1\left(\left(\underbrace{\rho\left(h\right)}_{=g}j\right)^{-1}p\right)f\left(\underbrace{\rho\left(h\right)}_{=g}j\right)=\sum_{j\in J}\sum_{g\in\overline{H}}\underbrace{\sum_{h\in\rho^{-1}(g)}\epsilon_1\left(\left(gj\right)^{-1}p\right)f\left(gj\right)}_{=\left|\rho^{-1}(g)\right|\cdot\epsilon_1\left((gj)^{-1}p\right)f(gj)} \\
&=\sum_{j\in J}\sum_{g\in\overline{H}}\underbrace{\left|\rho^{-1}\left(g\right)\right|}_{\substack{=|K|\\ \text{(by (13.111.70))}}}\cdot\epsilon_1\left(\left(gj\right)^{-1}p\right)f\left(gj\right)=\sum_{j\in J}\sum_{g\in\overline{H}}\left|K\right|\cdot\epsilon_1\left(\left(gj\right)^{-1}p\right)f\left(gj\right)
\end{aligned}$$

$$= |K| \cdot \sum_{j\in J} \sum_{g\in \overline{H}} \epsilon_1 \left((gj)^{-1} p \right) f(gj) = |K| \cdot \underbrace{\sum_{j\in J} \sum_{g\in \overline{H}j}}_{\substack{= \sum_{g\in G} \\ (\text{since } \bigsqcup_{j\in J} \overline{H}j = G)}} \epsilon_1 \left(g^{-1} p \right) f(g)$$

$$\left(\begin{array}{c} \text{here, we substituted } g \text{ for } gj \text{ in the second sum, since the map} \\ \overline{H} \to \overline{H}j,\ g \mapsto gj \text{ is a bijection (because } G \text{ is a group)} \end{array} \right)$$

$$= |K| \cdot \sum_{g\in G} \underbrace{\epsilon_1 \left(g^{-1} p \right)}_{\substack{= \epsilon_g(p) \\ (\text{by (13.111.63), applied to } q=p)}} f(g) = |K| \cdot \underbrace{\sum_{g\in G} \epsilon_g(p) f(g)}_{\substack{= f\left(\sum_{g\in G} \epsilon_g(p) g \right) \\ (\text{since } f \text{ is } \mathbb{C}\text{-linear})}} = |K| \cdot f\left(\sum_{g\in G} \epsilon_g(p) g \right).$$

Compared with

$$(|K| f)(p) = |K| \cdot f \left(\underbrace{p}_{\substack{= \sum_{g\in G} \epsilon_g(p) g \\ (\text{by (13.111.60), applied to } q=p)}} \right) = |K| \cdot f \left(\sum_{g\in G} \epsilon_g(p) g \right),$$

this yields $(\beta(\alpha(f)))(p) = (|K| f)(p)$.

Now, forget that we fixed p. We thus have proven that $(\beta(\alpha(f)))(p) = (|K| f)(p)$ for every $p \in \mathbb{C}G$. In other words, $\beta(\alpha(f)) = |K| f$. Hence, $(\beta \circ \alpha)(f) = \beta(\alpha(f)) = |K| f = (|K| \operatorname{id})(f)$.

Now, forget that we fixed f. We thus have shown that $(\beta \circ \alpha)(f) = (|K| \operatorname{id})(f)$ for every $f \in \operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, U)$. In other words, $\beta \circ \alpha = |K| \operatorname{id}$.

We now have $\beta \circ \alpha = |K| \operatorname{id}$ and thus $\left(\dfrac{1}{|K|} \beta \right) \circ \alpha = \dfrac{1}{|K|} \underbrace{(\beta \circ \alpha)}_{= |K| \operatorname{id}} = \dfrac{1}{|K|} |K| \operatorname{id} = \operatorname{id}$.

On the other hand, recall that $\alpha \circ \beta = |K| \operatorname{id}$. Hence, $\alpha \circ \left(\dfrac{1}{|K|} \beta \right) = \dfrac{1}{|K|} \underbrace{(\alpha \circ \beta)}_{= |K| \operatorname{id}} = \dfrac{1}{|K|} |K| \operatorname{id} = \operatorname{id}$.

Combining the equalities $\left(\dfrac{1}{|K|} \beta \right) \circ \alpha = \operatorname{id}$ and $\alpha \circ \left(\dfrac{1}{|K|} \beta \right) = \operatorname{id}$, we conclude that the maps α and $\dfrac{1}{|K|} \beta$ are mutually inverse. Hence, the map α is invertible.

Now, we know that the map α is an invertible left $\mathbb{C}G$-module homomorphism. Hence, α is a left $\mathbb{C}G$-module isomorphism. Therefore, there exists a left $\mathbb{C}G$-module isomorphism $\operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, U) \to \mathbb{C}G \otimes_{\mathbb{C}H} U$ (namely, α). Therefore, $\operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, U) \cong \mathbb{C}G \otimes_{\mathbb{C}H} U = \operatorname{Ind}_\rho U$ as left $\mathbb{C}G$-modules. This solves Exercise 4.1.14(i).

[*Remark:* In our above solution of Exercise 4.1.14(i), we explicitly constructed a $\mathbb{C}G$-module isomorphism $\alpha : \operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, U) \to \operatorname{Ind}_\rho U$. This isomorphism is functorial with respect to U. It is also independent on the choice of J (this is not immediately clear from its definition, but it can be shown very easily, by observing that the tensor $j^{-1} \otimes_{\mathbb{C}H} f(j)$ for $j \in G$ depends only on the coset $\overline{H}j$ and not on j itself).]

[*Remark:* Exercise 4.1.14(i) allows us to give yet another solution to Exercise 4.1.14(f):

Second solution to Exercise 4.1.14(f): Assume that $G = H/K$ for some normal subgroup K of H. Let $\rho : H \to G$ be the projection map. We want to prove that $\operatorname{Ind}_\rho U \cong U^K$ for every $\mathbb{C}H$-module U.

We know that ρ is the projection map from H to H/K. Hence, ρ is a surjective group homomorphism and has kernel $\ker \rho = K$.

Let U be a $\mathbb{C}H$-module. Recall that U^K is a $\mathbb{C}[H/K]$-module, thus a $\mathbb{C}G$-module (since $H/K = G$). This $\mathbb{C}G$-module structure has the property that

$$\rho(h) \cdot v = hv \qquad \text{for any } h \in H \text{ and } v \in U^K. \tag{13.111.71}$$

[836]

Let us make $\mathbb{C}G$ into a $(\mathbb{C}H, \mathbb{C}G)$-bimodule as in Exercise 4.1.14(i). From now on, we regard $\mathbb{C}G$ as endowed with this $(\mathbb{C}H, \mathbb{C}G)$-bimodule structure. Then, the $\mathbb{C}G$-module $\operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, U)$ (defined as in Exercise 4.1.4 using the $(\mathbb{C}H, \mathbb{C}G)$-bimodule structure on $\mathbb{C}G$) is isomorphic to $\operatorname{Ind}_\rho U$ (because of Exercise 4.1.14(i)). In other words, $\operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, U) \cong \operatorname{Ind}_\rho U$ as left $\mathbb{C}G$-modules.

We recall that $\operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, U)$ is the left $\mathbb{C}G$-module consisting of all left $\mathbb{C}H$-module homomorphisms from $\mathbb{C}G$ to U. This uses only the $(\mathbb{C}H, \mathbb{C}G)$-bimodule structure on $\mathbb{C}G$ that was introduced in the statement of Exercise 4.1.14(i) (but not the $(\mathbb{C}G, \mathbb{C}H)$-bimodule structure on $\mathbb{C}G$ that was introduced in the definition of $\operatorname{Ind}_\rho U$).

We have $f(1) \in U^K$ for every $f \in \operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, U)$ [837]. Thus, we can define a map $\phi : \operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, U) \to U^K$ by

$$(\phi(f) = f(1) \qquad \text{for all } f \in \operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, U)).$$

[836] *Proof of (13.111.71):* The proof of (13.111.71) is identical with the proof of (13.111.30) in the First solution to Exercise 4.1.14(f).

[837] *Proof.* Let $f \in \operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, U)$. Let $k \in K$. Thus, $k \in K = \ker \rho$, so that $\rho(k) = 1_G$.

But $k \in K \subset H$. Thus, (13.111.55) (applied to $\gamma = 1_{\mathbb{C}G}$ and $\eta = k$) yields $k 1_{\mathbb{C}G} = \underbrace{\rho(k)}_{=1_G} 1_{\mathbb{C}G} = 1_G 1_{\mathbb{C}G} = 1_G$.

But $f \in \operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, U)$. Hence, the map f is left $\mathbb{C}H$-linear. Thus, $f(k 1_{\mathbb{C}G}) = kf(1_{\mathbb{C}G})$ (since $k \in H \subset \mathbb{C}H$). Since $k 1_{\mathbb{C}G} = 1_G$, this rewrites as $f(1_G) = kf\left(\underbrace{1_{\mathbb{C}G}}_{=1}\right) = kf(1)$, so that $kf(1) = f\left(\underbrace{1_G}_{=1}\right) = f(1)$.

Now, let us forget that we fixed k. We thus have shown that $kf(1) = f(1)$ for every $k \in K$. In other words, $f(1)$ is an element y of U which satisfies $ky = y$ for every $k \in K$. In other words,

$$f(1) \in \{y \in U \mid ky = y \text{ for every } k \in K\} = U^K,$$

qed.

Consider this ϕ. This map ϕ is clearly $\mathbb{C}$-linear (since $f(1)$ depends $\mathbb{C}$-linearly on f). Also, the map ϕ is surjective[838] and injective[839]. Hence, the map ϕ is bijective. Thus, ϕ is a $\mathbb{C}$-vector space isomorphism (since ϕ is $\mathbb{C}$-linear).

Moreover, ϕ is a homomorphism of left G-sets[840], and therefore a left $\mathbb{C}G$-module homomorphism (since f is $\mathbb{C}$-linear). Thus, ϕ is a left $\mathbb{C}G$-module isomorphism (since ϕ is a $\mathbb{C}$-vector space isomorphism). We thus have found a left $\mathbb{C}G$-module isomorphism from $\operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, U)$ to U^K (namely, ϕ). Hence, $U^K \cong \operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, U)$ as left $\mathbb{C}G$-modules. Hence, $U^K \cong \operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, U) \cong \operatorname{Ind}_\rho U$ as left $\mathbb{C}G$-modules. This solves Exercise 4.1.14(f) again.]

(j) Let U be a $\mathbb{C}G$-module, and let V be a $\mathbb{C}H$-module.

[838] *Proof.* Let $u \in U^K$. We are going to construct a map $f \in \operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, U)$ which satisfies $\phi(f) = u$.

Indeed, let us define a map $f : \mathbb{C}G \to U$ by

$$(f(\gamma) = \gamma u \qquad \text{for every } \gamma \in \mathbb{C}G). \tag{13.111.72}$$

(This is well-defined, since U^K is a $\mathbb{C}G$-module.) This map f is $\mathbb{C}$-linear (since γu depends $\mathbb{C}$-linearly on γ). We shall now check that the map f is a homomorphism of left H-sets.

Indeed, let $h \in H$ and $\gamma \in \mathbb{C}G$. We have $f(h\gamma) = (h\gamma)u$ (by the definition of f).

On the other hand, (13.111.55) (applied to $\eta = h$) yields $h\gamma = \rho(h) \cdot \gamma$.

But $u \in U^K$ and therefore $\gamma u \in U^K$ (since $\gamma \in \mathbb{C}G$ and since U^K is a $\mathbb{C}G$-module). Thus, (13.111.71) (applied to γu instead of v) yields $\rho(h) \cdot (\gamma u) = h(\gamma u)$. Now,

$$f(h\gamma) = \underbrace{(h\gamma)}_{=\rho(h)\cdot\gamma} u = (\rho(h)\cdot\gamma)u = \rho(h)\cdot(\gamma u) = h \underbrace{(\gamma u)}_{\substack{=f(\gamma) \\ (\text{by } (13.111.72))}} = hf(\gamma).$$

Now, let us forget that we fixed h and γ. We thus have proven that $f(h\gamma) = hf(\gamma)$ for all $h \in H$ and $\gamma \in \mathbb{C}G$. In other words, f is a homomorphism of left H-sets. Hence, f is a left $\mathbb{C}H$-module homomorphism (since f is $\mathbb{C}$-linear). In other words, $f \in \operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, U)$. Now, the definition of ϕ yields

$$\begin{aligned} \phi(f) = f(1) &= 1u \qquad (\text{by } (13.111.72), \text{ applied to } \gamma = 1) \\ &= u. \end{aligned}$$

Thus, $u = \phi\left(\underbrace{f}_{\in \operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G,U)}\right) \in \phi(\operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, U))$.

Let us now forget that we fixed u. We thus have shown that $u \in \phi(\operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, U))$ for every $u \in U^K$. In other words, $U^K \subset \phi(\operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, U))$. This proves that the map ϕ is surjective. Qed.

[839] *Proof.* Let $f \in \ker\phi$. Then, $f \in \operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, U)$ and $\phi(f) = 0$ (since $f \in \ker\phi$). The definition of ϕ yields $\phi(f) = f\left(\underbrace{1}_{=1_{\mathbb{C}G}}\right) = f(1_{\mathbb{C}G})$, so that $f(1_{\mathbb{C}G}) = \phi(f) = 0$.

Now, let $g \in G$ be arbitrary. Then, there exists some $h \in H$ such that $g = \rho(h)$ (since ρ is surjective). Consider this h. Applying (13.111.55) to $\gamma = 1_{\mathbb{C}G}$ and $\eta = h$, we obtain $h1_{\mathbb{C}G} = \rho(h) \cdot 1_{\mathbb{C}G} = \rho(h) = g$. But the map f is left $\mathbb{C}H$-linear (since $f \in \operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, U)$). Thus, $f(h1_{\mathbb{C}G}) = hf(1_{\mathbb{C}G})$ (since $h \in H \subset \mathbb{C}H$), so that $f(h1_{\mathbb{C}G}) = h\underbrace{f(1_{\mathbb{C}G})}_{=0} = 0$. Compared with $f\left(\underbrace{h1_{\mathbb{C}G}}_{=g}\right) = f(g)$, this yields $f(g) = 0$.

Now, let us forget that we fixed g. Thus, we have shown that $f(g) = 0$ for every $g \in G$. In other words, the map f sends every element of the basis G of the $\mathbb{C}$-vector space $\mathbb{C}G$ to 0. Hence, $f = 0$ (since the map f is $\mathbb{C}$-linear).

Now, let us forget that we fixed f. We thus have shown that $f = 0$ for every $f \in \ker\phi$. In other words, $\ker\phi = 0$. Since ϕ is a $\mathbb{C}$-linear map, this shows that ϕ is injective. Qed.

[840] *Proof.* Let $g \in G$ and $f \in \operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, U)$. The definition of the left $\mathbb{C}G$-module structure on $\operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, U)$ yields

$$(gf)(1) = f\left(\underbrace{1g}_{=g}\right) = f(g).$$

Now, the definition of ϕ yields $\phi(gf) = (gf)(1) = f(g)$.

But there exists some $h \in H$ such that $g = \rho(h)$ (since ρ is surjective). Consider this h. The map f is left $\mathbb{C}H$-linear (since $f \in \operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, U)$). Thus, we have $f(h1_{\mathbb{C}G}) = hf(1_{\mathbb{C}G})$. But (13.111.55) (applied to $\gamma = 1_{\mathbb{C}G}$ and $\eta = h$) yields $h1_{\mathbb{C}G} = \underbrace{\rho(h)}_{=g} \cdot 1_{\mathbb{C}G} = g \cdot 1_{\mathbb{C}G} = g$, so that $g = h1_{\mathbb{C}G}$. Hence, $f\left(\underbrace{g}_{=h1_{\mathbb{C}G}}\right) = f(h1_{\mathbb{C}G}) = hf(1_{\mathbb{C}G})$.

Let us make $\mathbb{C}G$ into a $(\mathbb{C}H, \mathbb{C}G)$-bimodule as in Exercise 4.1.14(i). From now on, we regard $\mathbb{C}G$ as endowed with this $(\mathbb{C}H, \mathbb{C}G)$-bimodule structure. Then, the $\mathbb{C}G$-module $\operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, V)$ (defined as in Exercise 4.1.4 using the $(\mathbb{C}H, \mathbb{C}G)$-bimodule structure on $\mathbb{C}G$) is isomorphic to $\operatorname{Ind}_\rho V$ (because of Exercise 4.1.14(i), applied to V instead of U). In other words,

$$\operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, V) \cong \operatorname{Ind}_\rho V \qquad \text{as left } \mathbb{C}G\text{-modules.} \tag{13.111.73}$$

Now, (4.1.8) (applied to $R = \mathbb{C}G$, $S = \mathbb{C}H$, $A = \mathbb{C}G$, $B = U$ and $C = V$) yields

$$\operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G \otimes_{\mathbb{C}G} U, V) \cong \operatorname{Hom}_{\mathbb{C}G}(U, \operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, V)). \tag{13.111.74}$$

We shall now prove that $\mathbb{C}G \otimes_{\mathbb{C}G} U \cong \operatorname{Res}_\rho U$ as left $\mathbb{C}H$-modules.

Indeed, a fundamental fact in abstract algebra says the following: If $\mathfrak{A}$ is a $\mathbb{C}$-algebra, and if $\mathfrak{M}$ is a left $\mathfrak{A}$-module, then there exists a $\mathbb{C}$-vector space isomorphism $\Xi : \mathfrak{A} \otimes_{\mathfrak{A}} \mathfrak{M} \to \mathfrak{M}$ which satisfies $(\Xi(a \otimes_{\mathfrak{A}} m) = am$ for every $a \in \mathfrak{A}$ and $m \in \mathfrak{M})$. Applying this fact to $\mathfrak{A} = \mathbb{C}G$ and $\mathfrak{M} = U$, we conclude that there exists a $\mathbb{C}$-vector space isomorphism $\Xi : \mathbb{C}G \otimes_{\mathbb{C}G} U \to U$ which satisfies
$(\Xi(a \otimes_{\mathbb{C}G} m) = am$ for every $a \in \mathbb{C}G$ and $m \in U)$. Consider this Ξ.

The map Ξ is a $\mathbb{C}$-vector space isomorphism. Hence, its inverse Ξ^{-1} exists and also is a $\mathbb{C}$-vector space isomorphism, and therefore a $\mathbb{C}$-linear map. It furthermore satisfies

$$\Xi^{-1}(u) = 1 \otimes_{\mathbb{C}G} u \qquad \text{for every } u \in U. \tag{13.111.75}$$

[841]

The map Ξ^{-1} is a homomorphism of H-sets from $\operatorname{Res}_\rho U$ to $\mathbb{C}G \otimes_{\mathbb{C}G} U$ [842]. Therefore, Ξ^{-1} is a $\mathbb{C}H$-module homomorphism from $\operatorname{Res}_\rho U$ to $\mathbb{C}G \otimes_{\mathbb{C}G} U$ (since Ξ^{-1} is a $\mathbb{C}$-linear map). Consequently, Ξ^{-1} is a

The definition of ϕ yields $\phi(f) = f\left(\underbrace{1}_{=1_{\mathbb{C}G}}\right) = f(1_{\mathbb{C}G})$, so that $f(1_{\mathbb{C}G}) = \phi(f)$. Now, $f(g) = h\underbrace{f(1_{\mathbb{C}G})}_{=\phi(f)} = h\phi(f)$. Hence,

$$\phi(gf) = f(g) = h\phi(f).$$

The equality (13.111.71) (applied to $\phi(f)$ instead of v) yields $\rho(h) \cdot \phi(f) = h\phi(f)$. Thus, $h\phi(f) = \underbrace{\rho(h)}_{=g} \cdot \phi(f) = g \cdot \phi(f) = g\phi(f)$. Hence, $\phi(gf) = h\phi(f) = g\phi(f)$.

Now, let us forget that we fixed g and f. We thus have proven that $\phi(gf) = g\phi(f)$ for all $g \in G$ and $f \in \operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, U)$. In other words, the map ϕ is a homomorphism of left G-sets. Qed.

[841] *Proof of (13.111.75):* Let $u \in U$. Recall that $\Xi(a \otimes_{\mathbb{C}G} m) = am$ for every $a \in \mathbb{C}G$ and $m \in U$. Applying this to $a = 1$ and $m = u$, we obtain $\Xi(1 \otimes_{\mathbb{C}G} u) = 1u = u$. Hence, $1 \otimes_{\mathbb{C}G} u = \Xi^{-1}(u)$ (since Ξ is invertible). This proves (13.111.75).

[842] *Proof.* The map Ξ^{-1} is a map from U to $\mathbb{C}G \otimes_{\mathbb{C}G} U$, therefore a map from $\operatorname{Res}_\rho U$ to $\mathbb{C}G \otimes_{\mathbb{C}G} U$ (since $\operatorname{Res}_\rho U = U$ as sets).

Recall that the $\mathbb{C}H$-module structure on $\operatorname{Res}_\rho U$ is given by

$$h \cdot v = \rho(h) \cdot v \qquad \text{for every } h \in H \text{ and } v \in U. \tag{13.111.76}$$

Now, let $v \in \operatorname{Res}_\rho U$ and $h \in H$. We are going to prove that $\Xi^{-1}(h \cdot v) = h \cdot \Xi^{-1}(v)$, where $h \cdot \Xi^{-1}(v)$ is computed in the $\mathbb{C}H$-module $\mathbb{C}G \otimes_{\mathbb{C}G} U$.

Indeed, recall that the left $\mathbb{C}H$-module structure on $\mathbb{C}G$ is induced by the $\mathbb{C}$-algebra homomorphism $\mathbb{C}[\rho] : \mathbb{C}H \to \mathbb{C}G$. Thus, it is explicitly given by

$$\eta\gamma = (\mathbb{C}[\rho])\,\eta \cdot \gamma \qquad \text{for all } \gamma \in \mathbb{C}G \text{ and } \eta \in \mathbb{C}H.$$

Applying this to $\eta = h$ and $\gamma = 1_{\mathbb{C}G}$, we obtain $h1_{\mathbb{C}G} = (\mathbb{C}[\rho])(h) \cdot 1_{\mathbb{C}G}$ (since $h \in H \subset \mathbb{C}H$). But $h \in H$ and thus $(\mathbb{C}[\rho])(h) = \rho(h)$. Hence, $h1_{\mathbb{C}G} = \underbrace{(\mathbb{C}[\rho])(h)}_{=\rho(h)} \cdot 1_{\mathbb{C}G} = \rho(h) \cdot 1_{\mathbb{C}G} = \rho(h) = 1_{\mathbb{C}G} \cdot \rho(h)$. Now, $v \in \operatorname{Res}_\rho U = U$. Hence, (13.111.75) (applied to $u = v$) yields $\Xi^{-1}(v) = 1 \otimes_{\mathbb{C}G} v$. Thus,

$$\begin{aligned}
h \cdot \underbrace{\Xi^{-1}(v)}_{=1 \otimes_{\mathbb{C}G} v} &= h \cdot (1 \otimes_{\mathbb{C}G} v) = h \underbrace{1}_{=1_{\mathbb{C}G}} \otimes_{\mathbb{C}G} v = \underbrace{h1_{\mathbb{C}G}}_{=1_{\mathbb{C}G} \cdot \rho(h)} \otimes_{\mathbb{C}G} v = 1_{\mathbb{C}G} \cdot \rho(h) \otimes_{\mathbb{C}G} v \\
&= \underbrace{1_{\mathbb{C}G}}_{=1} \otimes_{\mathbb{C}G} \underbrace{\rho(h) \cdot v}_{\substack{=h \cdot v \\ \text{(by (13.111.76))}}} \qquad \left(\begin{array}{c}\text{here, we have moved } \rho(h) \text{ past the tensor sign} \\ \text{(this is allowed since } \rho(h) \in G \subset \mathbb{C}G)\end{array}\right) \\
&= 1 \otimes_{\mathbb{C}G} h \cdot v.
\end{aligned}$$

Compared with $\Xi^{-1}(h \cdot v) = 1 \otimes_{\mathbb{C}G} h \cdot v$ (by (13.111.75), applied to $u = h \cdot v$), this yields $\Xi^{-1}(h \cdot v) = h \cdot \Xi^{-1}(v)$.

Now, let us forget that we fixed v and h. We thus have shown that $\Xi^{-1}(h \cdot v) = h \cdot \Xi^{-1}(v)$ for every $v \in \operatorname{Res}_\rho U$ and $h \in H$. In other words, Ξ^{-1} is a homomorphism of H-sets from $\operatorname{Res}_\rho U$ to $\mathbb{C}G \otimes_{\mathbb{C}G} U$, qed.

$\mathbb{C}H$-module isomorphism from $\operatorname{Res}_\rho U$ to $\mathbb{C}G \otimes_{\mathbb{C}G} U$ (since Ξ^{-1} is a $\mathbb{C}$-vector space isomorphism). Therefore, $\mathbb{C}G \otimes_{\mathbb{C}G} U \cong \operatorname{Res}_\rho U$ as $\mathbb{C}H$-modules. Now, (4.1.8) (applied to $\mathbb{C}G$, $\mathbb{C}H$, $\mathbb{C}G$, U and V instead of R, S, A, B and C) yields

$$\operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G \otimes_{\mathbb{C}G} U, V) \cong \operatorname{Hom}_{\mathbb{C}G}(U, \operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, V)).$$

Thus,

$$\begin{aligned}\operatorname{Hom}_{\mathbb{C}G}(U, \operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, V)) &\cong \operatorname{Hom}_{\mathbb{C}H}\left(\underbrace{\mathbb{C}G \otimes_{\mathbb{C}G} U}_{\cong \operatorname{Res}_\rho U \text{ as } \mathbb{C}H\text{-modules}}, V\right) \\ &\cong \operatorname{Hom}_{\mathbb{C}H}(\operatorname{Res}_\rho U, V).\end{aligned}$$

Thus,

$$\operatorname{Hom}_{\mathbb{C}H}(\operatorname{Res}_\rho U, V) \cong \operatorname{Hom}_{\mathbb{C}G}\left(U, \underbrace{\operatorname{Hom}_{\mathbb{C}H}(\mathbb{C}G, V)}_{\substack{\cong \operatorname{Ind}_\rho V \text{ as left } \mathbb{C}G\text{-modules} \\ \text{(by (13.111.73))}}}\right) \cong \operatorname{Hom}_{\mathbb{C}G}(U, \operatorname{Ind}_\rho V).$$

This solves Exercise 4.1.14(j).

(k) *Alternative proof of (4.1.3):* Let G be a finite group, and let H be a subgroup of G. Let U be a finite-dimensional $\mathbb{C}H$-module. We need to prove the identity (4.1.3).

Let ρ denote the inclusion map $H \to G$. Clearly, ρ is a group homomorphism. But Exercise 4.1.14(d) yields $\operatorname{Ind}_\rho U = \operatorname{Ind}_H^G U$. Hence, $\chi_{\operatorname{Ind}_\rho U} = \chi_{\operatorname{Ind}_H^G U}$, so that $\chi_{\operatorname{Ind}_H^G U} = \chi_{\operatorname{Ind}_\rho U} = \operatorname{Ind}_\rho \chi_U$ (by Exercise 4.1.14(b)). But Exercise 4.1.14(c) (applied to $f = \chi_U$) yields $\operatorname{Ind}_\rho \chi_U = \operatorname{Ind}_H^G \chi_U$. Thus, $\chi_{\operatorname{Ind}_H^G U} = \operatorname{Ind}_\rho \chi_U = \operatorname{Ind}_H^G \chi_U$. Thus, every $g \in G$ satisfies

$$\underbrace{\chi_{\operatorname{Ind}_H^G U}}_{= \operatorname{Ind}_H^G \chi_U}(g) = \left(\operatorname{Ind}_H^G \chi_U\right)(g) = \frac{1}{|H|} \sum_{\substack{k \in G: \\ kgk^{-1} \in H}} \chi_U\left(kgk^{-1}\right) \qquad \left(\text{by the definition of } \operatorname{Ind}_H^G \chi_U\right).$$

This proves (4.1.3). Thus, Exercise 4.1.14(k) is solved.

(l) *Alternative proof of (4.1.12):* Let G be a finite group, and let K be a normal subgroup of G. Let V be a finite-dimensional $\mathbb{C}G$-module. We need to prove the identity (4.1.12).

Let ρ denote the projection map $G \to G/K$. Exercise 4.1.14(f) (applied to G/K, G and V instead of G, H and U) yields $\operatorname{Ind}_\rho V \cong V^K$ as $\mathbb{C}[G/K]$-modules. Hence, $\chi_{\operatorname{Ind}_\rho V} = \chi_{V^K}$ (since isomorphic $\mathbb{C}[G/K]$-modules have equal characters). Thus, $\chi_{V^K} = \chi_{\operatorname{Ind}_\rho V} = \operatorname{Ind}_\rho \chi_V$ (by Exercise 4.1.14(b), applied to G/K, G and V instead of G, H and U). But Exercise 4.1.14(e) (applied to G/K, G, V and χ_V instead of G, H, U and f) yields $\operatorname{Ind}_\rho \chi_V = (\chi_V)^K$. Hence, $\chi_{V^K} = \operatorname{Ind}_\rho \chi_V = (\chi_V)^K$. Thus, every $g \in G$ satisfies

$$\underbrace{\chi_{V^K}}_{=(\chi_V)^K}(gK) = (\chi_V)^K(gK) = \frac{1}{|K|} \sum_{k \in K} \chi_V(gk)$$

(by the definition of $(\chi_V)^K$). This proves (4.1.12). Thus, Exercise 4.1.14(l) is solved.

[*Remark:* Most parts of Exercise 4.1.14 work in a far greater generality than they are stated in. Let us briefly survey the straightforward generalizations. We begin with the generalizations of the notions of induction, restriction, inflation and fixed point constructions defined in Chapter 4.1:

- Recall that if G is a finite group and H is a subgroup of G, then we have defined a $\mathbb{C}G$-module $\operatorname{Ind}_H^G U$ for every $\mathbb{C}H$-module U, and we have also defined a $\mathbb{C}H$-module $\operatorname{Res}_H^G V$ for every $\mathbb{C}G$-module V. Both of these definitions are also valid when $\mathbb{C}$ is replaced by any commutative ring. Basic properties of induction and restriction (such as the equality (4.1.7), the statement of Exercise 4.1.2 and the statement of Exercise 4.1.3) are still true in this generality (and the proofs that we gave can be transferred to this generality with almost no changes). However, more advanced properties might fail in this generality, and some can not even be stated over a general commutative ring $\mathbb{C}$ [843].

[843] For instance, any properties of characters of modules can only be stated when these characters are well-defined. Since characters are defined as traces of certain endomorphisms of $\mathbb{C}$-vector spaces, they are not automatically well-defined when $\mathbb{C}$

- Recall that if G is a finite group and K is a normal subgroup of G, then we have defined a $\mathbb{C}G$-module $\operatorname{Infl}_{G/K}^{G} U$ for every $\mathbb{C}\left[G/K\right]$-module U, and we have also defined a $\mathbb{C}\left[G/K\right]$-module V^{K} for every $\mathbb{C}G$-module V. Both of these definitions are also valid when $\mathbb{C}$ is replaced by any commutative ring. Some basic properties of inflation and fixed points (such as the isomorphisms (4.1.10) and (4.1.11)) are still correct in this generality (again, they can be proven in the same way as above), but some others are not (e.g., Exercise 4.1.12(b) is generally false[844]).
- If G is a finite group, then we defined $R_{\mathbb{C}}\left(G\right)$ as the $\mathbb{C}$-vector space of all class functions $G \to \mathbb{C}$. This definition still applies when $\mathbb{C}$ is replaced by any commutative ring. (Of course, in this case, "$\mathbb{C}$-vector space" will have to be replaced by "$\mathbb{C}$-module".)
- Recall that if G is a finite group and H is a subgroup of G, then we have defined a class function $\operatorname{Ind}_{H}^{G} f \in R_{\mathbb{C}}\left(G\right)$ for every $f \in R_{\mathbb{C}}\left(H\right)$, and we have also defined a class function $\operatorname{Res}_{H}^{G} f \in R_{\mathbb{C}}\left(H\right)$ for every $f \in R_{\mathbb{C}}\left(G\right)$. The definition of $\operatorname{Res}_{H}^{G} f$ still works when $\mathbb{C}$ is replaced by any commutative ring. Our definition of $\operatorname{Ind}_{H}^{G} f$ cannot be reasonably interpreted in this generality (due to the denominator $\left|H\right|$ in (4.1.4)); however, Exercise 4.1.1(b) can be used as an alternative definition of $\operatorname{Ind}_{H}^{G} f$ when $\mathbb{C}$ is replaced by any commutative ring. (Of course, the necessity of using Exercise 4.1.1(b) as a definition of $\operatorname{Ind}_{H}^{G} f$ entails that most of our proofs concerning $\operatorname{Ind}_{H}^{G} f$ no longer are valid in this generality, because they use the equality (4.1.4) as the definition of $\operatorname{Ind}_{H}^{G} f$. Even the fact that $\operatorname{Ind}_{H}^{G} f$ is well-defined needs to be proven anew, since it is not immediately obvious that the sum $\sum\limits_{\substack{j \in J;\\ jgj^{-1} \in H}} f\left(jgj^{-1}\right)$ in Exercise 4.1.1(b) is independent on J.)

Now, let us move on to generalizing Remark 4.1.13 and Exercise 4.1.14. Let G and H be two finite groups, and let $\rho : H \to G$ be a group homomorphism.

- Remark 4.1.13 still holds if $\mathbb{C}$ is replaced by any commutative ring.
- In Exercise 4.1.14, we defined a map $\operatorname{Ind}_{\rho} f : G \to \mathbb{C}$ for every $f \in R_{\mathbb{C}}\left(H\right)$. This definition still works when $\mathbb{C}$ is replaced by any commutative ring in which $\left|H\right|$ is invertible. There is an alternative definition which works in even greater generality: Namely, as long as $\left|\ker\rho\right|$ is invertible[845], we can define $\operatorname{Ind}_{\rho} f : G \to \mathbb{C}$ by the equality

$$\left(\operatorname{Ind}_{\rho} f\right)\left(g\right) = \frac{1}{\left|\ker\rho\right|} \sum_{\substack{(h,k) \in H \times J;\\ k\rho(h)k^{-1}=g}} f\left(h\right), \tag{13.111.77}$$

where J is a system of left coset representatives for $G/\rho\left(H\right)$ (so that $G = \bigsqcup_{j \in J} j\rho\left(H\right)$). We leave it to the reader to show that this definition is still well-defined (i.e., that $\sum\limits_{\substack{(h,k) \in H \times J;\\ k\rho(h)k^{-1}=g}} f\left(h\right)$ is independent of the choice of J), and that it defines the same function $\operatorname{Ind}_{\rho} f$ as the definition made in Exercise 4.1.14 when $\left|H\right|$ is invertible in the base ring. To my knowledge, there exists no reasonable definition of $\operatorname{Ind}_{\rho} f$ that completely avoids making any requirements on the base ring.
- In Exercise 4.1.14, we defined a $\mathbb{C}G$-module $\operatorname{Ind}_{\rho} U$ for every $\mathbb{C}H$-module U. This definition is still valid when $\mathbb{C}$ is replaced by a commutative ring. (Notice the slightly surprising fact that if $\mathbb{C}$ is replaced by a commutative ring, then the ρ-induction of a $\mathbb{C}H$-module is always well-defined, whereas the ρ-induction of a class function might not be.)
- Exercise 4.1.14(a) is still valid when $\mathbb{C}$ is replaced by any commutative ring in which $\left|H\right|$ is invertible. And the solution that we gave still applies in this generality. More generally, if we define $\operatorname{Ind}_{\rho} f$ by (13.111.77), then Exercise 4.1.14(a) is still valid when $\mathbb{C}$ is replaced by any commutative ring in which $\left|\ker\rho\right|$ is invertible. However, this can no longer be proven by blindly copying our solution.

is replaced by a commutative ring. They are, however, well-defined if our modules are projective modules over this base ring; see the Remark after the proof of Proposition 13.111.3 for details.

[844] However, Exercise 4.1.12(b) holds if we additionally assume that $\left|K\right|$ is invertible in the ring that replaces $\mathbb{C}$. Again, this can be proven by repeating our solution of Exercise 4.1.12(b).

[845] This is a weaker condition than $\left|H\right|$ being invertible.

- Exercise 4.1.14(b) is still valid when $\mathbb{C}$ is replaced by any field in which $|H|$ is invertible. Again, our solution is still valid in this generality.[846] More generally, Exercise 4.1.14(b) is still valid when $\mathbb{C}$ is replaced by any field in which $|\ker\rho|$ is invertible (as long as $\operatorname{Ind}_\rho f$ is defined through (13.111.77)); but our solution is not sufficient to prove this[847]. Even more generally, Exercise 4.1.14(b) holds whenever $\mathbb{C}$ is replaced by any **commutative ring** A in which $|\ker\rho|$ is invertible, as long as U is assumed to be a finitely generated projective A-module. However, in order to make sense of this statement, one needs to know that $\operatorname{Ind}_\rho U$ is a finitely generated projective A-module as well, and one needs to know how the trace of an endomorphism of a finitely generated projective A-module is defined. We essentially did most of this in our solution above[848].
- Exercise 4.1.14(c) is still valid when $\mathbb{C}$ is replaced by any commutative ring in which $|H|$ is invertible. Again, our solution is still valid in this generality. Again, if we define $\operatorname{Ind}_\rho f$ by (13.111.77) and define $\operatorname{Ind}_H^G f$ by Exercise 4.1.1(b), then Exercise 4.1.14(c) even holds when $\mathbb{C}$ is replaced by any arbitrary commutative ring. (The definition of $\operatorname{Ind}_\rho f$ requires $|\ker\rho|$ to be invertible, but this is automatically satisfied since ρ is injective.)
- Exercise 4.1.14(d) is still valid when $\mathbb{C}$ is replaced by any commutative ring whatsoever. Again, our solution is still valid in this generality.
- Exercise 4.1.14(e) is still valid when $\mathbb{C}$ is replaced by any commutative ring in which $|H|$ is invertible. Again, our solution is still valid in this generality. Again, it is possible to prove Exercise 4.1.14(e) also when $\mathbb{C}$ is replaced by any commutative ring in which $|\ker\rho|$ is invertible (as long as $\operatorname{Ind}_\rho f$ is defined using (13.111.77)). (This is actually very easy to check – arguably even easier than our above solution of Exercise 4.1.14(e).)
- Exercise 4.1.14(f) is still valid when $\mathbb{C}$ is replaced by any commutative ring in which $|K|$ is invertible. Both solutions of Exercise 4.1.14(f) given above still remain valid in this generality.
- The second claim of Exercise 4.1.14(g) (that is, the equality (4.1.17)) is still valid when $\mathbb{C}$ is replaced by any commutative ring in which $|G|$ and $|H|$ are invertible. The first claim, a priori, makes no sense when $\mathbb{C}$ is replaced by an arbitrary commutative ring, because the definition of the Hermitian forms $(\cdot,\cdot)_G$ and $(\cdot,\cdot)_H$ involves complex conjugation (which is only defined on $\mathbb{C}$). However, it turns out that this complex conjugation can be replaced by **any map from the base ring to itself** and Exercise 4.1.14(g) remains valid[849].

[846] We can also replace $\mathbb{C}$ by a commutative ring A rather than by a field; but then we need to replace "finite-dimensional $\mathbb{C}H$-module U" by "AH-module U which is a finitely generated projective A-module" in the statement of the exercise. The "finitely generated projective A-module" condition is needed in order to make sure that χ_U and $\chi_{\operatorname{Ind}_\rho U}$ are well-defined. (See the Remark after the proof of Proposition 13.111.3 for details.)

[847] However, a certain modification of this solution does the trick. Namely, we replace the definition of $\widetilde{F}_g$ by

$$\left(\widetilde{F}_g(\gamma,u)=\frac{1}{|\ker\rho|}\sum_{h\in H}(g\rho(h))^*(\gamma)\,hu\qquad\text{for all }(\gamma,u)\in\mathbb{C}G\times U\right).$$

Then, we should pick a system J' of left coset representatives for $G/\rho(H)$ (so that $G=\bigsqcup_{j\in J'}j\rho(H)$). (Notice that we do not call it J because the letter J already has a different meaning in this solution.) Then, we define a family $(a_i)_{i\in J'\times J}$ of elements of $\mathbb{C}G\otimes_{\mathbb{C}H}U$ by

$$\left(a_{(k,j)}=k\otimes_{\mathbb{C}H}b_j\qquad\text{for all }(k,j)\in J'\times J\right),$$

and a family $(f_i)_{i\in J'\times J}$ of elements of $(\mathbb{C}G\otimes_{\mathbb{C}H}U)^*$ by

$$\left(f_{(k,j)}=g_j\circ F_k\qquad\text{for all }(k,j)\in J'\times J\right).$$

(Notice that these families $(a_i)_{i\in J'\times J}$ and $(f_i)_{i\in J'\times J}$ are subfamilies of the families $(a_i)_{i\in G\times J}$ and $(f_i)_{i\in G\times J}$ from the original solution to Exercise 4.1.14(b).) It can then be shown that $\left(J'\times J,(a_i)_{i\in J'\times J},(f_i)_{i\in J'\times J}\right)$ is a finite dual generating system for $\operatorname{Ind}_\rho U$ (though the proof is somewhat different from the way this was shown in the solution to Exercise 4.1.14(b)). This allows us to use Proposition 13.111.3 to derive $\chi_{\operatorname{Ind}_\rho U}=\operatorname{Ind}_\rho\chi_U$. All details are left to the reader.

[848] The only missing link is the fact that an A-module is finitely generated and projective if and only if it has a finite dual generating system; this fact is easy to prove.

[849] More precisely: Let A be any commutative ring in which $|G|$ and $|H|$ are invertible. Fix an arbitrary map $\operatorname{conj}:A\to A$ (not necessarily linear). For every $a\in A$, let $\overline{a}$ denote the image $\operatorname{conj}(a)$ of a under this map. For any $f_1\in R_A(G)$ and $f_2\in R_A(G)$, define $(f_1,f_2)_G\in A$ by $(f_1,f_2)_G=\dfrac{1}{|G|}\sum_{g\in G}f_1(g)\overline{f_2(g)}$ (using the notation $\overline{a}$ that we just introduced). Similarly, define $(f_1,f_2)_H\in A$ for any $f_1\in R_A(H)$ and $f_2\in R_A(H)$. Then, Exercise 4.1.14(g) is still valid when $\mathbb{C}$ is replaced by A. (And our solution to this exercise still applies.)

- Exercise 4.1.14(h) is still valid when $\mathbb{C}$ is replaced by any commutative ring whatsoever. Again, our solution is still valid in this generality.
- Exercise 4.1.14(i) is still valid when $\mathbb{C}$ is replaced by any commutative ring in which $|\ker \rho|$ is invertible. Again, our solution is still valid in this generality (because if we set $K = \ker \rho$ as we did in our solution, then $\left| \underbrace{K}_{= \ker \rho} \right| = |\ker \rho|$ is invertible).
- Exercise 4.1.14(j) is still valid when $\mathbb{C}$ is replaced by any commutative ring in which $|\ker \rho|$ is invertible.
- Exercise 4.1.14(k) still gives a proof of the formula (4.1.3) when $\mathbb{C}$ is replaced by any field in which $|H|$ is invertible.
- Exercise 4.1.14(l) still gives a proof of the formula (4.1.12) when $\mathbb{C}$ is replaced by any field in which $|G|$ and $|K|$ are invertible.

We actually tailored our above solutions to apply in a reasonably high generality (although we did not state Exercise 4.1.14 in this generality). Had we contented ourselves with solving no more and no less than what Exercise 4.1.14 demanded of us, we could have obtained much shorter solutions of some parts of it:

- Exercise 4.1.14(b) can be obtained by combining parts (g) and (h) of Exercise 4.1.14 with the following two observations[850]:
  - Every $\mathbb{C}G$-module V satisfies $\chi_{\operatorname{Res}_\rho V} = \operatorname{Res}_\rho \chi_V$. (This is obvious.)
  - If α and β are two elements of $R_{\mathbb{C}}(G)$ such that every finite-dimensional $\mathbb{C}G$-module V satisfies $(\alpha, \chi_V)_G = (\beta, \chi_V)_G$, then $\alpha = \beta$. (This follows from the fact that $R(G)$ spans the $\mathbb{C}$-vector space $R_{\mathbb{C}}(G)$ and from the fact that the Hermitian form $(\cdot, \cdot)_G$ is nondegenerate.)

  However, this quick argument only works when the base ring is $\mathbb{C}$. It cannot be easily generalized to the case when $\mathbb{C}$ is replaced by any field in which $|H|$ is invertible.
- Part (e) (at least in the case when U is finite-dimensional) and part (f) of Exercise 4.1.14 can be derived from each other using the correspondence between irreducible representations and irreducible characters. Again, this argument is quick but hard to generalize.

]

13.112. **Solution to Exercise 4.1.15.** *Solution to Exercise 4.1.15.* We shall use the following fact:

Proposition 13.112.1. *Let* $\mathbf{k}$ *be a commutative ring. Let* A, B, C, A', B' *and* C' *be six* $\mathbf{k}$*-algebras. Let* P *be an* (A, B)*-bimodule*[851]. *Let* Q *be a* (B, C)*-bimodule. Let* P' *be an* (A', B')*-bimodule. Let* Q' *be a* (B', C')*-bimodule. Then,*

$$(P \otimes P') \otimes_{B \otimes B'} (Q \otimes Q') \cong (P \otimes_B Q) \otimes (P' \otimes_{B'} Q')$$

as $(A \otimes A', C \otimes C')$*-bimodules. Here, all* $\otimes$ *signs without subscript stand for* $\otimes_{\mathbf{k}}$.

Proposition 13.112.1 is identical with the Proposition 13.104.1 that appeared in our solution to Exercise 4.1.3. It is proven by straightforward (repeated) use of the universal property of the tensor product.

We have a canonical $\mathbb{C}$-algebra isomorphism

$$\mathfrak{H} : \mathbb{C}[H_1 \times H_2] \to \mathbb{C}H_1 \otimes \mathbb{C}H_2$$

which satisfies

$$\mathfrak{H}\left(t_{(h_1,h_2)}\right) = t_{h_1} \otimes t_{h_2} \qquad \text{for every } (h_1, h_2) \in H_1 \times H_2.$$

We also have a canonical $\mathbb{C}$-algebra isomorphism

$$\mathfrak{G} : \mathbb{C}[G_1 \times G_2] \to \mathbb{C}G_1 \otimes \mathbb{C}G_2$$

which satisfies

$$\mathfrak{G}\left(t_{(g_1,g_2)}\right) = t_{g_1} \otimes t_{g_2} \qquad \text{for every } (g_1, g_2) \in G_1 \times G_2.$$

[850] We only sketch this argument; the details are left to the reader.

[851] As usual, we understand the notion of a bimodule to be defined over $\mathbf{k}$; that is, the left A-module structure and the right B-module structure of an (A, B)-bimodule must restrict to one and the same $\mathbf{k}$-module structure.

Recall how the left $\mathbb{C}[H_1 \times H_2]$-module $U_1 \otimes U_2$ is defined: Its left $\mathbb{C}[H_1 \times H_2]$-module structure is given by

$$\left(t_{(h_1,h_2)}(u_1 \otimes u_2) = t_{h_1}u_1 \otimes t_{h_2}u_2 \qquad \text{for all } (h_1,h_2) \in H_1 \times H_2 \text{ and } (u_1,u_2) \in U_1 \times U_2\right).$$

We shall now recall the definitions of $\operatorname{Ind}_{\rho_1} U_1$, $\operatorname{Ind}_{\rho_2} U_2$ and $\operatorname{Ind}_{\rho_1 \times \rho_2}(U_1 \otimes U_2)$:

- We defined $\operatorname{Ind}_{\rho_1}(U_1)$ as the $\mathbb{C}G_1$-module $\mathbb{C}G_1 \otimes_{\mathbb{C}H_1} U_1$, where $\mathbb{C}G_1$ is regarded as a $(\mathbb{C}G_1, \mathbb{C}H_1)$-bimodule according to the following rule: The left $\mathbb{C}G_1$-module structure on $\mathbb{C}G_1$ is plain multiplication inside $\mathbb{C}G_1$; the right $\mathbb{C}H_1$-module structure on $\mathbb{C}G_1$ is induced by the $\mathbb{C}$-algebra homomorphism $\mathbb{C}[\rho_1] : \mathbb{C}H_1 \to \mathbb{C}G_1$ (thus, it is explicitly given by $\gamma\eta = \gamma \cdot (\mathbb{C}[\rho_1])\eta$ for all $\gamma \in \mathbb{C}G_1$ and $\eta \in \mathbb{C}H_1$).
- We defined $\operatorname{Ind}_{\rho_2}(U_2)$ as the $\mathbb{C}G_2$-module $\mathbb{C}G_2 \otimes_{\mathbb{C}H_2} U_2$, where $\mathbb{C}G_2$ is regarded as a $(\mathbb{C}G_2, \mathbb{C}H_2)$-bimodule in a similar fashion.
- We defined $\operatorname{Ind}_{\rho_1 \times \rho_2}(U_1 \otimes U_2)$ as the $\mathbb{C}[G_1 \times G_2]$-module $\mathbb{C}[G_1 \times G_2] \otimes_{\mathbb{C}[H_1 \times H_2]} (U_1 \otimes U_2)$, where $\mathbb{C}[G_1 \times G_2]$ is regarded as a $(\mathbb{C}[G_1 \times G_2], \mathbb{C}[H_1 \times H_2])$-bimodule in a similar fashion.

Proposition 13.104.1 (applied to $\mathbf{k} = \mathbb{C}$, $A = \mathbb{C}G_1$, $B = \mathbb{C}H_1$, $C = \mathbb{C}$, $A' = \mathbb{C}G_2$, $B' = \mathbb{C}H_2$, $C' = \mathbb{C}$, $P = \mathbb{C}G_1$, $Q = U_1$, $P' = \mathbb{C}G_2$ and $Q' = U_2$) yields

$$(\mathbb{C}G_1 \otimes \mathbb{C}G_2) \otimes_{\mathbb{C}H_1 \otimes \mathbb{C}H_2} (U_1 \otimes U_2) \cong (\mathbb{C}G_1 \otimes_{\mathbb{C}H_1} U_1) \otimes (\mathbb{C}G_2 \otimes_{\mathbb{C}H_2} U_2)$$

as $(\mathbb{C}G_1 \otimes \mathbb{C}G_2, \mathbb{C} \otimes \mathbb{C})$-bimodules, hence also as left $\mathbb{C}G_1 \otimes \mathbb{C}G_2$-modules. Thus,

$$\begin{aligned}
&(\mathbb{C}G_1 \otimes \mathbb{C}G_2) \otimes_{\mathbb{C}H_1 \otimes \mathbb{C}H_2} (U_1 \otimes U_2) \\
&\cong \underbrace{(\mathbb{C}G_1 \otimes_{\mathbb{C}H_1} U_1)}_{=\operatorname{Ind}_{\rho_1} U_1} \otimes \underbrace{(\mathbb{C}G_2 \otimes_{\mathbb{C}H_2} U_2)}_{=\operatorname{Ind}_{\rho_2} U_2} \\
&= (\operatorname{Ind}_{\rho_1} U_1) \otimes (\operatorname{Ind}_{\rho_2} U_2)
\end{aligned} \tag{13.112.1}$$

as left $\mathbb{C}G_1 \otimes \mathbb{C}G_2$-modules.

At this point, there is a quick way to finish the solution using handwaving: We use the $\mathbb{C}$-algebra isomorphism $\mathfrak{H} : \mathbb{C}[H_1 \times H_2] \to \mathbb{C}H_1 \otimes \mathbb{C}H_2$ to identify the $\mathbb{C}$-algebra $\mathbb{C}[H_1 \times H_2]$ with the $\mathbb{C}$-algebra $\mathbb{C}H_1 \otimes \mathbb{C}H_2$, and we use the $\mathbb{C}$-algebra isomorphism $\mathfrak{G} : \mathbb{C}[G_1 \times G_2] \to \mathbb{C}G_1 \otimes \mathbb{C}G_2$ to identify the $\mathbb{C}$-algebra $\mathbb{C}[G_1 \times G_2]$ with the $\mathbb{C}$-algebra $\mathbb{C}G_1 \otimes \mathbb{C}G_2$. It is "easy to see" that these two identifications "play nicely with each other and with the module structures on $U_1 \otimes U_2$" (this is the part where we wave our hands). Now, (13.112.1) yields that $(\mathbb{C}G_1 \otimes \mathbb{C}G_2) \otimes_{\mathbb{C}H_1 \otimes \mathbb{C}H_2} (U_1 \otimes U_2) \cong (\operatorname{Ind}_{\rho_1} U_1) \otimes (\operatorname{Ind}_{\rho_2} U_2)$ as left $\mathbb{C}G_1 \otimes \mathbb{C}G_2$-modules, and therefore also as left $\mathbb{C}[G_1 \times G_2]$-modules (since $\mathbb{C}[G_1 \times G_2] = \mathbb{C}G_1 \otimes \mathbb{C}G_2$). Now,

$$\begin{aligned}
\operatorname{Ind}_{\rho_1 \times \rho_2}(U_1 \otimes U_2) &= \mathbb{C}[G_1 \times G_2] \otimes_{\mathbb{C}[H_1 \times H_2]} (U_1 \otimes U_2) \\
&= (\mathbb{C}G_1 \otimes \mathbb{C}G_2) \otimes_{\mathbb{C}H_1 \otimes \mathbb{C}H_2} (U_1 \otimes U_2) \\
&\qquad (\text{since } \mathbb{C}[G_1 \times G_2] = \mathbb{C}G_1 \otimes \mathbb{C}G_2 \text{ and } \mathbb{C}[H_1 \times H_2] = \mathbb{C}H_1 \otimes \mathbb{C}H_2) \\
&\cong (\operatorname{Ind}_{\rho_1} U_1) \otimes (\operatorname{Ind}_{\rho_2} U_2)
\end{aligned}$$

as left $\mathbb{C}[G_1 \times G_2]$-modules. This solves Exercise 4.1.15 if you believe the handwaving I have done above.

The handwaving we have done is slightly questionable, since we have made two identifications which turned two isomorphisms into identities. In reality, they are merely isomorphisms, not identities, and it is not immediately clear that regarding them as identities will not lead to contradictions. (Indeed, this is the meaning of our vague claim that the two identifications "play nicely with each other".)

Let us now show a way to formalize the above questionable argument. We shall focus on explaining how to do this proof in a clean fashion (in particular, we shall avoid identifying any things that are not already identical; instead, we will work with the isomorphisms $\mathfrak{G}$ and $\mathfrak{H}$ explicitly); we will leave straightforward computations and arguments to the reader.

We first state a general fact:

Proposition 13.112.2. *Let $\mathbf{k}$ be a commutative ring. In the following, all $\otimes$ signs without subscript stand for $\otimes_{\mathbf{k}}$.*

Let A, B, C, A', B' and C' be six $\mathbf{k}$-algebras. Let M be an (A,B)-bimodule[852]*. Let N be a (B,C)-bimodule. Let M' be an (A',B')-bimodule. Let N' be a (B',C')-bimodule. Let $\beta : B \to B'$, $\mu : M \to M'$*

[852] As usual, we understand the notion of a bimodule to be defined over $\mathbf{k}$; that is, the left A-module structure and the right B-module structure of an (A,B)-bimodule must restrict to one and the same $\mathbf{k}$-module structure.

and $\nu : N \to N'$ be three **k***-module homomorphisms. Assume that we have*

$$(\mu(mb) = \mu(m)\beta(b) \qquad \text{for all } b \in B \text{ and } m \in M) \tag{13.112.2}$$

and

$$(\nu(bn) = \beta(b)\nu(n) \qquad \text{for all } b \in B \text{ and } n \in N). \tag{13.112.3}$$

Then:

(a) *There exists a unique* **k***-module homomorphism*

$$\Omega : M \otimes_B N \to M' \otimes_{B'} N'$$

which satisfies

$$(\Omega(m \otimes_B n) = \mu(m) \otimes_{B'} \nu(n) \qquad \text{for all } (m,n) \in M \times N).$$

We shall denote this homomorphism Ω by $\Omega_{\beta,\mu,\nu}$.

(b) *If the maps β, μ and ν are invertible, then $\Omega_{\beta,\mu,\nu}$ is a* **k***-module isomorphism.*

(c) *Let $\alpha : A \to A'$ be a* **k***-module homomorphism. Assume that*

$$\mu(am) = \alpha(a)\mu(m) \qquad \text{for all } a \in A \text{ and } m \in M. \tag{13.112.4}$$

Assume also that the **k***-module M' is endowed with a left A-module structure. Assume that this left A-module structure on M' and the right B'-module structure on M' together form an (A,B')-bimodule structure on M'. Thus, $M' \otimes_{B'} N'$ becomes a left A-module. Assume furthermore that*

$$am = \alpha(a)m \qquad \text{for every } a \in A \text{ and } m \in M' \tag{13.112.5}$$

[853]*. Then, $\Omega_{\beta,\mu,\nu}$ is a left A-module homomorphism.*

The proof of Proposition 13.112.2 is straightforward[854] and is left to the reader. We could also add a part (d) to Proposition 13.112.2, which would give a criterion for $\Omega_{\beta,\mu,\nu}$ to be a right C-module homomorphism given an appropriate right C-module structure on N'.

Let us now return to solving Exercise 4.1.15. We have four bimodules:

- The $\mathbb{C}$-vector space $\mathbb{C}G_1 \otimes \mathbb{C}G_2$ is a $(\mathbb{C}G_1 \otimes \mathbb{C}G_2, \mathbb{C}H_1 \otimes \mathbb{C}H_2)$-bimodule (because it is the tensor product of the $(\mathbb{C}G_1, \mathbb{C}H_1)$-bimodule $\mathbb{C}G_1$ with the $(\mathbb{C}G_2, \mathbb{C}H_2)$-bimodule $\mathbb{C}G_2$). As a left $\mathbb{C}G_1 \otimes \mathbb{C}G_2$-module, it is thus the tensor product of the left $\mathbb{C}G_1$-module $\mathbb{C}G_1$ with the left $\mathbb{C}G_2$-module $\mathbb{C}G_2$. Hence, its left $\mathbb{C}G_1 \otimes \mathbb{C}G_2$-module structure is given by plain multiplication inside $\mathbb{C}G_1 \otimes \mathbb{C}G_2$ [855]. (This is straightforward to check.)
- The $\mathbb{C}$-vector space $U_1 \otimes U_2$ is a left $\mathbb{C}H_1 \otimes \mathbb{C}H_2$-module (because it is the tensor product of the left $\mathbb{C}H_1$-module U_1 with the left $\mathbb{C}H_2$-module U_2), and thus a $(\mathbb{C}H_1 \otimes \mathbb{C}H_2, \mathbb{C})$-bimodule.
- The $\mathbb{C}$-vector space $\mathbb{C}[G_1 \times G_2]$ is a $(\mathbb{C}[G_1 \times G_2], \mathbb{C}[H_1 \times H_2])$-bimodule (as we already know).
- The $\mathbb{C}$-vector space $U_1 \otimes U_2$ is a left $\mathbb{C}[H_1 \times H_2]$-module (since it is the tensor product of the left $\mathbb{C}H_1$-module U_1 with the left $\mathbb{C}H_2$-module U_2), and thus a $(\mathbb{C}[H_1 \times H_2], \mathbb{C})$-bimodule.

We have

$$(\mathfrak{G}(mb) = \mathfrak{G}(m)\mathfrak{H}(b) \qquad \text{for all } b \in \mathbb{C}[H_1 \times H_2] \text{ and } m \in \mathbb{C}[G_1 \times G_2]) \tag{13.112.6}$$

and

$$(\operatorname{id}(bn) = \mathfrak{H}(b)\operatorname{id}(n) \qquad \text{for all } b \in \mathbb{C}[H_1 \times H_2] \text{ and } n \in U_1 \otimes U_2). \tag{13.112.7}$$

(In fact, both of these equalities are easily checked on basis elements and pure tensors.) Hence, we can apply Proposition 13.112.2(a) to $\mathbf{k} = \mathbb{C}$, $A = \mathbb{C}[G_1 \times G_2]$, $B = \mathbb{C}[H_1 \times H_2]$, $C = \mathbb{C}$, $A' = \mathbb{C}G_1 \otimes \mathbb{C}G_2$, $B' = \mathbb{C}H_1 \otimes \mathbb{C}H_2$, $C' = \mathbb{C}$, $M = \mathbb{C}[G_1 \times G_2]$, $N = U_1 \otimes U_2$, $M' = \mathbb{C}G_1 \otimes \mathbb{C}G_2$, $N' = U_1 \otimes U_2$, $\beta = \mathfrak{H}$, $\mu = \mathfrak{G}$ and $\nu = \operatorname{id}$. As a consequence, we conclude that there exists a unique $\mathbb{C}$-module homomorphism

$$\Omega : \mathbb{C}[G_1 \times G_2] \otimes_{\mathbb{C}[H_1 \times H_2]} (U_1 \otimes U_2) \to (\mathbb{C}G_1 \otimes \mathbb{C}G_2) \otimes_{\mathbb{C}H_1 \otimes \mathbb{C}H_2} (U_1 \otimes U_2)$$

[853] Here the $\alpha(a)m$ on the right hand side is defined using the action of A' on the left A'-module M', whereas the am on the left hand side is defined using the action of A on the left A-module M'.

[854] For part (b), the inverse of $\Omega_{\beta,\mu,\nu}$ is $\Omega_{\beta^{-1},\mu^{-1},\nu^{-1}}$, of course.

[855] That is, the action of any $a \in \mathbb{C}G_1 \otimes \mathbb{C}G_2$ on any element m of the left $\mathbb{C}G_1 \otimes \mathbb{C}G_2$-module $\mathbb{C}G_1 \otimes \mathbb{C}G_2$ equals the product of a with m in the $\mathbb{C}$-algebra $\mathbb{C}G_1 \otimes \mathbb{C}G_2$.

which satisfies

$$\left(\Omega\left(m \otimes_{\mathbb{C}[H_1\times H_2]} n\right) = \mathfrak{G}(m) \otimes_{\mathbb{C}H_1\otimes\mathbb{C}H_2} \mathrm{id}(n) \qquad \text{for all } (m,n) \in \mathbb{C}[G_1\times G_2] \times (U_1\otimes U_2)\right).$$

According to Proposition 13.112.2(a), this homomorphism Ω is denoted by $\Omega_{\mathfrak{H},\mathfrak{G},\mathrm{id}}$.

The maps $\mathfrak{H}$, $\mathfrak{G}$ and id are $\mathbb{C}$-algebra isomorphisms, and therefore invertible. Hence, Proposition 13.112.2(b) (applied to $\mathbf{k} = \mathbb{C}$, $A = \mathbb{C}[G_1\times G_2]$, $B = \mathbb{C}[H_1\times H_2]$, $C = \mathbb{C}$, $A' = \mathbb{C}G_1\otimes\mathbb{C}G_2$, $B' = \mathbb{C}H_1\otimes\mathbb{C}H_2$, $C' = \mathbb{C}$, $M = \mathbb{C}[G_1\times G_2]$, $N = U_1\otimes U_2$, $M' = \mathbb{C}G_1\otimes\mathbb{C}G_2$, $N' = U_1\otimes U_2$, $\beta = \mathfrak{H}$, $\mu = \mathfrak{G}$ and $\nu = \mathrm{id}$) yields that $\Omega_{\mathfrak{H},\mathfrak{G},\mathrm{id}}$ is a $\mathbb{C}$-module isomorphism.

Next, we notice that the $\mathbb{C}$-vector space $\mathbb{C}G_1\otimes\mathbb{C}G_2$ is a left $\mathbb{C}[G_1\times G_2]$-module (since it is the tensor product of the left $\mathbb{C}G_1$-module $\mathbb{C}G_1$ with the left $\mathbb{C}G_2$-module $\mathbb{C}G_2$). This left $\mathbb{C}[G_1\times G_2]$-module structure is defined by the rule

$$(13.112.8)\quad t_{(g_1,g_2)}(u_1\otimes u_2) = t_{g_1}u_1\otimes t_{g_2}u_2 \qquad \text{for all } (g_1,g_2)\in G_1\times G_2 \text{ and } (u_1,u_2)\in\mathbb{C}G_1\times\mathbb{C}G_2.$$

This left $\mathbb{C}[G_1\times G_2]$-module structure on $\mathbb{C}G_1\otimes\mathbb{C}G_2$ and the right $\mathbb{C}H_1\otimes\mathbb{C}H_2$-module structure on $\mathbb{C}G_1\otimes\mathbb{C}G_2$ are connected by the equality

$$(13.112.9)\qquad ((gm)h = g(mh) \qquad \text{for all } g\in\mathbb{C}[G_1\times G_2],\ m\in\mathbb{C}G_1\otimes\mathbb{C}G_2 \text{ and } h\in\mathbb{C}H_1\otimes\mathbb{C}H_2)$$

(which is easily checked). Hence, these two structures together form an $(\mathbb{C}[G_1\times G_2], \mathbb{C}H_1\otimes\mathbb{C}H_2)$-bimodule structure on $\mathbb{C}G_1\otimes\mathbb{C}G_2$ (because both of these structures are $\mathbb{C}$-bilinear). Thus, the tensor product $(\mathbb{C}G_1\otimes\mathbb{C}G_2)\otimes_{\mathbb{C}H_1\otimes\mathbb{C}H_2}(U_1\otimes U_2)$ becomes a left $\mathbb{C}[G_1\times G_2]$-module.

Next, it is easy to check that

$$(13.112.10)\qquad \mathfrak{G}(am) = \mathfrak{G}(a)\,\mathfrak{G}(m) \qquad \text{for all } a\in\mathbb{C}[G_1\times G_2] \text{ and } m\in\mathbb{C}[G_1\times G_2].$$

It is also easy to see that

$$(13.112.11)\qquad am = \mathfrak{G}(a)\,m \qquad \text{for every } a\in\mathbb{C}[G_1\times G_2] \text{ and } m\in\mathbb{C}G_1\otimes\mathbb{C}G_2.$$

Thus, we can apply Proposition 13.112.2(c) to $\mathbf{k} = \mathbb{C}$, $A = \mathbb{C}[G_1\times G_2]$, $B = \mathbb{C}[H_1\times H_2]$, $C = \mathbb{C}$, $A' = \mathbb{C}G_1\otimes\mathbb{C}G_2$, $B' = \mathbb{C}H_1\otimes\mathbb{C}H_2$, $C' = \mathbb{C}$, $M = \mathbb{C}[G_1\times G_2]$, $N = U_1\otimes U_2$, $M' = \mathbb{C}G_1\otimes\mathbb{C}G_2$, $N' = U_1\otimes U_2$, $\beta = \mathfrak{H}$, $\mu = \mathfrak{G}$, $\nu = \mathrm{id}$ and $\alpha = \mathfrak{G}$. As a result, we obtain that $\Omega_{\mathfrak{H},\mathfrak{G},\mathrm{id}}$ is a left $\mathbb{C}[G_1\times G_2]$-module homomorphism. Hence, $\Omega_{\mathfrak{H},\mathfrak{G},\mathrm{id}}$ is a left $\mathbb{C}[G_1\times G_2]$-module isomorphism (since $\Omega_{\mathfrak{H},\mathfrak{G},\mathrm{id}}$ is invertible). Thus,

$$(13.112.12)\qquad \mathbb{C}[G_1\times G_2]\otimes_{\mathbb{C}[H_1\times H_2]}(U_1\otimes U_2) \cong (\mathbb{C}G_1\otimes\mathbb{C}G_2)\otimes_{\mathbb{C}H_1\otimes\mathbb{C}H_2}(U_1\otimes U_2)$$

as left $\mathbb{C}[G_1\times G_2]$-modules.

Now, let us recall the isomorphism (13.112.1). It is an isomorphism of left $\mathbb{C}G_1\otimes\mathbb{C}G_2$-modules, and thus cannot be immediately combined with (13.112.12). However, it is easy to see that the corresponding isomorphism of left $\mathbb{C}[G_1\times G_2]$-modules holds as well: Namely, we have

$$(13.112.13)\qquad (\mathbb{C}G_1\otimes\mathbb{C}G_2)\otimes_{\mathbb{C}H_1\otimes\mathbb{C}H_2}(U_1\otimes U_2) \cong (\mathrm{Ind}_{\rho_1} U_1)\otimes(\mathrm{Ind}_{\rho_2} U_2)$$

as left $\mathbb{C}[G_1\times G_2]$-modules. Before we prove this, let us make three auxiliary observations which connect the left $\mathbb{C}[G_1\times G_2]$-module structures on the modules appearing in (13.112.13) with the left $\mathbb{C}G_1\otimes\mathbb{C}G_2$-module structures on the same modules:

- We have

$$(13.112.14)\qquad t_{(g_1,g_2)}m = (t_{g_1}\otimes t_{g_2})\,m \qquad \text{for every } (g_1,g_2)\in G_1\times G_2 \text{ and } m\in\mathbb{C}G_1\otimes\mathbb{C}G_2$$

  (where the expression $t_{(g_1,g_2)}m$ on the left hand side is defined using the left $\mathbb{C}[G_1\times G_2]$-module structure on $\mathbb{C}G_1\otimes\mathbb{C}G_2$, whereas the expression $(t_{g_1}\otimes t_{g_2})\,m$ on the right hand side is defined using the left $\mathbb{C}G_1\otimes\mathbb{C}G_2$-module structure on $\mathbb{C}G_1\otimes\mathbb{C}G_2$). This is easy to prove.
- We have

(13.112.15)

$$t_{(g_1,g_2)}n = (t_{g_1}\otimes t_{g_2})\,n \qquad \text{for every } (g_1,g_2)\in G_1\times G_2 \text{ and } n\in(\mathbb{C}G_1\otimes\mathbb{C}G_2)\otimes_{\mathbb{C}H_1\otimes\mathbb{C}H_2}(U_1\otimes U_2)$$

  (where the expression $t_{(g_1,g_2)}n$ on the left hand side is defined using the left $\mathbb{C}[G_1\times G_2]$-module structure on $(\mathbb{C}G_1\otimes\mathbb{C}G_2)\otimes_{\mathbb{C}H_1\otimes\mathbb{C}H_2}(U_1\otimes U_2)$, whereas the expression $(t_{g_1}\otimes t_{g_2})\,n$ on the right hand side is defined using the left $\mathbb{C}G_1\otimes\mathbb{C}G_2$-module structure on $(\mathbb{C}G_1\otimes\mathbb{C}G_2)\otimes_{\mathbb{C}H_1\otimes\mathbb{C}H_2}(U_1\otimes U_2)$). This is easy to prove using (13.112.14).

- Finally, we have

$$t_{(g_1,g_2)}m = (t_{g_1} \otimes t_{g_2}) m \qquad \text{for every } (g_1, g_2) \in G_1 \times G_2 \text{ and } m \in (\operatorname{Ind}_{\rho_1} U_1) \otimes (\operatorname{Ind}_{\rho_2} U_2) \tag{13.112.16}$$

(where the expression $t_{(g_1,g_2)}m$ on the left hand side is defined using the left $\mathbb{C}[G_1 \times G_2]$-module structure on $(\operatorname{Ind}_{\rho_1} U_1) \otimes (\operatorname{Ind}_{\rho_2} U_2)$, whereas the expression $(t_{g_1} \otimes t_{g_2}) m$ on the right hand side is defined using the left $\mathbb{C}G_1 \otimes \mathbb{C}G_2$-module structure on $(\operatorname{Ind}_{\rho_1} U_1) \otimes (\operatorname{Ind}_{\rho_2} U_2)$). Again, this is easy to check.

Now, it is easy to see that (13.112.13) holds[856].

Now, (13.112.12) becomes

$$\begin{aligned}&\mathbb{C}[G_1 \times G_2] \otimes_{\mathbb{C}[H_1 \times H_2]} (U_1 \otimes U_2)\\ &\cong (\mathbb{C}G_1 \otimes \mathbb{C}G_2) \otimes_{\mathbb{C}H_1 \otimes \mathbb{C}H_2} (U_1 \otimes U_2) \cong (\operatorname{Ind}_{\rho_1} U_1) \otimes (\operatorname{Ind}_{\rho_2} U_2) \qquad (\text{by } (13.112.13))\end{aligned}$$

as left $\mathbb{C}[G_1 \times G_2]$-modules. Therefore,

$$\operatorname{Ind}_{\rho_1 \times \rho_2} (U_1 \otimes U_2) = \mathbb{C}[G_1 \times G_2] \otimes_{\mathbb{C}[H_1 \times H_2]} (U_1 \otimes U_2) \cong (\operatorname{Ind}_{\rho_1} U_1) \otimes (\operatorname{Ind}_{\rho_2} U_2)$$

as left $\mathbb{C}[G_1 \times G_2]$-modules. This finishes the solution of Exercise 4.1.15.

13.113. **Solution to Exercise 4.1.16.** *Solution to Exercise 4.1.16.* In the following, we will use the following convention: Whenever K is a group, and k is an element of K, we shall write k for the element t_k of $\mathbb{C}K$. This is a relatively common abuse of notation, and it is harmless because the map $K \to \mathbb{C}K,\ k \mapsto t_k$ is an injective homomorphism of multiplicative monoids (so $t_{gh} = t_g t_h$ and $t_1 = 1$, which means that we won't run into ambiguities denoting t_k by k) and because every $\mathbb{C}K$-module M, every $m \in M$ and every $k \in K$ satisfy $km = t_k m$.

We solve the four parts of Exercise 4.1.16 in the following order: first, part (c); then, part (d); then, part (a); finally, part (b).

(c) The definition of $\operatorname{Res}_\rho V$ yields $\operatorname{Res}_\rho V = V$ as $\mathbb{C}$-vector spaces. Similarly, $\operatorname{Res}_\tau \operatorname{Res}_\rho V = \operatorname{Res}_\rho V$ as $\mathbb{C}$-vector spaces, and $\operatorname{Res}_{\rho\circ\tau} V = V$ as $\mathbb{C}$-vector spaces. Thus, $\operatorname{Res}_\tau \operatorname{Res}_\rho V = \operatorname{Res}_\rho V = V = \operatorname{Res}_{\rho\circ\tau} V$ as $\mathbb{C}$-vector spaces. But our goal is to show that $\operatorname{Res}_\tau \operatorname{Res}_\rho V = \operatorname{Res}_{\rho\circ\tau} V$ as $\mathbb{C}I$-modules. Thus, it suffices to prove that the left $\mathbb{C}I$-module structures on $\operatorname{Res}_\tau \operatorname{Res}_\rho V$ and $\operatorname{Res}_{\rho\circ\tau} V$ are identical. In other words, it

[856] *Proof of (13.112.13):* From (13.112.1), we conclude that there exists an isomorphism

$$\mathbf{T} : (\mathbb{C}G_1 \otimes \mathbb{C}G_2) \otimes_{\mathbb{C}H_1 \otimes \mathbb{C}H_2} (U_1 \otimes U_2) \to (\operatorname{Ind}_{\rho_1} U_1) \otimes (\operatorname{Ind}_{\rho_2} U_2)$$

of left $\mathbb{C}G_1 \otimes \mathbb{C}G_2$-modules. We shall now show that this map $\mathbf{T}$ is an isomorphism of left $\mathbb{C}[G_1 \times G_2]$-modules as well. In order to do so, it is sufficient to show that the map $\mathbf{T}$ is a homomorphism of left $\mathbb{C}[G_1 \times G_2]$-modules (because we already know that $\mathbf{T}$ is invertible). In other words, it is sufficient to show that

$$\mathbf{T}(an) = a\mathbf{T}(n) \tag{13.112.17}$$

for every $a \in \mathbb{C}[G_1 \times G_2]$ and $n \in (\mathbb{C}G_1 \otimes \mathbb{C}G_2) \otimes_{\mathbb{C}H_1 \otimes \mathbb{C}H_2} (U_1 \otimes U_2)$ (since we already know that the map $\mathbf{T}$ is $\mathbb{C}$-linear).

Proof of (13.112.17): Let $a \in \mathbb{C}[G_1 \times G_2]$ and $n \in (\mathbb{C}G_1 \otimes \mathbb{C}G_2) \otimes_{\mathbb{C}H_1 \otimes \mathbb{C}H_2} (U_1 \otimes U_2)$. We need to prove the equality $\mathbf{T}(an) = a\mathbf{T}(n)$. This equality is $\mathbb{C}$-linear in a. Hence, we can WLOG assume that a belongs to the basis $\left(t_{(g_1,g_2)}\right)_{(g_1,g_2)\in G_1\times G_2}$ of the $\mathbb{C}$-vector space $\mathbb{C}[G_1 \times G_2]$. Assume this. Thus, $a = t_{(g_1,g_2)}$ for some $(g_1, g_2) \in G_1 \times G_2$. Consider this (g_1, g_2). We have $\underbrace{a}_{=t_{(g_1,g_2)}} n = t_{(g_1,g_2)}n = (t_{g_1} \otimes t_{g_2}) n$ (by (13.112.15)). Applying the map $\mathbf{T}$ to both sides of this equality, we obtain

$$\mathbf{T}(an) = \mathbf{T}((t_{g_1} \otimes t_{g_2}) n) = (t_{g_1} \otimes t_{g_2}) \mathbf{T}(n) \qquad (\text{since } \mathbf{T} \text{ is a homomorphism of left } \mathbb{C}G_1 \otimes \mathbb{C}G_2\text{-modules}).$$

Compared with

$$\underbrace{a}_{=t_{(g_1,g_2)}} \mathbf{T}(n) = t_{(g_1,g_2)}\mathbf{T}(n) = (t_{g_1} \otimes t_{g_2}) \mathbf{T}(n) \qquad (\text{by } (13.112.16), \text{ applied to } m = \mathbf{T}(n)),$$

this yields $\mathbf{T}(an) = a\mathbf{T}(n)$. This proves (13.112.17).

Now, (13.112.17) shows that $\mathbf{T}$ is a homomorphism of left $\mathbb{C}[G_1 \times G_2]$-modules (since $\mathbf{T}$ is $\mathbb{C}$-linear), and therefore an isomorphism of left $\mathbb{C}[G_1 \times G_2]$-modules (since $\mathbf{T}$ is invertible). Thus, $(\mathbb{C}G_1 \otimes \mathbb{C}G_2) \otimes_{\mathbb{C}H_1 \otimes \mathbb{C}H_2} (U_1 \otimes U_2) \cong (\operatorname{Ind}_{\rho_1} U_1) \otimes (\operatorname{Ind}_{\rho_2} U_2)$ as left $\mathbb{C}[G_1 \times G_2]$-modules. This proves (13.112.13).

suffices to show that every $i \in \mathbb{C}I$ and $v \in V$ satisfy the equality

$$\begin{aligned}&(\text{the action of } i \in \mathbb{C}I \text{ on the element } v \text{ of the left } \mathbb{C}I\text{-module } \operatorname{Res}_\tau \operatorname{Res}_\rho V) \\ &= (\text{the action of } i \in \mathbb{C}I \text{ on the element } v \text{ of the left } \mathbb{C}I\text{-module } \operatorname{Res}_{\rho\circ\tau} V).\end{aligned} \tag{13.113.1}$$

(Of course, both sides of the equality (13.113.1) could be rewritten as iv, but this notation is ambiguous, because i simultaneously belongs to two $\mathbb{C}I$-modules $\operatorname{Res}_\tau \operatorname{Res}_\rho V$ and $\operatorname{Res}_{\rho\circ\tau} V$ which are not yet known to be identical.)

Proof of (13.113.1): Let $i \in \mathbb{C}I$ and $v \in V$. We need to prove the equality (13.113.1). Since this equality is $\mathbb{C}$-linear in i, we can WLOG assume that i belongs to the basis I of the $\mathbb{C}$-vector space $\mathbb{C}I$. Assume this. Now, the definition of the left $\mathbb{C}I$-module $\operatorname{Res}_\tau \operatorname{Res}_\rho V$ yields

$$\begin{aligned}&(\text{the action of } i \in \mathbb{C}I \text{ on the element } v \text{ of the left } \mathbb{C}I\text{-module } \operatorname{Res}_\tau \operatorname{Res}_\rho V) \\ &= (\text{the action of } \tau(i) \in \mathbb{C}H \text{ on the element } v \text{ of the left } \mathbb{C}H\text{-module } \operatorname{Res}_\rho V) \\ &= \left(\text{the action of } \underbrace{\rho(\tau(i))}_{=(\rho\circ\tau)(i)} \in \mathbb{C}G \text{ on the element } v \text{ of the left } \mathbb{C}G\text{-module } V\right) \\ &\qquad (\text{by the definition of the left } \mathbb{C}H\text{-module } \operatorname{Res}_\rho V) \\ &= (\text{the action of } (\rho\circ\tau)(i) \in \mathbb{C}G \text{ on the element } v \text{ of the left } \mathbb{C}G\text{-module } V).\end{aligned}$$

Compared with

$$\begin{aligned}&(\text{the action of } i \in \mathbb{C}I \text{ on the element } v \text{ of the left } \mathbb{C}I\text{-module } \operatorname{Res}_{\rho\circ\tau} V) \\ &= (\text{the action of } (\rho\circ\tau)(i) \in \mathbb{C}G \text{ on the element } v \text{ of the left } \mathbb{C}G\text{-module } V) \\ &\qquad (\text{by the definition of the left } \mathbb{C}I\text{-module } \operatorname{Res}_{\rho\circ\tau} V),\end{aligned}$$

this yields

$$\begin{aligned}&(\text{the action of } i \in \mathbb{C}I \text{ on the element } v \text{ of the left } \mathbb{C}I\text{-module } \operatorname{Res}_\tau \operatorname{Res}_\rho V) \\ &= (\text{the action of } i \in \mathbb{C}I \text{ on the element } v \text{ of the left } \mathbb{C}I\text{-module } \operatorname{Res}_{\rho\circ\tau} V).\end{aligned}$$

This proves (13.113.1).

Now, we know that $\operatorname{Res}_\tau \operatorname{Res}_\rho V = \operatorname{Res}_{\rho\circ\tau} V$ as $\mathbb{C}$-vector spaces. Thus, (13.113.1) shows that $\operatorname{Res}_\tau \operatorname{Res}_\rho V = \operatorname{Res}_{\rho\circ\tau} V$ as left $\mathbb{C}I$-modules as well. Exercise 4.1.16(c) is thus solved.

(d) Let $f \in R_{\mathbb{C}}(G)$. The definition of $\operatorname{Res}_{\rho\circ\tau} f$ yields $\operatorname{Res}_{\rho\circ\tau} f = f \circ (\rho \circ \tau)$. But the definition of $\operatorname{Res}_\rho f$ yields $\operatorname{Res}_\rho f = f \circ \rho$. The definition of $\operatorname{Res}_\tau \operatorname{Res}_\rho f$ yields

$$\operatorname{Res}_\tau \operatorname{Res}_\rho f = \underbrace{(\operatorname{Res}_\rho f)}_{=f\circ\rho} \circ \tau = (f\circ\rho)\circ\tau = f\circ(\rho\circ\tau) = \operatorname{Res}_{\rho\circ\tau} f.$$

This solves Exercise 4.1.16(d).

(a) Let U be any $\mathbb{C}I$-module.

Recall that $\operatorname{Ind}_\tau U$ is defined as the $\mathbb{C}H$-module $\mathbb{C}H \otimes_{\mathbb{C}I} U$, where $\mathbb{C}H$ is regarded as a $(\mathbb{C}H, \mathbb{C}I)$-bimodule according to the following rule: The left $\mathbb{C}H$-module structure on $\mathbb{C}H$ is plain multiplication inside $\mathbb{C}H$; the right $\mathbb{C}I$-module structure on $\mathbb{C}H$ is induced by the $\mathbb{C}$-algebra homomorphism $\mathbb{C}[\tau] : \mathbb{C}I \to \mathbb{C}H$ (thus, it is explicitly given by $\gamma\eta = \gamma \cdot (\mathbb{C}[\tau])\,\eta$ for all $\gamma \in \mathbb{C}H$ and $\eta \in \mathbb{C}I$).

Furthermore, $\operatorname{Ind}_\rho(\operatorname{Ind}_\tau U)$ is defined as the $\mathbb{C}G$-module $\mathbb{C}G \otimes_{\mathbb{C}H} (\operatorname{Ind}_\tau U)$, where $\mathbb{C}G$ is regarded as a $(\mathbb{C}G, \mathbb{C}H)$-bimodule according to the following rule: The left $\mathbb{C}G$-module structure on $\mathbb{C}G$ is plain multiplication inside $\mathbb{C}G$; the right $\mathbb{C}H$-module structure on $\mathbb{C}G$ is induced by the $\mathbb{C}$-algebra homomorphism $\mathbb{C}[\rho] : \mathbb{C}H \to \mathbb{C}G$ (thus, it is explicitly given by $\gamma\eta = \gamma \cdot (\mathbb{C}[\rho])\,\eta$ for all $\gamma \in \mathbb{C}G$ and $\eta \in \mathbb{C}H$).

Finally, $\operatorname{Ind}_{\rho\circ\tau} U$ is defined as the $\mathbb{C}G$-module $\mathbb{C}G \otimes_{\mathbb{C}I} U$, where $\mathbb{C}G$ is regarded as a $(\mathbb{C}G, \mathbb{C}I)$-bimodule according to the following rule: The left $\mathbb{C}G$-module structure on $\mathbb{C}G$ is plain multiplication inside $\mathbb{C}G$; the right $\mathbb{C}I$-module structure on $\mathbb{C}G$ is induced by the $\mathbb{C}$-algebra homomorphism $\mathbb{C}[\rho\circ\tau] : \mathbb{C}I \to \mathbb{C}G$ (thus, it is explicitly given by $\gamma\eta = \gamma \cdot (\mathbb{C}[\rho\circ\tau])\,\eta$ for all $\gamma \in \mathbb{C}G$ and $\eta \in \mathbb{C}I$).

Thus, we have introduced a $(\mathbb{C}H, \mathbb{C}I)$-bimodule structure on $\mathbb{C}H$, a $(\mathbb{C}G, \mathbb{C}H)$-bimodule structure on $\mathbb{C}G$ and a $(\mathbb{C}G, \mathbb{C}I)$-bimodule structure on $\mathbb{C}G$. The left $\mathbb{C}G$-module structure that underlies the $(\mathbb{C}G, \mathbb{C}H)$-bimodule structure on $\mathbb{C}G$ is identical with the left $\mathbb{C}G$-module structure that underlies the $(\mathbb{C}G, \mathbb{C}I)$-bimodule structure on $\mathbb{C}G$ (because both of these left $\mathbb{C}G$-module structures are defined to be plain multiplication inside $\mathbb{C}G$). Therefore, we will not run into ambiguities if we write expressions such as ab for $a \in \mathbb{C}G$ and $b \in \mathbb{C}G$.

We shall now show that

$$\mathbb{C}G \otimes_{\mathbb{C}H} \mathbb{C}H \cong \mathbb{C}G \qquad \text{as } (\mathbb{C}G, \mathbb{C}I)\text{-bimodules.} \tag{13.113.2}$$

(Here, on the left hand side, $\mathbb{C}G$ is regarded as a $(\mathbb{C}G, \mathbb{C}H)$-bimodule and $\mathbb{C}H$ is regarded as $(\mathbb{C}H, \mathbb{C}I)$-bimodule, whereas on the right hand side, $\mathbb{C}G$ is regarded as a $(\mathbb{C}G, \mathbb{C}I)$-bimodule.)

Proof of (13.113.2): There is clearly a unique $\mathbb{C}$-vector space isomorphism $\Phi : \mathbb{C}G \to \mathbb{C}G \otimes_{\mathbb{C}H} \mathbb{C}H$ which satisfies

$$(\Phi(m) = m \otimes_{\mathbb{C}H} 1_{\mathbb{C}H} \qquad \text{for all } m \in \mathbb{C}G). \tag{13.113.3}$$

[857] Consider this Φ. It is straightforward to see that Φ is a homomorphism of left G-sets and a homomorphism of right I-sets. Hence, Φ is a homomorphism of $(\mathbb{C}G, \mathbb{C}I)$-bimodules (since Φ is $\mathbb{C}$-linear), thus an isomorphism of $(\mathbb{C}G, \mathbb{C}I)$-bimodules (since Φ is invertible). Therefore, there exists an isomorphism of $(\mathbb{C}G, \mathbb{C}I)$-bimodules from $\mathbb{C}G$ to $\mathbb{C}G \otimes_{\mathbb{C}H} \mathbb{C}H$ (namely, Φ). In other words, $\mathbb{C}G \otimes_{\mathbb{C}H} \mathbb{C}H \cong \mathbb{C}G$ as $(\mathbb{C}G, \mathbb{C}I)$-bimodules. This proves (13.113.2).

Now,

$$\begin{aligned}
\operatorname{Ind}_\rho (\operatorname{Ind}_\tau U) &= \mathbb{C}G \otimes_{\mathbb{C}H} \underbrace{(\operatorname{Ind}_\tau U)}_{=\mathbb{C}H \otimes_{\mathbb{C}I} U} = \mathbb{C}G \otimes_{\mathbb{C}H} (\mathbb{C}H \otimes_{\mathbb{C}I} U) \\
&\cong \underbrace{(\mathbb{C}G \otimes_{\mathbb{C}H} \mathbb{C}H)}_{\substack{\cong \mathbb{C}G \text{ as } (\mathbb{C}G, \mathbb{C}I)\text{-bimodules} \\ \text{(by (13.113.2))}}} \otimes_{\mathbb{C}I} U \qquad \text{(by the associativity of the tensor product)} \\
&\cong \mathbb{C}G \otimes_{\mathbb{C}I} U = \operatorname{Ind}_{\rho \circ \tau} U \qquad \left(\text{since } \operatorname{Ind}_{\rho \circ \tau} U = \mathbb{C}G \otimes_{\mathbb{C}I} U \text{ as left } \mathbb{C}G\text{-modules}\right)
\end{aligned}$$

as left $\mathbb{C}G$-modules. This solves Exercise 4.1.16(a).

(b) *First solution to Exercise 4.1.16(b).* Let $f \in R_{\mathbb{C}}(I)$. Every $r \in H$ satisfies

$$\begin{aligned}
(\operatorname{Ind}_\tau f)(r) &= \frac{1}{|I|} \sum_{\substack{(h,k) \in I \times H; \\ k\tau(h)k^{-1} = r}} f(h) \qquad \left(\text{by the definition of } \operatorname{Ind}_\tau f\right) \\
&= \frac{1}{|I|} \underbrace{\sum_{\substack{(i,v) \in I \times H; \\ v\tau(i)v^{-1} = r}}}_{= \sum_{i \in I} \sum_{\substack{v \in H; \\ v\tau(i)v^{-1} = r}}} f(i) \qquad \left(\text{here, we renamed the summation index } (h,k) \text{ as } (i,v)\right) \\
&= \frac{1}{|I|} \sum_{i \in I} \sum_{\substack{v \in H; \\ v\tau(i)v^{-1} = r}} f(i).
\end{aligned} \tag{13.113.4}$$

[857] Indeed, this is a particular case of the following fundamental fact from linear algebra: If A is a $\mathbb{C}$-algebra, and if M is a right A-module, then there is a unique $\mathbb{C}$-vector space isomorphism $\Phi : M \to M \otimes_A A$ which satisfies

$$(\Phi(m) = m \otimes_A 1_A \qquad \text{for all } m \in M).$$

Now, let $g \in G$. Then, the definition of $\operatorname{Ind}_\rho\left(\operatorname{Ind}_\tau f\right)$ yields

$$\left(\operatorname{Ind}_\rho\left(\operatorname{Ind}_\tau f\right)\right)(g) = \frac{1}{|H|} \sum_{\substack{(h,k)\in H\times G;\\ k\rho(h)k^{-1}=g}} \left(\operatorname{Ind}_\tau f\right)(h) = \frac{1}{|H|} \underbrace{\sum_{\substack{(r,k)\in H\times G;\\ k\rho(r)k^{-1}=g}}}_{=\sum_{k\in G} \sum_{\substack{r\in H;\\ k\rho(r)k^{-1}=g}}} \underbrace{\left(\operatorname{Ind}_\tau f\right)(r)}_{\substack{=\frac{1}{|I|}\sum_{i\in I}\sum_{\substack{v\in H;\\ v\tau(i)v^{-1}=r}} f(i)\\ \text{(by (13.113.4))}}}$$

$$\text{(here, we renamed the summation index } (h,k) \text{ as } (r,k)\text{)}$$

$$= \frac{1}{|H|} \sum_{k\in G} \sum_{\substack{r\in H;\\ k\rho(r)k^{-1}=g}} \frac{1}{|I|} \sum_{i\in I} \sum_{\substack{v\in H;\\ v\tau(i)v^{-1}=r}} f(i)$$

$$= \frac{1}{|I|}\frac{1}{|H|} \sum_{k\in G} \underbrace{\sum_{\substack{r\in H;\\ k\rho(r)k^{-1}=g}} \sum_{i\in I}}_{=\sum_{i\in I} \sum_{\substack{r\in H;\\ k\rho(r)k^{-1}=g}}} \sum_{\substack{v\in H;\\ v\tau(i)v^{-1}=r}} f(i)$$

$$= \frac{1}{|I|}\frac{1}{|H|} \sum_{k\in G} \sum_{i\in I} \sum_{\substack{r\in H;\\ k\rho(r)k^{-1}=g}} \sum_{\substack{v\in H;\\ v\tau(i)v^{-1}=r}} f(i). \tag{13.113.5}$$

But every $k \in G$ and $i \in I$ satisfy

$$\sum_{\substack{r\in H;\\ k\rho(r)k^{-1}=g}} \sum_{\substack{v\in H;\\ v\tau(i)v^{-1}=r}} f(i) = \sum_{\substack{v\in H;\\ k\rho(v)\cdot\rho(\tau(i))\cdot(k\rho(v))^{-1}=g}} f(i) \tag{13.113.6}$$

[858].

[858] *Proof of (13.113.6):* Let $k \in G$ and $i \in I$. We must prove (13.113.6). We have

$$\underbrace{\sum_{\substack{r\in H;\\ k\rho(r)k^{-1}=g}} \sum_{\substack{v\in H;\\ v\tau(i)v^{-1}=r}}}_{=\sum_{v\in H} \sum_{\substack{r\in H;\\ k\rho(r)k^{-1}=g;\\ v\tau(i)v^{-1}=r}}} f(i)$$

$$= \sum_{v\in H} \sum_{\substack{r\in H;\\ k\rho(r)k^{-1}=g;\\ v\tau(i)v^{-1}=r}} f(i) = \sum_{\substack{v\in H;\\ k\rho(v)\cdot\rho(\tau(i))\cdot(k\rho(v))^{-1}=g}} \sum_{\substack{r\in H;\\ k\rho(r)k^{-1}=g;\\ v\tau(i)v^{-1}=r}} f(i) + \sum_{\substack{v\in H;\\ k\rho(v)\cdot\rho(\tau(i))\cdot(k\rho(v))^{-1}\neq g}} \sum_{\substack{r\in H;\\ k\rho(r)k^{-1}=g;\\ v\tau(i)v^{-1}=r}} f(i) \tag{13.113.7}$$

(because every $v \in H$ satisfies either $k\rho(v)\cdot\rho(\tau(i))\cdot(k\rho(v))^{-1} = g$ or $k\rho(v)\cdot\rho(\tau(i))\cdot(k\rho(v))^{-1} \neq g$, but never both).

Now, we shall show that

$$\sum_{\substack{r\in H;\\ k\rho(r)k^{-1}=g;\\ v\tau(i)v^{-1}=r}} f(i) = 0 \tag{13.113.8}$$

for every $v \in H$ which satisfies $k\rho(v)\cdot\rho(\tau(i))\cdot(k\rho(v))^{-1} \neq g$.

Proof of (13.113.8): Let $v \in H$ be such that $k\rho(v)\cdot\rho(\tau(i))\cdot(k\rho(v))^{-1} \neq g$. If some $r \in H$ satisfies $k\rho(r)k^{-1} = g$ and $v\tau(i)v^{-1} = r$, then this r must satisfy

$$k\rho(v)\cdot\rho(\tau(i))\cdot\underbrace{(k\rho(v))^{-1}}_{=(\rho(v))^{-1}k^{-1}} = k\underbrace{\rho(v)\cdot\rho(\tau(i))\cdot(\rho(v))^{-1}}_{\substack{=\rho\left(v\tau(i)v^{-1}\right)\\ \text{(since } \rho \text{ is a group}\\ \text{homomorphism)}}}k^{-1}$$

$$= k\rho\left(\underbrace{v\tau(i)v^{-1}}_{=r}\right)k^{-1} = k\rho(r)k^{-1} = g,$$

which contradicts $k\rho(v)\cdot\rho(\tau(i))\cdot(k\rho(v))^{-1}\neq g$. Hence, we have obtained a contradiction for every $r\in H$ which satisfies $k\rho(r)k^{-1}=g$ and $v\tau(i)v^{-1}=r$. Thus, there exists no $r\in H$ which satisfies $k\rho(r)k^{-1}=g$ and $v\tau(i)v^{-1}=r$. Therefore, the sum $\sum\limits_{\substack{r\in H;\\ k\rho(r)k^{-1}=g;\\ v\tau(i)v^{-1}=r}} f(i)$ is an empty sum, and thus its value is 0. This proves (13.113.8).

On the other hand, let us show that

$$\sum_{\substack{r\in H;\\ k\rho(r)k^{-1}=g;\\ v\tau(i)v^{-1}=r}} f(i)=f(i) \tag{13.113.9}$$

for every $v\in H$ which satisfies $k\rho(v)\cdot\rho(\tau(i))\cdot(k\rho(v))^{-1}=g$.

Proof of (13.113.9): Let $v\in H$ be such that $k\rho(v)\cdot\rho(\tau(i))\cdot(k\rho(v))^{-1}=g$. If some $r\in H$ satisfies $k\rho(r)k^{-1}=g$ and $v\tau(i)v^{-1}=r$, then this r must equal $v\tau(i)v^{-1}$ (because $v\tau(i)v^{-1}=r$). Hence, there exists **at most one** $r\in H$ which satisfies $k\rho(r)k^{-1}=g$ and $v\tau(i)v^{-1}=r$.

On the other hand, the element $v\tau(i)v^{-1}$ of H satisfies

$$k\underbrace{\rho\left(v\tau(i)v^{-1}\right)}_{\substack{=\rho(v)\cdot\rho(\tau(i))\cdot\rho(v)^{-1}\\ \text{(since }\rho\text{ is a group}\\ \text{homomorphism)}}}k^{-1}=k\rho(v)\cdot\rho(\tau(i))\cdot\underbrace{\rho(v)^{-1}k^{-1}}_{=(k\rho(v))^{-1}}=k\rho(v)\cdot\rho(\tau(i))\cdot(k\rho(v))^{-1}=g$$

and $v\tau(i)v^{-1}=v\tau(i)v^{-1}$. In other words, $v\tau(i)v^{-1}$ is an element $r\in H$ which satisfies $k\rho(r)k^{-1}=g$ and $v\tau(i)v^{-1}=r$. Hence, there exists **at least one** $r\in H$ which satisfies $k\rho(r)k^{-1}=g$ and $v\tau(i)v^{-1}=r$ (namely, $v\tau(i)v^{-1}$). Since we also have shown that there exists **at most one** such $r\in H$, we can thus conclude that there exists **exactly one** $r\in H$ which satisfies $k\rho(r)k^{-1}=g$ and $v\tau(i)v^{-1}=r$. In other words, the sum $\sum\limits_{\substack{r\in H;\\ k\rho(r)k^{-1}=g;\\ v\tau(i)v^{-1}=r}} f(i)$ has precisely one addend. Hence, this sum rewrites as follows: $\sum\limits_{\substack{r\in H;\\ k\rho(r)k^{-1}=g;\\ v\tau(i)v^{-1}=r}} f(i)=f(i)$. This proves (13.113.9).

Now, (13.113.7) becomes

$$\begin{aligned}
&\sum_{\substack{r\in H;\\ k\rho(r)k^{-1}=g}}\ \sum_{\substack{v\in H;\\ v\tau(i)v^{-1}=r}} f(i)\\
&=\sum_{\substack{v\in H;\\ k\rho(v)\cdot\rho(\tau(i))\cdot(k\rho(v))^{-1}=g}}\ \underbrace{\sum_{\substack{r\in H;\\ k\rho(r)k^{-1}=g;\\ v\tau(i)v^{-1}=r}} f(i)}_{\substack{=f(i)\\ \text{(by (13.113.9))}}}+\sum_{\substack{v\in H;\\ k\rho(v)\cdot\rho(\tau(i))\cdot(k\rho(v))^{-1}\neq g}}\ \underbrace{\sum_{\substack{r\in H;\\ k\rho(r)k^{-1}=g;\\ v\tau(i)v^{-1}=r}} f(i)}_{\substack{=0\\ \text{(by (13.113.8))}}}\\
&=\sum_{\substack{v\in H;\\ k\rho(v)\cdot\rho(\tau(i))\cdot(k\rho(v))^{-1}=g}} f(i)+\underbrace{\sum_{\substack{v\in H;\\ k\rho(v)\cdot\rho(\tau(i))\cdot(k\rho(v))^{-1}\neq g}} 0}_{=0}=\sum_{\substack{v\in H;\\ k\rho(v)\cdot\rho(\tau(i))\cdot(k\rho(v))^{-1}=g}} f(i)\,.
\end{aligned}$$

This proves (13.113.6).

Thus, (13.113.5) becomes

$$\begin{aligned}
&\left(\operatorname{Ind}_\rho\left(\operatorname{Ind}_\tau f\right)\right)(g)\\
&= \frac{1}{|I|}\frac{1}{|H|} \underbrace{\sum_{k\in G}\sum_{i\in I}}_{=\sum_{i\in I}\sum_{k\in G}} \underbrace{\sum_{\substack{r\in H;\\ k\rho(r)k^{-1}=g}} \sum_{\substack{v\in H;\\ v\tau(i)v^{-1}=r}} f(i)}_{\substack{=\sum_{\substack{v\in H;\\ k\rho(v)\cdot\rho(\tau(i))\cdot(k\rho(v))^{-1}=g}} f(i)\\ \text{(by (13.113.6))}}}\\
&= \frac{1}{|I|}\frac{1}{|H|}\sum_{i\in I} \underbrace{\sum_{k\in G}\sum_{\substack{v\in H;\\ k\rho(v)\cdot\rho(\tau(i))\cdot(k\rho(v))^{-1}=g}}}_{=\sum_{v\in H}\sum_{\substack{k\in G;\\ k\rho(v)\cdot\rho(\tau(i))\cdot(k\rho(v))^{-1}=g}}} f(i)\\
&= \frac{1}{|I|}\frac{1}{|H|}\sum_{i\in I}\sum_{v\in H}\sum_{\substack{k\in G;\\ k\rho(v)\cdot\rho(\tau(i))\cdot(k\rho(v))^{-1}=g}} f(i).
\end{aligned}\tag{13.113.10}$$

However, for every $i\in I$ and $v\in H$, we have

$$\sum_{\substack{k\in G;\\ k\rho(v)\cdot\rho(\tau(i))\cdot(k\rho(v))^{-1}=g}} f(i) = \sum_{\substack{k\in G;\\ k\rho(\tau(i))k^{-1}=g}} f(i) \tag{13.113.11}$$

[859].

Thus, (13.113.10) becomes

$$\begin{aligned}
&\left(\operatorname{Ind}_\rho\left(\operatorname{Ind}_\tau f\right)\right)(g)\\
&= \frac{1}{|I|}\frac{1}{|H|}\sum_{i\in I}\sum_{v\in H} \underbrace{\sum_{\substack{k\in G;\\ k\rho(v)\cdot\rho(\tau(i))\cdot(k\rho(v))^{-1}=g}} f(i)}_{\substack{=\sum_{\substack{k\in G;\\ k\rho(\tau(i))k^{-1}=g}} f(i)\\ \text{(by (13.113.11))}}} = \frac{1}{|I|}\frac{1}{|H|}\sum_{i\in I} \underbrace{\sum_{v\in H}\sum_{\substack{k\in G;\\ k\rho(\tau(i))k^{-1}=g}} f(i)}_{=|H|\sum_{\substack{k\in G;\\ k\rho(\tau(i))k^{-1}=g}} f(i)}\\
&= \frac{1}{|I|}\frac{1}{|H|}\sum_{i\in I}|H|\sum_{\substack{k\in G;\\ k\rho(\tau(i))k^{-1}=g}} f(i) = \frac{1}{|I|}\underbrace{\frac{1}{|H|}|H|}_{=1} \underbrace{\sum_{i\in I}\sum_{\substack{k\in G;\\ k\rho(\tau(i))k^{-1}=g}}}_{\substack{=\sum_{\substack{(i,k)\in I\times G;\\ k\rho(\tau(i))k^{-1}=g}} = \sum_{\substack{(i,k)\in I\times G;\\ k(\rho\circ\tau)(i)k^{-1}=g}}\\ \text{(since } \rho(\tau(i))=(\rho\circ\tau)(i))}} f(i)\\
&= \frac{1}{|I|}\sum_{\substack{(i,k)\in I\times G;\\ k(\rho\circ\tau)(i)k^{-1}=g}} f(i) = \frac{1}{|I|}\sum_{\substack{(h,k)\in I\times G;\\ k(\rho\circ\tau)(h)k^{-1}=g}} f(h)\\
&\qquad\text{(here, we renamed the summation index } (i,k) \text{ as } (h,k)).
\end{aligned}$$

[859] *Proof of (13.113.11):* Let $i\in I$ and $v\in H$. We have $\rho(v)\in G$. Therefore, the map $G\to G,\ k\mapsto k\rho(v)$ is a bijection (since G is a group). Therefore, we can substitute $k\rho(v)$ for k in the sum $\sum_{\substack{k\in G;\\ k\rho(\tau(i))k^{-1}=g}} f(i)$. We thus obtain

$$\sum_{\substack{k\in G;\\ k\rho(\tau(i))k^{-1}=g}} f(i) = \sum_{\substack{k\in G;\\ k\rho(v)\rho(\tau(i))(k\rho(v))^{-1}=g}} f(i) = \sum_{\substack{k\in G;\\ k\rho(v)\cdot\rho(\tau(i))\cdot(k\rho(v))^{-1}=g}} f(i).$$

This proves (13.113.11).

Compared with

$$\left(\operatorname{Ind}_{\rho\circ\tau} f\right)(g) = \frac{1}{|I|} \sum_{\substack{(h,k)\in I\times G;\\ k(\rho\circ\tau)(h)k^{-1}=g}} f(h) \qquad \left(\text{by the definition of } \left(\operatorname{Ind}_{\rho\circ\tau} f\right)(g)\right),$$

this yields $\left(\operatorname{Ind}_\rho\left(\operatorname{Ind}_\tau f\right)\right)(g) = \left(\operatorname{Ind}_{\rho\circ\tau} f\right)(g)$.

Let us now forget that we fixed g. We thus have shown that $\left(\operatorname{Ind}_\rho\left(\operatorname{Ind}_\tau f\right)\right)(g) = \left(\operatorname{Ind}_{\rho\circ\tau} f\right)(g)$ for every $g \in G$. In other words, $\operatorname{Ind}_\rho \operatorname{Ind}_\tau f = \operatorname{Ind}_{\rho\circ\tau} f$. This solves Exercise 4.1.16(b).

Second solution to Exercise 4.1.16(b). We shall now give an alternative solution of Exercise 4.1.16(b) which relies on Exercise 4.1.16(d) and a certain fact about class functions:

Lemma 13.113.1. *Let G be a finite group. Let $u \in R_{\mathbb{C}}(G)$. Assume that every $v \in R_{\mathbb{C}}(G)$ satisfies $\langle u, v\rangle_G = 0$. Then, $u = 0$.*

Proof of Lemma 13.113.1. Let us use the Iverson bracket notation; that is, for any statement $\mathcal{A}$, we define $[\mathcal{A}]$ to be the integer $\begin{cases} 1, & \text{if } \mathcal{A} \text{ is true;} \\ 0, & \text{if } \mathcal{A} \text{ is false} \end{cases}$.

Fix an element $h \in G$. We define a map $\alpha_h : G \to \mathbb{C}$ by

$$\left(\alpha_h(g) = \sum_{k\in G} \left[khk^{-1} = g\right] \qquad \text{for every } g \in G\right).$$

(Notice that this map α_h is identical with the map $\alpha_{G,h}$ defined in Exercise 4.4.3, but we will not use this.) Then, $\alpha_h \in R_{\mathbb{C}}(G)$ [860]. Hence, $\langle u, \alpha_h\rangle_G$ is well-defined. But recall that every $v \in R_{\mathbb{C}}(G)$ satisfies

[860] *Proof.* Let p and q be two conjugate elements of G. Then, there exists some $r \in G$ such that $p = rqr^{-1}$ (since p and q are conjugate). Consider this r. The definition of α_h yields

$$\alpha_h(p) = \sum_{k\in G} \left[khk^{-1} = \underbrace{p}_{=rqr^{-1}}\right] = \sum_{k\in G} \left[\underbrace{khk^{-1} = rqr^{-1}}_{\substack{\text{this is equivalent to} \\ (r^{-1}khk^{-1}r=q)}}\right] = \sum_{k\in G} \left[r^{-1}khk^{-1}r = q\right].$$

Now, recall that $r \in G$. Hence, the map $G \to G,\ k \mapsto rk$ is a bijection (since G is a group). Hence, we can substitute rk for k in the sum $\sum_{k\in G} \left[r^{-1}khk^{-1}r = q\right]$. We thus obtain

$$\begin{aligned} \sum_{k\in G} \left[r^{-1}khk^{-1}r = q\right] &= \sum_{k\in G} \left[r^{-1}rkh \underbrace{(rk)^{-1}}_{=k^{-1}r^{-1}} r = q\right] = \sum_{k\in G} \left[\underbrace{r^{-1}r}_{=1} khk^{-1} \underbrace{r^{-1}r}_{=1} = q\right] \\ &= \sum_{k\in G} \left[khk^{-1} = q\right] = \alpha_h(q) \end{aligned}$$

(since $\alpha_h(q) = \sum_{k\in G} \left[khk^{-1} = q\right]$ (by the definition of α_h)). Hence, $\alpha_h(p) = \sum_{k\in G} \left[r^{-1}khk^{-1}r = q\right] = \alpha_h(q)$.

Let us now forget that we fixed p and q. We thus have shown that $\alpha_h(p) = \alpha_h(q)$ whenever p and q are two conjugate elements of G. In other words, the function α_h is constant on G-conjugacy classes. In other words, α_h is a class function on G (because the class functions on G are defined to be the functions $G \to \mathbb{C}$ which are constant on G-conjugacy classes). In other words, $\alpha_h \in R_{\mathbb{C}}(G)$ (since $R_{\mathbb{C}}(G)$ is the set of all class functions on G). Qed.

$\langle u, v\rangle_G = 0$. Applying this to $v = \alpha_h$, we obtain $\langle u, \alpha_h\rangle_G = 0$. Thus,

$$0 = \langle u, \alpha_h\rangle_G = \frac{1}{|G|} \sum_{g\in G} u(g) \underbrace{\alpha_h\left(g^{-1}\right)}_{\substack{=\sum_{k\in G}\left[khk^{-1}=g^{-1}\right] \\ \text{(by the definition of } \alpha_h\text{)}}} \qquad \text{(by the definition of the bilinear form } \langle\cdot,\cdot\rangle_G\text{)}$$

$$= \frac{1}{|G|} \sum_{g\in G} u(g) \sum_{k\in G} \left[khk^{-1} = g^{-1}\right] = \frac{1}{|G|} \sum_{k\in G} \sum_{g\in G} u(g) \left[\underbrace{khk^{-1} = g^{-1}}_{\text{this is equivalent to } \left(g=\left(khk^{-1}\right)^{-1}\right)}\right]$$

$$= \frac{1}{|G|} \sum_{k\in G} \sum_{g\in G} u(g) \left[g = \left(khk^{-1}\right)^{-1}\right]. \tag{13.113.12}$$

But every $k \in G$ satisfies

$$\sum_{g\in G} u(g) \left[g = \left(khk^{-1}\right)^{-1}\right] = u\left(h^{-1}\right) \tag{13.113.13}$$

[861]. Hence, (13.113.12) becomes

$$0 = \frac{1}{|G|} \sum_{k\in G} \underbrace{\sum_{g\in G} u(g) \left[g = \left(khk^{-1}\right)^{-1}\right]}_{\substack{=u\left(h^{-1}\right) \\ \text{(by (13.113.13))}}} = \frac{1}{|G|} \underbrace{\sum_{k\in G} u\left(h^{-1}\right)}_{=|G|u\left(h^{-1}\right)} = \frac{1}{|G|} |G| u\left(h^{-1}\right) = u\left(h^{-1}\right).$$

Thus, $u\left(h^{-1}\right) = 0$.

Let us now forget that we fixed h. We thus have shown that

$$u\left(h^{-1}\right) = 0 \qquad \text{for every } h \in G. \tag{13.113.15}$$

Let us now fix $h \in G$. Then, (13.113.15) (applied to h^{-1} instead of h) yields $u\left(\left(h^{-1}\right)^{-1}\right) = 0$. In other words, $u(h) = 0$ (since $\left(h^{-1}\right)^{-1} = h$).

Let us now forget that we fixed h. We thus have shown that $u(h) = 0$ for every $h \in G$. In other words, $u = 0$. This proves Lemma 13.113.1. □

[861] *Proof of (13.113.13):* Let us first recall that u belongs to the set $R_{\mathbb{C}}(G)$. In other words, u is a class function on G (since $R_{\mathbb{C}}(G)$ is the set of all class functions on G). In other words, the function u is constant on G-conjugacy classes (because the class functions on G are defined to be the functions $G \to \mathbb{C}$ which are constant on G-conjugacy classes). In other words,

$$u(p) = u(q) \tag{13.113.14}$$

whenever p and q are two conjugate elements of G.

Now, let $k \in G$. Then, the elements $kh^{-1}k^{-1}$ and h^{-1} of G are conjugate. Hence, $u\left(kh^{-1}k^{-1}\right) = u\left(h^{-1}\right)$ (by (13.113.14), applied to $p = kh^{-1}k^{-1}$ and $q = h^{-1}$).

All addends of the sum $\sum_{g\in G} u(g) \left[g = \left(khk^{-1}\right)^{-1}\right]$ are zero except for the addend for $g = \left(khk^{-1}\right)^{-1}$ (because the factor $\left[g = \left(khk^{-1}\right)^{-1}\right]$ is zero unless $g = \left(khk^{-1}\right)^{-1}$). Hence, this sum simplifies as follows:

$$\sum_{g\in G} u(g) \left[g = \left(khk^{-1}\right)^{-1}\right] = u\left(\underbrace{\left(khk^{-1}\right)^{-1}}_{=kh^{-1}k^{-1}}\right) \underbrace{\left[\left(khk^{-1}\right)^{-1} = \left(khk^{-1}\right)^{-1}\right]}_{=1} = u\left(kh^{-1}k^{-1}\right) = u\left(h^{-1}\right).$$

This proves (13.113.13).

Now, let us return to solving Exercise 4.1.16(b). Let $f \in R_{\mathbb{C}}(I)$. Let $v \in R_{\mathbb{C}}(G)$. Then,

$$\begin{aligned}
\left\langle \operatorname{Ind}_\rho \operatorname{Ind}_\tau f, v \right\rangle_G &= \left\langle \operatorname{Ind}_\tau f, \operatorname{Res}_\rho v \right\rangle_H \qquad \text{(by (4.1.17), applied to } \alpha = \operatorname{Ind}_\tau f \text{ and } \beta = v\text{)} \\
&= \left\langle f, \underbrace{\operatorname{Res}_\tau \operatorname{Res}_\rho v}_{\substack{= \operatorname{Res}_{\rho\circ\tau} v \\ \text{(by Exercise 4.1.16(d),} \\ \text{applied to } v \text{ instead of } f\text{)}}} \right\rangle_I \\
&\qquad \text{(by (4.1.17), applied to } H,\ I,\ \tau,\ f \text{ and } \operatorname{Res}_\rho v \text{ instead of } G,\ H,\ \rho,\ \alpha \text{ and } \beta\text{)} \\
&= \left\langle f, \operatorname{Res}_{\rho\circ\tau} v \right\rangle_I .
\end{aligned}$$

Compared with

$$\left\langle \operatorname{Ind}_{\rho\circ\tau} f, v \right\rangle_G = \left\langle f, \operatorname{Res}_{\rho\circ\tau} v \right\rangle_I \qquad \text{(by (4.1.17), applied to } I,\ \rho\circ\tau,\ f \text{ and } v \text{ instead of } H,\ \rho,\ \alpha \text{ and } \beta\text{)},$$

this yields $\left\langle \operatorname{Ind}_\rho \operatorname{Ind}_\tau f, v \right\rangle_G = \left\langle \operatorname{Ind}_{\rho\circ\tau} f, v \right\rangle_G$. Now, the form $\left\langle \cdot, \cdot \right\rangle_G$ is $\mathbb{C}$-bilinear, and therefore we have

$$\left\langle \operatorname{Ind}_\rho \operatorname{Ind}_\tau f - \operatorname{Ind}_{\rho\circ\tau} f, v \right\rangle_G = \underbrace{\left\langle \operatorname{Ind}_\rho \operatorname{Ind}_\tau f, v \right\rangle_G}_{= \left\langle \operatorname{Ind}_{\rho\circ\tau} f, v \right\rangle_G} - \left\langle \operatorname{Ind}_{\rho\circ\tau} f, v \right\rangle_G = \left\langle \operatorname{Ind}_{\rho\circ\tau} f, v \right\rangle_G - \left\langle \operatorname{Ind}_{\rho\circ\tau} f, v \right\rangle_G = 0.$$

Now, let us forget that we fixed v. We thus have shown that every $v \in R_{\mathbb{C}}(G)$ satisfies $\left\langle \operatorname{Ind}_\rho \operatorname{Ind}_\tau f - \operatorname{Ind}_{\rho\circ\tau} f, v \right\rangle_G = 0$. Hence, Lemma 13.113.1 (applied to $u = \operatorname{Ind}_\rho \operatorname{Ind}_\tau f - \operatorname{Ind}_{\rho\circ\tau} f$) yields $\operatorname{Ind}_\rho \operatorname{Ind}_\tau f - \operatorname{Ind}_{\rho\circ\tau} f = 0$. Thus, $\operatorname{Ind}_\rho \operatorname{Ind}_\tau f = \operatorname{Ind}_{\rho\circ\tau} f$. This solves Exercise 4.1.16(b) again.

[*Remark:* Recall that Exercise 4.1.14 (specifically, its parts (c), (d), (e) and (f)) shows that ρ-induction (of modules and of class functions) generalizes both the usual notion of induction and the fixed point construction. More precisely, ρ-induction becomes usual induction (at least up to isomorphism) when ρ is injective, and becomes fixed point construction when ρ is surjective. Exercise 4.1.16 (specifically, its parts (a) and (b)), on the other hand, shows that in the general case, ρ-induction can be reduced to a composition of usual induction and fixed point construction, because every group homomorphism $\rho : H \to G$ can be factored as a composition $\alpha \circ \beta$ of a surjective group homomorphism $\beta : H \to \overline{H}$ with an injective group homomorphism $\alpha : \overline{H} \to G$. This allows some alternative proofs of some parts of Exercise 4.1.14, which the interested reader can find.]

13.114. **Solution to Exercise 4.2.3.** *Solution to Exercise 4.2.3.* We need to prove that

$$\operatorname{Ind}_{\mathfrak{S}_{i+j}\times\mathfrak{S}_k}^{\mathfrak{S}_{i+j+k}} \left(\operatorname{Ind}_{\mathfrak{S}_i\times\mathfrak{S}_j}^{\mathfrak{S}_{i+j}} (U \otimes V) \otimes W \right) \cong \operatorname{Ind}_{\mathfrak{S}_i\times\mathfrak{S}_j\times\mathfrak{S}_k}^{\mathfrak{S}_{i+j+k}} (U \otimes V \otimes W) \tag{13.114.1}$$

and

$$\operatorname{Ind}_{\mathfrak{S}_i\times\mathfrak{S}_{j+k}}^{\mathfrak{S}_{i+j+k}} \left(U \otimes \operatorname{Ind}_{\mathfrak{S}_j\times\mathfrak{S}_k}^{\mathfrak{S}_{j+k}} (V \otimes W) \right) \cong \operatorname{Ind}_{\mathfrak{S}_i\times\mathfrak{S}_j\times\mathfrak{S}_k}^{\mathfrak{S}_{i+j+k}} (U \otimes V \otimes W) \tag{13.114.2}$$

as $\mathbb{C}[\mathfrak{S}_i \times \mathfrak{S}_j \times \mathfrak{S}_k]$-modules. We will only prove (13.114.1), since (13.114.2) is analogous.

It is easy to see that every finite group G and every $\mathbb{C}G$-module P satisfy $\operatorname{Ind}_G^G P \cong P$. Applied to $G = \mathfrak{S}_k$ and $P = W$, this yields $\operatorname{Ind}_{\mathfrak{S}_k}^{\mathfrak{S}_k} W \cong W$.

Now, Exercise 4.1.2 (applied to $\mathfrak{S}_{i+j+k}$, $\mathfrak{S}_{i+j} \times \mathfrak{S}_k$, $\mathfrak{S}_i \times \mathfrak{S}_j \times \mathfrak{S}_k$ and $U \otimes V \otimes W$ instead of G, H, I and U) yields

$$\operatorname{Ind}_{\mathfrak{S}_{i+j}\times\mathfrak{S}_k}^{\mathfrak{S}_{i+j+k}} \left(\operatorname{Ind}_{\mathfrak{S}_i\times\mathfrak{S}_j\times\mathfrak{S}_k}^{\mathfrak{S}_{i+j}\times\mathfrak{S}_k} (U \otimes V \otimes W) \right) \cong \operatorname{Ind}_{\mathfrak{S}_i\times\mathfrak{S}_j\times\mathfrak{S}_k}^{\mathfrak{S}_{i+j+k}} (U \otimes V \otimes W) .$$

Hence,

$$\begin{aligned}
\operatorname{Ind}_{\mathfrak{S}_i\times\mathfrak{S}_j\times\mathfrak{S}_k}^{\mathfrak{S}_{i+j+k}}\left(U\otimes V\otimes W\right) &\cong \operatorname{Ind}_{\mathfrak{S}_{i+j}\times\mathfrak{S}_k}^{\mathfrak{S}_{i+j+k}}\left(\underbrace{\operatorname{Ind}_{\mathfrak{S}_i\times\mathfrak{S}_j\times\mathfrak{S}_k}^{\mathfrak{S}_{i+j}\times\mathfrak{S}_k}\left(U\otimes V\otimes W\right)}_{\substack{\cong\left(\operatorname{Ind}_{\mathfrak{S}_i\times\mathfrak{S}_j}^{\mathfrak{S}_{i+j}}(U\otimes V)\right)\otimes\left(\operatorname{Ind}_{\mathfrak{S}_k}^{\mathfrak{S}_k}W\right)\\ \text{(by (4.1.6), applied to } G_1=\mathfrak{S}_{i+j},\ G_2=\mathfrak{S}_k,\\ H_1=\mathfrak{S}_i\times\mathfrak{S}_j,\ H_2=\mathfrak{S}_k,\ U_1=U\otimes V \text{ and } U_2=W)}}\right)\\
&\cong \operatorname{Ind}_{\mathfrak{S}_{i+j}\times\mathfrak{S}_k}^{\mathfrak{S}_{i+j+k}}\left(\left(\operatorname{Ind}_{\mathfrak{S}_i\times\mathfrak{S}_j}^{\mathfrak{S}_{i+j}}\left(U\otimes V\right)\right)\otimes\underbrace{\left(\operatorname{Ind}_{\mathfrak{S}_k}^{\mathfrak{S}_k}W\right)}_{\cong W}\right)\\
&\cong \operatorname{Ind}_{\mathfrak{S}_{i+j}\times\mathfrak{S}_k}^{\mathfrak{S}_{i+j+k}}\left(\operatorname{Ind}_{\mathfrak{S}_i\times\mathfrak{S}_j}^{\mathfrak{S}_{i+j}}\left(U\otimes V\right)\otimes W\right).
\end{aligned}$$

This proves (13.114.1). Thus, the solution of Exercise 4.2.3 is complete.

13.115. **Solution to Exercise 4.3.9.** *Solution to Exercise 4.3.9.* (a) Let $n\in\mathbb{N}$. Let B denote the subgroup of $GL_n(\mathbb{F})$ consisting of all upper-triangular matrices. Then, $GL_n(\mathbb{F})=\bigsqcup_{w\in\mathfrak{S}_n}BwB$. (Indeed, this can be proven in the same way as we have shown the equality $GL_n=\bigsqcup_{w\in\mathfrak{S}_n}BwB$ in our above proof of Proposition 4.3.7(c), with the only difference that we are now working over $\mathbb{F}$ instead of $\mathbb{F}_q$.)

Let w_0 denote the permutation in $\mathfrak{S}_n$ which sends every $i\in\{1,2,\ldots,n\}$ to $n+1-i$; then, $w_0^2=\operatorname{id}$. (The permutation w_0 is written as $(n,n-1,\ldots,1)$ in one-line notation.)

Let $A\in GL_n(\mathbb{F})$. We have $w_0A\in GL_n(\mathbb{F})=\bigsqcup_{w\in\mathfrak{S}_n}BwB$. Thus, there exists some $w\in\mathfrak{S}_n$ such that $w_0A\in BwB$. Consider this w. There exist two upper-triangular invertible matrices L' and U such that $w_0A=L'wU$ (since $w_0A\in BwB$). Consider these L' and U. The matrix $w_0L'w_0$ is lower-triangular (since L' is upper-triangular, and since w_0 is what it is). Set $L=w_0L'w_0$. Then, $\underbrace{L}_{=w_0L'w_0}w_0wU=w_0L'\underbrace{w_0w_0}_{=w_0^2=\operatorname{id}}wU=w_0\underbrace{L'wU}_{=w_0A}=\underbrace{w_0w_0}_{=w_0^2=\operatorname{id}}A=A$. In other words, $LPU=A$ with $P=w_0w$. Since P is clearly a permutation matrix (being the product of the permutation matrices w_0 and w), we thus have shown that $A=LPU$ for a lower-triangular matrix $L\in GL_n(\mathbb{F})$, an upper-triangular matrix $U\in GL_n(\mathbb{F})$ and a permutation matrix $P\in\mathfrak{S}_n\subset GL_n(\mathbb{F})$. This solves Exercise 4.3.9(a).

(b) In the proof that follows, we shall essentially mimic the arguments used to prove $GL_n=\bigsqcup_{w\in\mathfrak{S}_n}BwB$ in our proof of Proposition 4.3.7(c) (but we will add some details in a few places).

We recall that B_n is a subgroup of $GL_n(\mathbb{F})$. Hence, every element of B_n is invertible in B_n. Likewise, every element of B_m is invertible in B_m.

Let us first show that the disjoint union $\bigsqcup_{f\in F_{n,m}}B_nfB_m$ is well-defined. This means proving that the sets B_nfB_m for $f\in F_{n,m}$ are disjoint.

For every $A\in\mathbb{F}^{n\times m}$, every $i\in\{1,2,\ldots,n+1\}$ and every $j\in\{0,1,\ldots,m\}$, let $r_{i,j}(A)$ denote the rank of the matrix obtained by restricting A to the rows $i,i+1,\ldots,n$ and columns $1,2,\ldots,j$. This rank $r_{i,j}(A)$ does not change when we replace A by XA for some $X\in B_n$ (because replacing A by XA for some $X\in B_n$ means that we add to each row of A a linear combination of the rows further below; but this row transformation does not change $r_{i,j}(A)$); neither does it change when we replace A by AY for some $Y\in B_m$ (for a similar reason). Hence, for every $f\in F_{n,m}$, $X\in B_n$, $Y\in B_m$, $i\in\{1,2,\ldots,n+1\}$ and $j\in\{0,1,\ldots,m\}$, we have

$$r_{i,j}(f)=r_{i,j}(Xf)=r_{i,j}(XfY). \tag{13.115.1}$$

Now, we claim that if $f\in F_{n,m}$ and $A\in B_nfB_m$, then we can reconstruct f from A. Indeed, let $f\in F_{n,m}$ and $A\in B_nfB_m$. Then, there exist $X\in B_n$ and $Y\in B_m$ such that $A=XfY$ (since $A\in B_nfB_m$). Consider these X and Y. Then, (13.115.1) yields $r_{i,j}(f) = r_{i,j}\left(\underbrace{XfY}_{=A}\right) = r_{i,j}(A)$ for every $i \in \{1, 2, \ldots, n+1\}$ and $j \in \{0, 1, \ldots, m\}$. Thus, from A we can reconstruct the ranks $r_{i,j}(f)$ for all $i \in \{1, 2, \ldots, n+1\}$ and $j \in \{0, 1, \ldots, m\}$.

Notice that $f \in F_{n,m}$. Thus, f is a matrix in $\{0,1\}^{n\times m}$ such that each row of f contains at most one 1 and each column of f contains at most one 1.

Now, fix $i \in \{1, 2, \ldots, n\}$ and $j \in \{1, 2, \ldots, m\}$. Let $f_{i,j}$ be the entry of f in row i and column j. We will now show that

$$f_{i,j} = r_{i,j}(f) - r_{i,j-1}(f) - r_{i+1,j}(f) + r_{i+1,j-1}(f). \tag{13.115.2}$$

Proof of (13.115.2). First of all, if the first j entries of the i-th row of f are all 0, then (13.115.2) holds for obvious reasons (in fact, in this case, it is clear that $f_{i,j} = 0$, $r_{i,j}(f) = r_{i+1,j}(f)$ and $r_{i,j-1}(f) = r_{i+1,j-1}(f)$). We thus WLOG assume that the first j entries of the i-th row of f are not all 0. Then, there must be a 1 among these entries. It must lie in a different column than the 1's appearing in all other rows of f (because each column of f contains at most one 1). Therefore, the first row of the matrix obtained by restricting f to the rows $i, i+1, \ldots, n$ and columns $1, 2, \ldots, j$ is linearly independent from its other rows. Thus,

$$r_{i,j}(f) = r_{i+1,j}(f) + 1.$$

If $f_{i,j} = 0$, then the same argument yields $r_{i,j-1}(f) = r_{i+1,j-1}(f) + 1$ (because if $f_{i,j} = 0$, then the 1 among the first j entries of the i-th row of f must not be the last of these entries, and so it is one of the first $j-1$ entries of the i-th row of f). On the other hand, if $f_{i,j} = 1$, then we have $r_{i,j-1}(f) = r_{i+1,j-1}(f)$ (because if $f_{i,j} = 1$, then the first $j-1$ entries of the i-th row of f must be all 0 (since the j-th entry of this row is a 1, but each row of f contains at most one 1)). We can subsume both of these statements in one equation, which holds both if $f_{i,j} = 0$ and if $f_{i,j} = 1$: namely, we have

$$r_{i,j-1}(f) = \begin{cases} r_{i+1,j-1}(f) + 1 & \text{if } f_{i,j} = 0; \\ r_{i+1,j-1}(f) & \text{if } f_{i,j} = 1 \end{cases} = r_{i+1,j-1}(f) + (1 - f_{i,j}).$$

Subtracting this equality from $r_{i,j}(f) = r_{i+1,j}(f) + 1$, we obtain

$$r_{i,j}(f) - r_{i,j-1}(f) = (r_{i+1,j}(f) + 1) - (r_{i+1,j-1}(f) + (1 - f_{i,j})) = r_{i+1,j}(f) - r_{i+1,j-1}(f) + f_{i,j}.$$

This is easily seen to be equivalent to (13.115.2). Thus, (13.115.2) is proven.

Now, we can reconstruct the ranks $r_{i,j}(f)$, $r_{i,j-1}(f)$, $r_{i+1,j}(f)$ and $r_{i+1,j-1}(f)$ from A (since from A we can reconstruct the ranks $r_{i,j}(f)$ for all $i \in \{1, 2, \ldots, n+1\}$ and $j \in \{0, 1, \ldots, m\}$). Hence, we can reconstruct $f_{i,j}$ from A (due to (13.115.2)).

Let us now forget that we fixed i and j. We thus have seen that we can reconstruct $f_{i,j}$ from A for every $i \in \{1, 2, \ldots, n\}$ and $j \in \{1, 2, \ldots, m\}$. In other words, we can reconstruct f from A.

Thus we have proven that if $f \in F_{n,m}$ and $A \in B_n f B_m$, then we can reconstruct f from A. In other words, the sets $B_n f B_m$ for $f \in F_{n,m}$ are disjoint. Thus, the disjoint union $\bigsqcup_{f\in F_{n,m}} B_n f B_m$ is well-defined.

It remains to prove that this disjoint union $\bigsqcup_{f\in F_{n,m}} B_n f B_m$ is $\mathbb{F}^{n\times m}$. In order to do so, it is clearly enough to show that every $g \in \mathbb{F}^{n\times m}$ belongs to $\bigsqcup_{f\in F_{n,m}} B_n f B_m$.

Let $g \in \mathbb{F}^{n\times m}$. We need to show that $g \in \bigsqcup_{f\in F_{n,m}} B_n f B_m$. In other words, we need to find some $f \in F_{n,m}$ such that $g \in B_n f B_m$. In order to do so, we shall find a matrix $f \in F_{n,m}$ which lies in $B_n g B_m$. Once this is done, it will follow that $g \in B_n f B_m$ (since every element of B_n is invertible in B_n, and every element of B_m is invertible in B_m), and we will be done.

We refer to gB_m and $B_n g$ as cosets, despite the fact that they are not subsets of a group.

The freedom to alter g within the coset gB_m allows one to scale columns and add scalar multiples of earlier columns to later columns. We claim that using such column operations, one can always find a representative g' for coset gB_m in which

- the bottommost nonzero entry of each nonzero column is 1 (call this entry a *pivot*),
- the entries to right of each pivot within its row are all 0, and
- there is at most one pivot in each row and at most one pivot in each column (so that their positions are the positions of the 1's in some matrix $f \in F_{n,m}$).

In fact, we will see below that $B_n g B_m = B_n f B_m$ in this case. The algorithm which produces g' from g is simple: starting with the leftmost nonzero column, find its bottommost nonzero entry, and scale the column to make this entry a 1, creating the pivot in this column. Now use this pivot to clear out all entries in its row to its right, using column operations that subtract multiples of this column from later columns. Having done this, move on to the next nonzero column to the right, and repeat, scaling to create a pivot, and using it to eliminate entries to its right.[862]

Having found this g' in gB_m, a similar algorithm using left multiplication by B_n shows that f lies in $B_n g' \subset B_n g' B_m = B_n g B_m$. This time no scalings are required to create the pivot entries: starting with the bottommost nonzero row, one uses its pivot to eliminate all the entries above it in the same column by adding multiples of this row to higher rows. Then do the same using the pivot in the next-to-bottom nonzero row, etc.[863] The result is the matrix f.

13.116. **Solution to Exercise 4.3.11.** *Solution to Exercise 4.3.11.* (a) Let us begin this solution by stating some trivialities. It is easy to see that every finite group G and every $\mathbb{C}G$-module P satisfy

$$\operatorname{Ind}_G^G P \cong P \tag{13.116.1}$$

and

$$\operatorname{Infl}_G^G P \cong P. \tag{13.116.2}$$

[862] To see that this works, we need to check three facts:

- **(a)** We will find a nonzero entry in every nonzero column during our algorithm.
- **(b)** Our column operations preserve the zeroes lying to the right of already existing pivots.
- **(c)** Every row contains at most one pivot at the end of the algorithm.

But fact **(a)** is a tautology. Fact **(b)** holds because all our operations either scale columns (which clearly preserves zero entries) or subtract a multiple of the column c containing the current pivot from a later column d (which will preserve every zero lying to the right of an already existing pivot, because any already existing pivot must lie in a column $b < c$ and therefore both columns c and d have zeroes in its row). Fact **(c)** follows from noticing that the entries to the right of a pivot in its row are 0.

[863] One thing that requires verification is the fact that these row operations preserve the following three properties of g':

- the bottommost nonzero entry of each nonzero column is 1 (call this entry a *pivot*),
- the entries to right of each pivot within its row are all 0, and
- there is at most one pivot in each row and at most one pivot in each column (so that their positions are the positions of the 1's in some matrix $f \in F_{n,m}$).

Let us show this, and also show that the positions of the pivots are preserved. Indeed, it is clear that the positions of the pivots are preserved (because zero columns stay zero, nonzero columns stay nonzero, and the bottommost entries of nonzero columns do not move); therefore it is enough to prove the following fact:

- **(d)** Consider an *elimination step*, by which we mean a step in which a pivot in some position a is used to eliminate all the entries above it in the same column. Assume that, before the elimination step, the entries to the right of a within its row were all 0. Let b be the position of another pivot that existed before this elimination step. Assume that, before the elimination step, the pivot at b equalled 1, and the entries to the right of b within its row were all 0. Then, after the elimination step, b is still a position of a pivot and this pivot still equals 1, and the entries to the right of b within its row are still 0.

So let us prove fact **(d)**. We prove it by contradiction: Assume that it is not the case. Then, the elimination step must have changed at least one of the entries weakly to the right of b within its row (because we already know that b is still a position of a pivot after the elimination step). In particular, the elimination step must have changed at least one entry in the row of b. Thus, b must lie in a row strictly above a (since otherwise, the elimination step would have not changed any entry in the row of b). Let c be the position in the same row as b and in the same column as a. Then, the entry at c before the elimination step must not have been 0 (since otherwise, the elimination step would not have changed any entry in the row of b). As a consequence, c must not lie to the right of b within its row (because before the elimination step, the entries to the right of b within its row were all 0). Thus, c lies weakly to the left of b within its row. In other words, the column containing b lies weakly right of the column containing c. In other words, the column containing b lies weakly right of the column containing a (since the column containing c is the column containing a). Hence, the column containing b must lie **strictly** right of the column containing a (since otherwise, a and b would lie in the same column, which is absurd because a and b are two distinct pivots). But let us recall that, before the elimination step, the entries to the right of a within its row were all 0. Let us call these entries the *silent entries*. Now, what did the elimination step do to the entries weakly to the right of b within its row? It changed them by adding multiples of the corresponding entries of the row containing a. But all these corresponding entries were silent entries (because the column containing b lies strictly right of the column containing a) and thus were 0. Hence, the elimination step did not change the entries weakly to the right of b within its row. This contradicts the fact that the elimination step must have changed at least one of the entries weakly to the right of b within its row. This contradiction finishes the proof of fact **(d)**.

Furthermore, if $\mathcal{G}_1$ and $\mathcal{G}_2$ be two groups, if $\mathcal{K}_1 \triangleleft \mathcal{G}_1$ and $\mathcal{K}_2 \triangleleft \mathcal{G}_2$ are normal subgroups, if $\mathcal{U}_1$ is a $\mathbb{C}[\mathcal{G}_1/\mathcal{K}_1]$-module, and if $\mathcal{U}_2$ is a $\mathbb{C}[\mathcal{G}_2/\mathcal{K}_2]$-module, then

$$\operatorname{Infl}_{(\mathcal{G}_1/\mathcal{K}_1)\times(\mathcal{G}_2/\mathcal{K}_2)}^{\mathcal{G}_1\times\mathcal{G}_2} (\mathcal{U}_1 \otimes \mathcal{U}_2) \cong \left(\operatorname{Infl}_{\mathcal{G}_1/\mathcal{K}_1}^{\mathcal{G}_1} \mathcal{U}_1\right) \otimes \left(\operatorname{Infl}_{\mathcal{G}_2/\mathcal{K}_2}^{\mathcal{G}_2} \mathcal{U}_2\right) \tag{13.116.3}$$

as $\mathbb{C}[\mathcal{G}_1 \times \mathcal{G}_2]$-modules. (This is an analogue of (4.1.6), but is trivial to prove.) Also, if $\mathcal{G}$, $\mathcal{H}$ and $\mathcal{I}$ are three groups such that $\mathcal{I} < \mathcal{H} < \mathcal{G}$, and if $\mathcal{U}$ is a $\mathbb{C}\mathcal{I}$-module, then

$$\operatorname{Infl}_{\mathcal{H}}^{\mathcal{G}} \operatorname{Infl}_{\mathcal{I}}^{\mathcal{H}} \mathcal{U} = \operatorname{Infl}_{\mathcal{I}}^{\mathcal{G}} \mathcal{U}. \tag{13.116.4}$$

(This is an analogue of Exercise 4.1.2, and is again trivial[864].)

Let $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_\ell)$ be an almost-composition of an $n \in \mathbb{N}$ satisfying $\ell \geq 1$. Let V_i be a $\mathbb{C}G_i$-module for every $i \in \{1, 2, \ldots, \ell\}$. We need to prove the two isomorphisms

$$\begin{aligned} &\operatorname{ind}_{\alpha_1+\alpha_2+\cdots+\alpha_{\ell-1},\alpha_\ell}^{n} \left(\operatorname{ind}_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})}^{\alpha_1+\alpha_2+\cdots+\alpha_{\ell-1}} (V_1 \otimes V_2 \otimes \cdots \otimes V_{\ell-1}) \otimes V_\ell\right) \\ &\cong \operatorname{ind}_{\alpha}^{n} (V_1 \otimes V_2 \otimes \cdots \otimes V_\ell) \end{aligned} \tag{13.116.5}$$

and

$$\begin{aligned} &\operatorname{ind}_{\alpha_1,\alpha_2+\alpha_3+\cdots+\alpha_\ell}^{n} \left(V_1 \otimes \operatorname{ind}_{(\alpha_2,\alpha_3,\ldots,\alpha_\ell)}^{\alpha_2+\alpha_3+\cdots+\alpha_\ell} (V_2 \otimes V_3 \otimes \cdots \otimes V_\ell)\right) \\ &\cong \operatorname{ind}_{\alpha}^{n} (V_1 \otimes V_2 \otimes \cdots \otimes V_\ell). \end{aligned} \tag{13.116.6}$$

We will only prove (13.116.5), since the proof of (13.116.6) is analogous.

Let m be the nonnegative integer $\alpha_1 + \alpha_2 + \cdots + \alpha_{\ell-1}$. Then, $(\alpha_1, \alpha_2, \ldots, \alpha_{\ell-1})$ is an almost-composition of m, and we have $\alpha_1 + \alpha_2 + \cdots + \alpha_{\ell-1} = m$.

Let W denote the $\mathbb{C}[G_1 \times G_2 \times \cdots \times G_{\ell-1}]$-module $V_1 \otimes V_2 \otimes \cdots \otimes V_{\ell-1}$. Then, $V_1 \otimes V_2 \otimes \cdots \otimes V_{\ell-1} = W$ and $V_1 \otimes V_2 \otimes \cdots \otimes V_\ell = \underbrace{(V_1 \otimes V_2 \otimes \cdots \otimes V_{\ell-1})}_{=W} \otimes V_\ell = W \otimes V_\ell$ and $\alpha_1 + \alpha_2 + \cdots + \alpha_{\ell-1} = m$. Hence, the relation (13.116.5) (which we want to prove) rewrites as

$$\operatorname{ind}_{m,\alpha_\ell}^{n} \left(\operatorname{ind}_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})}^{m} W \otimes V_\ell\right) \cong \operatorname{ind}_{\alpha}^{n} (W \otimes V_\ell). \tag{13.116.7}$$

It thus remains to prove (13.116.7).

We distinguish between two cases:

Case 1: We have $G_* = \mathfrak{S}_*$ or $G_* = \mathfrak{S}_*[\Gamma]$.

Case 2: We have $G_* = GL_*$.

Let us consider Case 1 first. In this case, we have $G_* = \mathfrak{S}_*$ or $G_* = \mathfrak{S}_*[\Gamma]$. Thus, $\operatorname{ind}_{\beta}^{N} = \operatorname{Ind}_{G_\beta}^{G_N}$ for every $N \in \mathbb{N}$ and every almost-composition β of N (by the definition of $\operatorname{ind}_{\beta}^{N}$), and $\operatorname{ind}_{i,j}^{N} = \operatorname{Ind}_{G_i \times G_j}^{G_N}$ for every $N \in \mathbb{N}$ and every $i, j \in \mathbb{N}$ satisfying $i + j = N$ (by the definition of $\operatorname{ind}_{i,j}^{N}$).

[864] Note that the equality sign in (13.116.4) is a honest equality, not just a canonical isomorphism.

Since $\operatorname{ind}_\beta^N = \operatorname{Ind}_{G_\beta}^{G_N}$ for every $N \in \mathbb{N}$ and every almost-composition β of N, we have $\operatorname{ind}_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})}^m = \operatorname{Ind}_{G_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})}}^{G_m}$. Thus,

$$\begin{aligned}
&\operatorname{ind}_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})}^m W \otimes V_\ell \\
&= \underbrace{\operatorname{Ind}_{G_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})}}^{G_m} W}_{\substack{=\operatorname{Ind}_{G_{\alpha_1}\times G_{\alpha_2}\times\cdots\times G_{\alpha_{\ell-1}}}^{G_m} W \\ (\text{since } G_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})}=G_{\alpha_1}\times G_{\alpha_2}\times\cdots\times G_{\alpha_{\ell-1}})}} \otimes \underbrace{V_\ell}_{\substack{\cong\operatorname{Ind}_{G_{\alpha_\ell}}^{G_{\alpha_\ell}} V_\ell \\ (\text{because (13.116.1) (applied} \\ \text{to } G=G_{\alpha_\ell} \text{ and } P=V_\ell) \text{ yields } \operatorname{Ind}_{G_{\alpha_\ell}}^{G_{\alpha_\ell}} V_\ell\cong V_\ell)}} \\
&\cong \left(\operatorname{Ind}_{G_{\alpha_1}\times G_{\alpha_2}\times\cdots\times G_{\alpha_{\ell-1}}}^{G_m} W\right) \otimes \left(\operatorname{Ind}_{G_{\alpha_\ell}}^{G_{\alpha_\ell}} V_\ell\right) \cong \operatorname{Ind}_{G_{\alpha_1}\times G_{\alpha_2}\times\cdots\times G_{\alpha_{\ell-1}}\times G_{\alpha_\ell}}^{G_m\times G_{\alpha_\ell}} (W\otimes V_\ell) \\
&\qquad \left(\begin{array}{c} \text{since Exercise 4.1.3 (applied to } G_m,\ G_{\alpha_\ell},\ G_{\alpha_1}\times G_{\alpha_2}\times\cdots\times G_{\alpha_{\ell-1}},\ G_{\alpha_\ell}, \\ W \text{ and } V_\ell \text{ instead of } G_1,\ G_2,\ H_1,\ H_2,\ U_1 \text{ and } U_2\text{) yields} \\ \operatorname{Ind}_{G_{\alpha_1}\times G_{\alpha_2}\times\cdots\times G_{\alpha_{\ell-1}}\times G_{\alpha_\ell}}^{G_{\alpha_1+\alpha_2+\cdots+\alpha_{\ell-1}}\times G_{\alpha_\ell}} (W\otimes V_\ell) \\ \cong \left(\operatorname{Ind}_{G_{\alpha_1}\times G_{\alpha_2}\times\cdots\times G_{\alpha_{\ell-1}}}^{G_{\alpha_1+\alpha_2+\cdots+\alpha_{\ell-1}}} W\right)\otimes\left(\operatorname{Ind}_{G_{\alpha_\ell}}^{G_{\alpha_\ell}} V_\ell\right) \end{array}\right) \\
&= \operatorname{Ind}_{G_{\alpha_1}\times G_{\alpha_2}\times\cdots\times G_{\alpha_\ell}}^{G_m\times G_{\alpha_\ell}} (W\otimes V_\ell). \qquad (13.116.8)
\end{aligned}$$

But from $m = \alpha_1 + \alpha_2 + \cdots + \alpha_{\ell-1}$, we obtain $m + \alpha_\ell = (\alpha_1 + \alpha_2 + \cdots + \alpha_{\ell-1}) + \alpha_\ell = \alpha_1 + \alpha_2 + \cdots + \alpha_\ell = n$ (since $(\alpha_1, \alpha_2, \ldots, \alpha_\ell)$ is an almost-composition of n). Thus, since $\operatorname{ind}_{i,j}^N = \operatorname{Ind}_{G_i\times G_j}^{G_N}$ for every $N \in \mathbb{N}$ and every $i, j \in \mathbb{N}$ satisfying $i + j = N$, we have $\operatorname{ind}_{m,\alpha_\ell}^n = \operatorname{Ind}_{G_m\times G_{\alpha_\ell}}^{G_n}$. Thus,

$$\begin{aligned}
&\operatorname{ind}_{m,\alpha_\ell}^n \left(\operatorname{ind}_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})}^m W\otimes V_\ell\right) \\
&= \operatorname{Ind}_{G_m\times G_{\alpha_\ell}}^{G_n} \left(\underbrace{\operatorname{ind}_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})}^m W\otimes V_\ell}_{\substack{\cong\operatorname{Ind}_{G_{\alpha_1}\times G_{\alpha_2}\times\cdots\times G_{\alpha_\ell}}^{G_m\times G_{\alpha_\ell}} (W\otimes V_\ell) \\ (\text{by (13.116.8)})}}\right) \cong \operatorname{Ind}_{G_m\times G_{\alpha_\ell}}^{G_n} \operatorname{Ind}_{G_{\alpha_1}\times G_{\alpha_2}\times\cdots\times G_{\alpha_\ell}}^{G_m\times G_{\alpha_\ell}} (W\otimes V_\ell) \\
&\cong \operatorname{Ind}_{G_{\alpha_1}\times G_{\alpha_2}\times\cdots\times G_{\alpha_\ell}}^{G_n} (W\otimes V_\ell) \qquad \left(\begin{array}{c} \text{by Exercise 4.1.2, applied to } G=G_n,\ H=G_m\times G_{\alpha_\ell}, \\ I=G_{\alpha_1}\times G_{\alpha_2}\times\cdots\times G_{\alpha_\ell} \text{ and } U=W\otimes V_\ell \end{array}\right) \\
&= \operatorname{ind}_\alpha^n (W\otimes V_\ell) \qquad \left(\text{since } \operatorname{Ind}_{G_{\alpha_1}\times G_{\alpha_2}\times\cdots\times G_{\alpha_\ell}}^{G_n} = \operatorname{ind}_\alpha^n \text{ (by the definition of } \operatorname{ind}_\alpha^n)\right).
\end{aligned}$$

This proves (13.116.5). Hence, (13.116.5) is proven in Case 1.

Let us now consider Case 2. In this case, we have $G_* = GL_*$. Hence, $\operatorname{ind}_\beta^N = \operatorname{Ind}_{P_\beta}^{G_N} \operatorname{Infl}_{G_\beta}^{P_\beta}$ for every $N \in \mathbb{N}$ and every almost-composition β of N (by the definition of $\operatorname{ind}_\beta^N$), and $\operatorname{ind}_{i,j}^N = \operatorname{Ind}_{P_{i,j}}^{G_N} \operatorname{Infl}_{G_i\times G_j}^{P_{i,j}}$ for every $N \in \mathbb{N}$ and every $i, j \in \mathbb{N}$ satisfying $i + j = N$ (by the definition of $\operatorname{ind}_{i,j}^N$).

Since $\operatorname{ind}_\beta^N = \operatorname{Ind}_{P_\beta}^{G_N} \operatorname{Infl}_{G_\beta}^{P_\beta}$ for every $N \in \mathbb{N}$ and every almost-composition β of N, we have

$$\operatorname{ind}_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})}^m = \operatorname{Ind}_{P_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})}}^{G_m} \operatorname{Infl}_{G_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})}}^{P_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})}} = \operatorname{Ind}_{P_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})}}^{G_m} \operatorname{Infl}_{G_{\alpha_1}\times G_{\alpha_2}\times\cdots\times G_{\alpha_{\ell-1}}}^{P_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})}}$$

(since $G_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})} = G_{\alpha_1} \times G_{\alpha_2} \times \cdots \times G_{\alpha_{\ell-1}}$). Therefore,

$$\begin{align*}
&\operatorname{ind}^m_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})} W \otimes V_\ell \\
&= \operatorname{Ind}^{G_m}_{P_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})}} \operatorname{Infl}^{P_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})}}_{G_{\alpha_1}\times G_{\alpha_2}\times\cdots\times G_{\alpha_{\ell-1}}} W \otimes \underbrace{V_\ell}_{\substack{\cong \operatorname{Ind}^{G_{\alpha_\ell}}_{G_{\alpha_\ell}} V_\ell \\ \text{(because (13.116.1) (applied} \\ \text{to } G=G_{\alpha_\ell} \text{ and } P=V_\ell\text{) yields } \operatorname{Ind}^{G_{\alpha_\ell}}_{G_{\alpha_\ell}} V_\ell \cong V_\ell\text{)}}} \\
&\cong \left(\operatorname{Ind}^{G_m}_{P_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})}} \operatorname{Infl}^{P_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})}}_{G_{\alpha_1}\times G_{\alpha_2}\times\cdots\times G_{\alpha_{\ell-1}}} W \right) \otimes \left(\operatorname{Ind}^{G_{\alpha_\ell}}_{G_{\alpha_\ell}} V_\ell \right) \\
&\cong \operatorname{Ind}^{G_m\times G_{\alpha_\ell}}_{P_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})}\times G_{\alpha_\ell}} \left(\operatorname{Infl}^{P_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})}}_{G_{\alpha_1}\times G_{\alpha_2}\times\cdots\times G_{\alpha_{\ell-1}}} W \otimes \underbrace{V_\ell}_{\substack{\cong \operatorname{Infl}^{G_{\alpha_\ell}}_{G_{\alpha_\ell}} V_\ell \\ \text{(because (13.116.1) (applied} \\ \text{to } G=G_{\alpha_\ell} \text{ and } P=V_\ell\text{) yields } \operatorname{Infl}^{G_{\alpha_\ell}}_{G_{\alpha_\ell}} V_\ell \cong V_\ell\text{)}}} \right) \\
&\qquad \left(\begin{array}{c} \text{since Exercise 4.1.3 (applied to } G_m,\ G_{\alpha_\ell},\ P_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})},\ G_{\alpha_\ell}, \\ \operatorname{Infl}^{P_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})}}_{G_{\alpha_1}\times G_{\alpha_2}\times\cdots\times G_{\alpha_{\ell-1}}} W \text{ and } V_\ell \text{ instead of } G_1,\ G_2,\ H_1,\ H_2,\ U_1 \text{ and } U_2\text{) yields} \\ \operatorname{Ind}^{G_m\times G_{\alpha_\ell}}_{P_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})}\times G_{\alpha_\ell}} \left(\operatorname{Infl}^{P_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})}}_{G_{\alpha_1}\times G_{\alpha_2}\times\cdots\times G_{\alpha_{\ell-1}}} W \otimes V_\ell \right) \\ = \left(\operatorname{Ind}^{G_m}_{P_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})}} \operatorname{Infl}^{P_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})}}_{G_{\alpha_1}\times G_{\alpha_2}\times\cdots\times G_{\alpha_{\ell-1}}} W \right) \otimes \left(\operatorname{Ind}^{G_{\alpha_\ell}}_{G_{\alpha_\ell}} V_\ell \right) \end{array} \right) \\
&\cong \operatorname{Ind}^{G_m\times G_{\alpha_\ell}}_{P_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})}\times G_{\alpha_\ell}} \underbrace{\left(\operatorname{Infl}^{P_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})}}_{G_{\alpha_1}\times G_{\alpha_2}\times\cdots\times G_{\alpha_{\ell-1}}} W \otimes \operatorname{Infl}^{G_{\alpha_\ell}}_{G_{\alpha_\ell}} V_\ell \right)}_{\substack{\cong \operatorname{Infl}^{P_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})}\times G_{\alpha_\ell}}_{G_{\alpha_1}\times G_{\alpha_2}\times\cdots\times G_{\alpha_{\ell-1}}\times G_{\alpha_\ell}} (W\otimes V_\ell) \\ \text{(by (13.116.3), applied to} \\ \mathcal{G}_1 = P_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})},\ \mathcal{G}_2 = G_{\alpha_\ell},\ \mathcal{K}_1 = K_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})}, \\ \mathcal{K}_2 = G_{\alpha_\ell},\ \mathcal{U}_1 = W \text{ and } \mathcal{U}_2 = V_\ell \\ \text{(because } K_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})} \text{ is a normal subgroup} \\ \text{of } P_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})} \text{ and the quotient is} \\ P_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})}/K_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})} = G_{\alpha_1}\times G_{\alpha_2}\times\cdots\times G_{\alpha_{\ell-1}}\text{))}}} \\
&\cong \operatorname{Ind}^{G_m\times G_{\alpha_\ell}}_{P_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})}\times G_{\alpha_\ell}} \operatorname{Infl}^{P_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})}\times G_{\alpha_\ell}}_{G_{\alpha_1}\times G_{\alpha_2}\times\cdots\times G_{\alpha_{\ell-1}}\times G_{\alpha_\ell}} (W \otimes V_\ell) . \tag{13.116.9}
\end{align*}$$

But from $m = \alpha_1 + \alpha_2 + \cdots + \alpha_{\ell-1}$, we obtain $m + \alpha_\ell = (\alpha_1 + \alpha_2 + \cdots + \alpha_{\ell-1}) + \alpha_\ell = \alpha_1 + \alpha_2 + \cdots + \alpha_\ell = n$ (since $(\alpha_1, \alpha_2, \ldots, \alpha_\ell)$ is an almost-composition of n). Thus, since $\operatorname{ind}^N_{i,j} = \operatorname{Ind}^{G_N}_{P_{i,j}} \operatorname{Infl}^{P_{i,j}}_{G_i\times G_j}$ for every $N \in \mathbb{N}$ and every $i, j \in \mathbb{N}$ satisfying $i + j = N$, we have $\operatorname{ind}^n_{m,\alpha_\ell} = \operatorname{Ind}^{G_n}_{P_{m,\alpha_\ell}} \operatorname{Infl}^{P_{m,\alpha_\ell}}_{G_m\times G_{\alpha_\ell}}$, so that

$$\begin{align*}
&\operatorname{ind}^n_{m,\alpha_\ell} \left(\operatorname{ind}^m_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})} W \otimes V_\ell \right) \\
&= \operatorname{Ind}^{G_n}_{P_{m,\alpha_\ell}} \operatorname{Infl}^{P_{m,\alpha_\ell}}_{G_m\times G_{\alpha_\ell}} \left(\underbrace{\operatorname{ind}^m_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})} W \otimes V_\ell}_{\substack{\cong \operatorname{Ind}^{G_m\times G_{\alpha_\ell}}_{P_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})}\times G_{\alpha_\ell}} \operatorname{Infl}^{P_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})}\times G_{\alpha_\ell}}_{G_{\alpha_1}\times G_{\alpha_2}\times\cdots\times G_{\alpha_{\ell-1}}\times G_{\alpha_\ell}} (W\otimes V_\ell) \\ \text{(by (13.116.9))}}} \right) \\
&\cong \operatorname{Ind}^{G_n}_{P_{m,\alpha_\ell}} \operatorname{Infl}^{P_{m,\alpha_\ell}}_{G_m\times G_{\alpha_\ell}} \operatorname{Ind}^{G_m\times G_{\alpha_\ell}}_{P_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})}\times G_{\alpha_\ell}} \operatorname{Infl}^{P_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})}\times G_{\alpha_\ell}}_{G_{\alpha_1}\times G_{\alpha_2}\times\cdots\times G_{\alpha_{\ell-1}}\times G_{\alpha_\ell}} (W \otimes V_\ell) \tag{13.116.10}
\end{align*}$$

Now, we are going to prove that

$$\operatorname{Infl}_{G_m \times G_{\alpha_\ell}}^{P_{m,\alpha_\ell}} \operatorname{Ind}_{P_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})} \times G_{\alpha_\ell}}^{G_m \times G_{\alpha_\ell}} Z \cong \operatorname{Ind}_{P_\alpha}^{P_{m,\alpha_\ell}} \operatorname{Infl}_{P_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})} \times G_{\alpha_\ell}}^{P_\alpha} Z \tag{13.116.11}$$

for every $\mathbb{C}\left[P_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})} \times G_{\alpha_\ell}\right]$-module Z. In order to do so, we try to apply Exercise 4.1.11 to $G = P_{m,\alpha_\ell}$, $H = P_\alpha$, $K = K_{m,\alpha_\ell}$ and $V = Z$. This yields (13.116.11) if we can prove the following two statements:

(1) We have $K_{m,\alpha_\ell} < P_\alpha < P_{m,\alpha_\ell}$ and $K_{m,\alpha_\ell} \triangleleft P_{m,\alpha_\ell}$.
(2) The quotient $P_{m,\alpha_\ell}/K_{m,\alpha_\ell}$ is canonically identified with $G_m \times G_{\alpha_\ell}$ in such a way that its subgroup $P_\alpha/K_{m,\alpha_\ell}$ is canonically identified with $P_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})} \times G_{\alpha_\ell}$.

The first of these two statements is clear (using $m = \alpha_1 + \alpha_2 + \cdots + \alpha_{\ell-1}$ and $\alpha = (\alpha_1, \alpha_2, ..., \alpha_\ell)$). It remains to prove the second statement. We know how $P_{m,\alpha_\ell}/K_{m,\alpha_\ell}$ is identified with $G_m \times G_{\alpha_\ell}$ already. The thing that we need to prove is that the subgroup $P_\alpha/K_{m,\alpha_\ell}$ of $P_{m,\alpha_\ell}/K_{m,\alpha_\ell}$ is canonically identified with $P_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})} \times G_{\alpha_\ell}$. In other words, we need to prove that the projection of the subgroup P_α of P_{m,α_ℓ} onto $G_m \times G_{\alpha_\ell}$ is precisely $P_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})} \times G_{\alpha_\ell}$. But this is obvious. Hence, the second statement is proven, and so we are able to apply Exercise 4.1.11 and therefore obtain (13.116.11).

Now, (13.116.10) becomes

$$\begin{aligned}
&\operatorname{ind}_{m,\alpha_\ell}^{n}\left(\operatorname{ind}_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})}^{m} W \otimes V_\ell\right) \\
&\cong \operatorname{Ind}_{P_{m,\alpha_\ell}}^{G_n} \underbrace{\operatorname{Infl}_{G_m \times G_{\alpha_\ell}}^{P_{m,\alpha_\ell}} \operatorname{Ind}_{P_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})} \times G_{\alpha_\ell}}^{G_m \times G_{\alpha_\ell}} \operatorname{Infl}_{G_{\alpha_1} \times G_{\alpha_2} \times \cdots \times G_{\alpha_{\ell-1}} \times G_{\alpha_\ell}}^{P_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})} \times G_{\alpha_\ell}} (W \otimes V_\ell)}_{\substack{\cong \operatorname{Ind}_{P_\alpha}^{P_{m,\alpha_\ell}} \operatorname{Infl}_{P_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})} \times G_{\alpha_\ell}}^{P_\alpha} \operatorname{Infl}_{G_{\alpha_1} \times G_{\alpha_2} \times \cdots \times G_{\alpha_{\ell-1}} \times G_{\alpha_\ell}}^{P_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})} \times G_{\alpha_\ell}} (W \otimes V_\ell) \\ \text{(by (13.116.11), applied to } Z = \operatorname{Infl}_{G_{\alpha_1} \times G_{\alpha_2} \times \cdots \times G_{\alpha_{\ell-1}} \times G_{\alpha_\ell}}^{P_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})} \times G_{\alpha_\ell}} (W \otimes V_\ell))}} \\
&\cong \operatorname{Ind}_{P_{m,\alpha_\ell}}^{G_n} \operatorname{Ind}_{P_\alpha}^{P_{m,\alpha_\ell}} \operatorname{Infl}_{P_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})} \times G_{\alpha_\ell}}^{P_\alpha} \operatorname{Infl}_{G_{\alpha_1} \times G_{\alpha_2} \times \cdots \times G_{\alpha_{\ell-1}} \times G_{\alpha_\ell}}^{P_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})} \times G_{\alpha_\ell}} (W \otimes V_\ell) \\
&\cong \operatorname{Ind}_{P_\alpha}^{G_n} \underbrace{\operatorname{Infl}_{P_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})} \times G_{\alpha_\ell}}^{P_\alpha} \operatorname{Infl}_{G_{\alpha_1} \times G_{\alpha_2} \times \cdots \times G_{\alpha_{\ell-1}} \times G_{\alpha_\ell}}^{P_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})} \times G_{\alpha_\ell}} (W \otimes V_\ell)}_{\substack{= \operatorname{Infl}_{G_{\alpha_1} \times G_{\alpha_2} \times \cdots \times G_{\alpha_{\ell-1}} \times G_{\alpha_\ell}}^{P_\alpha} (W \otimes V_\ell) \\ \text{(by (13.116.4), applied to } \mathcal{G} = P_{(\alpha_1,\alpha_2,\ldots,\alpha_\ell)}, \\ \mathcal{H} = P_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})} \times G_{\alpha_\ell},\ \mathcal{I} = G_{\alpha_1} \times G_{\alpha_2} \times \cdots \times G_{\alpha_{\ell-1}} \times G_{\alpha_\ell} \text{ and } \mathcal{U} = W \otimes V_\ell)}} \\
&\qquad \left(\begin{array}{c} \text{by Exercise 4.1.2, applied to } G = G_n,\ H = P_{m,\alpha_\ell}, \\ I = P_\alpha \text{ and } U = \operatorname{Infl}_{P_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})} \times G_{\alpha_\ell}}^{P_\alpha} \operatorname{Infl}_{G_{\alpha_1} \times G_{\alpha_2} \times \cdots \times G_{\alpha_{\ell-1}} \times G_{\alpha_\ell}}^{P_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})} \times G_{\alpha_\ell}} (W \otimes V_\ell) \end{array}\right) \\
&= \operatorname{Ind}_{P_\alpha}^{G_n} \operatorname{Infl}_{G_{\alpha_1} \times G_{\alpha_2} \times \cdots \times G_{\alpha_{\ell-1}} \times G_{\alpha_\ell}}^{P_\alpha} (W \otimes V_\ell) \\
&= \operatorname{Ind}_{P_\alpha}^{G_n} \operatorname{Infl}_{G_\alpha}^{P_\alpha} (W \otimes V_\ell) \qquad \left(\text{since } G_{\alpha_1} \times G_{\alpha_2} \times \cdots \times G_{\alpha_{\ell-1}} \times G_{\alpha_\ell} = G_{\alpha_1} \times G_{\alpha_2} \times \cdots \times G_{\alpha_\ell} = G_\alpha\right) \\
&= \operatorname{ind}_\alpha^n (W \otimes V_\ell) \qquad \left(\text{since } \operatorname{Ind}_{P_\alpha}^{G_n} \operatorname{Infl}_{G_\alpha}^{P_\alpha} = \operatorname{ind}_\alpha^n \text{ (by the definition of } \operatorname{ind}_\alpha^n\text{)}\right).
\end{aligned}$$

This proves (13.116.5). Hence, (13.116.5) is proven in Case 2.

Now, (13.116.5) is proven in both Cases 1 and 2. Hence, (13.116.5) always holds. This completes the solution to Exercise 4.3.11(a).

(b) *Alternative solution to Exercise 4.2.3:* Let G_* be the tower $\mathfrak{S}_*$ of groups. We can then apply Exercise 4.3.11(a) to $\ell = 3$, $n = i + j + k$, $\alpha = (i, j, k)$, $V_1 = U$, $V_2 = V$ and $V_3 = W$. As the result, we obtain

$$\begin{aligned}
\operatorname{ind}_{i+j,k}^{i+j+k}\left(\operatorname{ind}_{(i,j)}^{i+j} (U \otimes V) \otimes W\right) &\cong \operatorname{ind}_{(i,j,k)}^{i+j+k} (U \otimes V \otimes W) \\
&\cong \operatorname{ind}_{i,j+k}^{i+j+k}\left(U \otimes \operatorname{ind}_{(j,k)}^{j+k} (V \otimes W)\right).
\end{aligned}$$

Since $\operatorname{ind}_{i+j,k}^{i+j+k} = \operatorname{Ind}_{\mathfrak{S}_{i+j}\times\mathfrak{S}_k}^{\mathfrak{S}_{i+j+k}}$, $\operatorname{ind}_{(i,j)}^{i+j} = \operatorname{ind}_{i,j}^{i+j} = \operatorname{Ind}_{\mathfrak{S}_i\times\mathfrak{S}_j}^{\mathfrak{S}_{i+j}}$, $\operatorname{ind}_{(i,j,k)}^{i+j+k} = \operatorname{Ind}_{G_{i,j,k}}^{\mathfrak{S}_{i+j+k}} = \operatorname{Ind}_{\mathfrak{S}_i\times\mathfrak{S}_j\times\mathfrak{S}_k}^{\mathfrak{S}_{i+j+k}}$ (because $G_{i,j,k} = \mathfrak{S}_i\times\mathfrak{S}_j\times\mathfrak{S}_k$), $\operatorname{ind}_{i,j+k}^{i+j+k} = \operatorname{Ind}_{\mathfrak{S}_i\times\mathfrak{S}_{j+k}}^{\mathfrak{S}_{i+j+k}}$ and $\operatorname{ind}_{(j,k)}^{j+k} = \operatorname{ind}_{j,k}^{j+k} = \operatorname{Ind}_{\mathfrak{S}_j\times\mathfrak{S}_k}^{\mathfrak{S}_{j+k}}$, this rewrites as follows:

$$\begin{aligned}\operatorname{Ind}_{\mathfrak{S}_{i+j}\times\mathfrak{S}_k}^{\mathfrak{S}_{i+j+k}}\left(\operatorname{Ind}_{\mathfrak{S}_i\times\mathfrak{S}_j}^{\mathfrak{S}_{i+j}}(U\otimes V)\otimes W\right) &\cong \operatorname{Ind}_{\mathfrak{S}_i\times\mathfrak{S}_j\times\mathfrak{S}_k}^{\mathfrak{S}_{i+j+k}}(U\otimes V\otimes W)\\ &\cong \operatorname{Ind}_{\mathfrak{S}_i\times\mathfrak{S}_{j+k}}^{\mathfrak{S}_{i+j+k}}\left(U\otimes \operatorname{Ind}_{\mathfrak{S}_j\times\mathfrak{S}_k}^{\mathfrak{S}_{j+k}}(V\otimes W)\right).\end{aligned}$$

Thus, Exercise 4.2.3 is solved again. Hence, we have solved Exercise 4.3.11(b).

(c) Let $\Sigma = \bigsqcup_{n\geq 0}\operatorname{Irr}(G_n)$. We have to prove that the map m is associative. In other words, we have to prove that $m(m(\alpha\otimes\beta)\otimes\gamma) = m(\alpha\otimes m(\beta\otimes\gamma))$ for any three elements α, β and γ of A. In order to do so, it is clearly enough to show that $m(m(\alpha\otimes\beta)\otimes\gamma) = m(\alpha\otimes m(\beta\otimes\gamma))$ for any three elements α, β and γ of Σ (because Σ is a $\mathbb{Z}$-module basis of A, and the equality $m(m(\alpha\otimes\beta)\otimes\gamma) = m(\alpha\otimes m(\beta\otimes\gamma))$ is $\mathbb{Z}$-linear in each of α, β and γ). So let α, β and γ be three elements of Σ. Then, there exists $i\in\mathbb{N}$, $j\in\mathbb{N}$ and $k\in\mathbb{N}$ satisfying $\alpha\in\operatorname{Irr}(G_i)$, $\beta\in\operatorname{Irr}(G_j)$ and $\gamma\in\operatorname{Irr}(G_k)$ (since α, β and γ belong to $\Sigma = \bigsqcup_{n\geq 0}\operatorname{Irr}(G_n)$). Consider these i, j and k.

There exists an irreducible $\mathbb{C}G_i$-module U satisfying $\alpha = \chi_U$ (since $\alpha\in\operatorname{Irr}(G_i)$). Similarly, there exists an irreducible $\mathbb{C}G_j$-module V satisfying $\beta = \chi_V$, and an irreducible $\mathbb{C}G_k$-module W satisfying $\gamma = \chi_W$. Consider these U, V and W.

We can apply Exercise 4.3.11(a) to $\ell = 3$, $n = i+j+k$, $\alpha = (i,j,k)$, $V_1 = U$, $V_2 = V$ and $V_3 = W$. As the result, we obtain

$$\begin{aligned}\operatorname{ind}_{i+j,k}^{i+j+k}\left(\operatorname{ind}_{(i,j)}^{i+j}(U\otimes V)\otimes W\right) &\cong \operatorname{ind}_{(i,j,k)}^{i+j+k}(U\otimes V\otimes W)\\ &\cong \operatorname{ind}_{i,j+k}^{i+j+k}\left(U\otimes\operatorname{ind}_{(j,k)}^{j+k}(V\otimes W)\right).\end{aligned}$$

Thus,

$$\operatorname{ind}_{i+j,k}^{i+j+k}\left(\operatorname{ind}_{(i,j)}^{i+j}(U\otimes V)\otimes W\right) \cong \operatorname{ind}_{i,j+k}^{i+j+k}\left(U\otimes\operatorname{ind}_{(j,k)}^{j+k}(V\otimes W)\right).$$

Since isomorphic representations have equal characters, this yields

$$\chi_{\operatorname{ind}_{i+j,k}^{i+j+k}\left(\operatorname{ind}_{(i,j)}^{i+j}(U\otimes V)\otimes W\right)} = \chi_{\operatorname{ind}_{i,j+k}^{i+j+k}\left(U\otimes\operatorname{ind}_{(j,k)}^{j+k}(V\otimes W)\right)}. \tag{13.116.12}$$

Since $\alpha = \chi_U$ and $\beta = \chi_V$, we have

$$\begin{aligned} m(\alpha\otimes\beta) &= m(\chi_U\otimes\chi_V) = \operatorname{ind}_{i,j}^{i+j}(\chi_U\otimes\chi_V) && (\text{since } U \text{ is a } \mathbb{C}G_i\text{-module and } V \text{ is a } \mathbb{C}G_j\text{-module})\\ &= \chi_{\operatorname{ind}_{i,j}^{i+j}(U\otimes V)} = \chi_{\operatorname{ind}_{(i,j)}^{i+j}(U\otimes V)} && \left(\text{since } \operatorname{ind}_{i,j}^{i+j} = \operatorname{ind}_{(i,j)}^{i+j}\right).\end{aligned}$$

Thus,

$$\begin{aligned} m\left(\underbrace{m(\alpha\otimes\beta)}_{=\chi_{\operatorname{ind}_{(i,j)}^{i+j}(U\otimes V)}} \otimes \underbrace{\gamma}_{=\chi_W}\right) &\\ = m\left(\chi_{\operatorname{ind}_{(i,j)}^{i+j}(U\otimes V)}\otimes\chi_W\right) &= \operatorname{ind}_{i+j,k}^{i+j+k}\left(\chi_{\operatorname{ind}_{(i,j)}^{i+j}(U\otimes V)}\otimes\chi_W\right)\\ &\qquad\left(\text{since } \operatorname{ind}_{(i,j)}^{i+j}(U\otimes V) \text{ is a } \mathbb{C}[G_{i+j}]\text{-module and } W \text{ is a } \mathbb{C}G_k\text{-module}\right)\\ = \chi_{\operatorname{ind}_{i+j,k}^{i+j+k}\left(\operatorname{ind}_{(i,j)}^{i+j}(U\otimes V)\otimes W\right)}. & \qquad (13.116.13)\end{aligned}$$

Similarly,

$$m(\alpha\otimes m(\beta\otimes\gamma)) = \chi_{\operatorname{ind}_{i,j+k}^{i+j+k}\left(U\otimes\operatorname{ind}_{(j,k)}^{j+k}(V\otimes W)\right)}. \tag{13.116.14}$$

Now, (13.116.13) becomes

$$\begin{aligned} m(m(\alpha\otimes\beta)\otimes\gamma) &= \chi_{\operatorname{ind}_{i+j,k}^{i+j+k}\left(\operatorname{ind}_{(i,j)}^{i+j}(U\otimes V)\otimes W\right)} = \chi_{\operatorname{ind}_{i,j+k}^{i+j+k}\left(U\otimes\operatorname{ind}_{(j,k)}^{j+k}(V\otimes W)\right)} && (\text{by } (13.116.12))\\ &= m(\alpha\otimes m(\beta\otimes\gamma)) && (\text{by } (13.116.14)).\end{aligned}$$

This is what we wanted to prove. Thus, Exercise 4.3.11(c) is solved.

(d) Let $\Sigma = \bigsqcup_{n\geq 0} \operatorname{Irr}(G_n)$.

We will solve Exercise 4.3.11(d) by induction over ℓ. The induction base (the case $\ell = 0$) is trivial, so we come to the induction step. Fix a positive integer ℓ. We assume that Exercise 4.3.11(d) is already solved for $\ell - 1$ instead of ℓ. We now need to solve Exercise 4.3.11(d) for our ℓ.

Let $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_\ell)$ be an almost-composition of an $n \in \mathbb{N}$. Let $\chi_i \in R(G_{\alpha_i})$ for every $i \in \{1, 2, \ldots, \ell\}$. We will show that

$$\chi_1\chi_2\cdots\chi_\ell = \operatorname{ind}^n_\alpha (\chi_1 \otimes \chi_2 \otimes \cdots \otimes \chi_\ell).$$

Since this equality is $\mathbb{Z}$-linear in each of χ_1, χ_2, ..., χ_ℓ, we can WLOG assume that χ_1, χ_2, ..., χ_ℓ all lie in Σ (since Σ is a basis of A). Assume this. Then, $\chi_i \in \Sigma \cap R(G_{\alpha_i}) = \operatorname{Irr}(G_{\alpha_i})$ for every $i \in \{1, 2, ..., \ell\}$. Hence, for every $i \in \{1, 2, ..., \ell\}$, there exists an irreducible $\mathbb{C}G_{\alpha_i}$-module V_i such that $\chi_i = \chi_{V_i}$. Consider this V_i.

Let $m = \alpha_1 + \alpha_2 + \cdots + \alpha_{\ell-1}$. By the induction hypothesis, we can apply Exercise 4.3.11(d) to the almost-composition $(\alpha_1, \alpha_2, ..., \alpha_{\ell-1})$ of m. As a result, we obtain

$$\begin{aligned}
\chi_1\chi_2\cdots\chi_{\ell-1} &= \operatorname{ind}^m_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})}\left(\underbrace{\chi_1}_{=\chi_{V_1}} \otimes \underbrace{\chi_2}_{=\chi_{V_2}} \otimes \cdots \otimes \underbrace{\chi_{\ell-1}}_{=\chi_{V_{\ell-1}}}\right)\\
&= \operatorname{ind}^m_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})}\underbrace{\left(\chi_{V_1} \otimes \chi_{V_2} \otimes \cdots \otimes \chi_{V_{\ell-1}}\right)}_{=\chi_{V_1\otimes V_2\otimes\cdots\otimes V_{\ell-1}}} = \operatorname{ind}^m_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})}\left(\chi_{V_1\otimes V_2\otimes\cdots\otimes V_{\ell-1}}\right)\\
&= \chi_{\operatorname{ind}^m_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})}(V_1\otimes V_2\otimes\cdots\otimes V_{\ell-1})}.
\end{aligned}$$

Now,

$$\begin{aligned}
\chi_1\chi_2\cdots\chi_\ell &= \underbrace{(\chi_1\chi_2\cdots\chi_{\ell-1})}_{=\chi_{\operatorname{ind}^m_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})}(V_1\otimes V_2\otimes\cdots\otimes V_{\ell-1})}} \underbrace{\chi_\ell}_{=\chi_{V_\ell}}\\
&= \chi_{\operatorname{ind}^m_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})}(V_1\otimes V_2\otimes\cdots\otimes V_{\ell-1})}\chi_{V_\ell} = \operatorname{ind}^n_{m,\alpha_\ell}\left(\chi_{\operatorname{ind}^m_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})}(V_1\otimes V_2\otimes\cdots\otimes V_{\ell-1})} \otimes \chi_{V_\ell}\right)\\
&\qquad \left(\begin{array}{c}\text{since } \operatorname{ind}^m_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})}(V_1\otimes V_2\otimes\cdots\otimes V_{\ell-1}) \text{ is a } \mathbb{C}G_m\text{-module, and}\\ V_\ell \text{ is a } \mathbb{C}G_{\alpha_\ell}\text{-module}\end{array}\right)\\
&= \chi_{\operatorname{ind}^n_{m,\alpha_\ell}\left(\operatorname{ind}^m_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})}(V_1\otimes V_2\otimes\cdots\otimes V_{\ell-1})\otimes V_\ell\right)}\\
&= \chi_{\operatorname{ind}^n_{\alpha_1+\alpha_2+\cdots+\alpha_{\ell-1},\alpha_\ell}\left(\operatorname{ind}^{\alpha_1+\alpha_2+\cdots+\alpha_{\ell-1}}_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})}(V_1\otimes V_2\otimes\cdots\otimes V_{\ell-1})\otimes V_\ell\right)} \qquad (\text{since } m = \alpha_1+\alpha_2+\cdots+\alpha_{\ell-1})\\
&= \chi_{\operatorname{ind}^n_\alpha(V_1\otimes V_2\otimes\cdots\otimes V_\ell)}
\end{aligned}$$

(since Exercise 4.3.11(a) yields $\operatorname{ind}^n_{\alpha_1+\alpha_2+\cdots+\alpha_{\ell-1},\alpha_\ell}\left(\operatorname{ind}^{\alpha_1+\alpha_2+\cdots+\alpha_{\ell-1}}_{(\alpha_1,\alpha_2,\ldots,\alpha_{\ell-1})}(V_1\otimes V_2\otimes\cdots\otimes V_{\ell-1})\otimes V_\ell\right) \cong \operatorname{ind}^n_\alpha(V_1\otimes V_2\otimes\cdots\otimes V_\ell)$). Compared with

$$\operatorname{ind}^n_\alpha\left(\underbrace{\chi_1}_{=\chi_{V_1}} \otimes \underbrace{\chi_2}_{=\chi_{V_2}} \otimes \cdots \otimes \underbrace{\chi_\ell}_{=\chi_{V_\ell}}\right) = \operatorname{ind}^n_\alpha(\chi_{V_1}\otimes\chi_{V_2}\otimes\cdots\otimes\chi_{V_\ell}) = \chi_{\operatorname{ind}^n_\alpha(V_1\otimes V_2\otimes\cdots\otimes V_\ell)},$$

this yields $\chi_1\chi_2\cdots\chi_\ell = \operatorname{ind}^n_\alpha(\chi_1\otimes\chi_2\otimes\cdots\otimes\chi_\ell)$. Thus, Exercise 4.3.11(d) is solved for our ℓ. This completes the induction step, and so Exercise 4.3.11(d) is solved.

(e) Let $n \in \mathbb{N}$, $\ell \in \mathbb{N}$ and $\chi \in R(G_n)$. We need to prove that $\Delta^{(\ell-1)}\chi = \sum \operatorname{res}^n_\alpha \chi$. Since this equality is $\mathbb{Z}$-linear in χ, we can WLOG assume that $\chi \in \operatorname{Irr}(G_n)$ (since $\operatorname{Irr}(G_n)$ is a $\mathbb{Z}$-module basis of $R(G_n)$). Assume this. Then, $\chi = \chi_P$ for some irreducible $\mathbb{C}G_n$-module P. Consider this P.

Similarly to how we showed (4.2.1), we can prove that

$$\operatorname{Hom}_{\mathbb{C}G_n}(\operatorname{ind}^n_\alpha U, V) \cong \operatorname{Hom}_{\mathbb{C}G_\alpha}(U, \operatorname{res}^n_\alpha V) \tag{13.116.15}$$

for every almost-composition α of n, every $\mathbb{C}G_\alpha$-module U and every $\mathbb{C}G_n$-module V. Thus,

$$(\operatorname{ind}^n_\alpha \varphi, \psi)_{R(G_n)} = (\varphi, \operatorname{res}^n_\alpha \psi)_{R(G_\alpha)} \tag{13.116.16}$$

for every almost-composition α of n, every $\varphi \in R(G_\alpha)$ and every $\psi \in R(G_n)$ [865].

The bilinear form $(\cdot, \cdot)_A$ on A induces a bilinear form $(\cdot, \cdot)_{A^{\otimes \ell}}$ on $A^{\otimes \ell}$. It is easy to see that the maps $m^{(\ell-1)}$ and $\Delta^{(\ell-1)}$ are adjoint with respect to the forms $(\cdot, \cdot)_A$ and $(\cdot, \cdot)_{A^{\otimes \ell}}$. [866] Hence, any $\varphi \in A$ and $\rho \in A^{\otimes \ell}$ satisfy

$$\left(\Delta^{(\ell-1)}\varphi, \rho\right)_{A^{\otimes \ell}} = \left(\varphi, m^{(\ell-1)}\rho\right)_A . \tag{13.116.17}$$

Since $A = \bigoplus_{n \geq 0} R(G_n)$, we have

$$\begin{aligned}
A^{\otimes \ell} &= \left(\bigoplus_{n \geq 0} R(G_n)\right)^{\otimes \ell} = \bigoplus_{n_1, n_2, \ldots, n_\ell \geq 0} \underbrace{R(G_{n_1}) \otimes R(G_{n_2}) \otimes \cdots \otimes R(G_{n_\ell})}_{= R\left(G_{n_1} \times G_{n_2} \times \cdots \times G_{n_\ell}\right)} \\
&= \bigoplus_{n_1, n_2, \ldots, n_\ell \geq 0} R(G_{n_1} \times G_{n_2} \times \cdots \times G_{n_\ell}) \\
&= \bigoplus_{\substack{\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_\ell) \text{ is an} \\ \text{almost-composition of length } \ell}} R\left(\underbrace{G_{\alpha_1} \times G_{\alpha_2} \times \cdots \times G_{\alpha_\ell}}_{= G_\alpha}\right) \\
&= \bigoplus_{\substack{\alpha \text{ is an almost-composition} \\ \text{of length } \ell}} R(G_\alpha) . \qquad (13.116.18)
\end{aligned}$$

This direct sum decomposition of $A^{\otimes \ell}$ is orthogonal with respect to the bilinear form $(\cdot, \cdot)_{A^{\otimes \ell}}$; that is, if α and β are two distinct almost-compositions of length ℓ, then

$$(R(G_\alpha), R(G_\beta))_{A^{\otimes \ell}} = 0. \tag{13.116.19}$$

[867] Moreover, for every almost-composition α of length ℓ, we have

$$(\varphi, \psi)_{A^{\otimes \ell}} = (\varphi, \psi)_{R(G_\alpha)} \qquad \text{for every } \varphi \in R(G_\alpha) \text{ and } \psi \in R(G_\alpha) . \tag{13.116.20}$$

[868]

[865] *Proof of (13.116.16):* Let α be an almost-composition of n. Let $\varphi \in R(G_\alpha)$ and $\psi \in R(G_n)$. We need to prove the equality (13.116.16). Since this equality is $\mathbb{Z}$-linear in each of φ and ψ, we can WLOG assume that $\varphi \in \operatorname{Irr}(G_\alpha)$ and $\psi \in \operatorname{Irr}(G_n)$ (since $\operatorname{Irr}(G_\alpha)$ and $\operatorname{Irr}(G_n)$ are $\mathbb{Z}$-module bases of $R(G_\alpha)$ and $R(G_n)$, respectively). Assume this. Thus, there exist an irreducible $\mathbb{C}G_\alpha$-module U and an irreducible $\mathbb{C}G_n$-module V such that $\varphi = \chi_U$ and $\psi = \chi_V$. Consider these U and V. We have

$$\begin{aligned}
\left(\operatorname{ind}_\alpha^n \underbrace{\varphi}_{= \chi_U}, \underbrace{\psi}_{= \chi_V}\right)_{R(G_n)} &= \left(\underbrace{\operatorname{ind}_\alpha^n \chi_U}_{= \chi_{\operatorname{ind}_\alpha^n U}}, \chi_V\right)_{R(G_n)} = \left(\chi_{\operatorname{ind}_\alpha^n U}, \chi_V\right)_{R(G_n)} \\
&= \dim_{\mathbb{C}} \operatorname{Hom}_{\mathbb{C}G_n}(\operatorname{ind}_\alpha^n U, V) = \dim_{\mathbb{C}} \operatorname{Hom}_{\mathbb{C}G_\alpha}(U, \operatorname{res}_\alpha^n V) \qquad (\text{by } (13.116.15)) .
\end{aligned}$$

Compared with

$$\left(\underbrace{\varphi}_{= \chi_U}, \operatorname{res}_\alpha^n \underbrace{\psi}_{= \chi_V}\right)_{R(G_\alpha)} = \left(\chi_U, \underbrace{\operatorname{res}_\alpha^n \chi_V}_{= \chi_{\operatorname{res}_\alpha^n V}}\right)_{R(G_\alpha)} = \left(\chi_U, \chi_{\operatorname{res}_\alpha^n V}\right)_{R(G_\alpha)} = \dim_{\mathbb{C}} \operatorname{Hom}_{\mathbb{C}G_\alpha}(U, \operatorname{res}_\alpha^n V) ,$$

this yields $(\operatorname{ind}_\alpha^n \varphi, \psi)_{R(G_n)} = (\varphi, \operatorname{res}_\alpha^n \psi)_{R(G_\alpha)}$. This proves (13.116.16).

[866] Indeed, this follows from the inductive definitions of $m^{(\ell-1)}$ and $\Delta^{(\ell-1)}$, if one recalls that the maps m and Δ are adjoint with respect to the forms $(\cdot, \cdot)_A$ and $(\cdot, \cdot)_{A \otimes A}$.

[867] This follows from the definition of $(\cdot, \cdot)_{A^{\otimes \ell}}$ and the fact that the direct sum decomposition $A = \bigoplus_{n \geq 0} R(G_n)$ is orthogonal with respect to the bilinear form $(\cdot, \cdot)_A$.

[868] *Proof of (13.116.20):* Let α be an almost-composition of length ℓ. Let $\varphi \in R(G_\alpha)$ and $\psi \in R(G_\alpha)$. We need to prove the equality (13.116.20). Since this equality is $\mathbb{Z}$-linear in each of φ and ψ, we can WLOG assume that $\varphi \in \operatorname{Irr}(G_\alpha)$ and $\psi \in \operatorname{Irr}(G_\alpha)$ (since $\operatorname{Irr}(G_\alpha)$ is a $\mathbb{Z}$-module basis of $R(G_\alpha)$). Assume this. Thus, there exist an irreducible $\mathbb{C}G_\alpha$-module V and an irreducible $\mathbb{C}G_\alpha$-module W such that $\varphi = \chi_V$ and $\psi = \chi_W$. Consider these V and W.

Write the almost-composition α in the form $(\alpha_1, \alpha_2, \ldots, \alpha_\ell)$. Since V is an irreducible representation of $G_\alpha = G_{\alpha_1} \times G_{\alpha_2} \times \cdots \times G_{\alpha_\ell}$, we can write V in the form $V = V_1 \otimes V_2 \otimes \cdots \otimes V_\ell$, where each V_i is an irreducible representation of G_{α_i}. Similarly, We need to show that $\Delta^{(\ell-1)}\chi = \sum \operatorname{res}^n_\alpha \chi$ (where the sum ranges over all almost-compositions α of n having length ℓ). In order to do so, it is clearly enough to prove that

$$\left(\Delta^{(\ell-1)}\chi, \rho\right)_{A^{\otimes \ell}} = \left(\sum \operatorname{res}^n_\alpha \chi, \rho\right)_{A^{\otimes \ell}} \qquad \text{for every } \rho \in A^{\otimes \ell} \tag{13.116.21}$$

(since the bilinear form $(\cdot,\cdot)_{A^{\otimes \ell}}$ is nondegenerate). So, let $\rho \in A^{\otimes \ell}$. It remains to prove (13.116.21).

The equality (13.116.21) is $\mathbb{Z}$-linear in ρ. Since $\bigsqcup_{\substack{\alpha \text{ is an almost-composition} \\ \text{of length } \ell}} \operatorname{Irr}(G_\alpha)$ is a $\mathbb{Z}$-module basis of $A^{\otimes \ell}$ (this follows from (13.116.18) and the fact that each $R(G_\alpha)$ has $\mathbb{Z}$-module basis $\operatorname{Irr}(G_\alpha)$), we can therefore WLOG assume that $\rho \in \bigsqcup_{\substack{\alpha \text{ is an almost-composition} \\ \text{of length } \ell}} \operatorname{Irr}(G_\alpha)$. Assume this. Then, there exists an almost-composition β of length ℓ such that $\rho \in \operatorname{Irr}(G_\beta)$. Consider this β, and notice that $\rho \in \operatorname{Irr}(G_\beta) \subset R(G_\beta)$.

For every almost-composition α of n having length ℓ satisfying $\alpha \neq \beta$, we have

$$\left(\operatorname{res}^n_\alpha \chi, \rho\right)_{A^{\otimes \ell}} = 0 \tag{13.116.22}$$

we can write W in the form $W = W_1 \otimes W_2 \otimes \cdots \otimes W_\ell$, where each W_i is an irreducible representation of G_{α_i}. Consider these V_i and W_i.

Now, there exists a $\mathbb{C}$-vector space isomorphism

$$\begin{aligned} &\operatorname{Hom}_{\mathbb{C}G_{\alpha_1}}(V_1, W_1) \otimes \operatorname{Hom}_{\mathbb{C}G_{\alpha_2}}(V_2, W_2) \otimes \cdots \otimes \operatorname{Hom}_{\mathbb{C}G_{\alpha_\ell}}(V_\ell, W_\ell) \\ &\to \operatorname{Hom}_{\mathbb{C}[G_{\alpha_1} \times G_{\alpha_2} \times \cdots \times G_{\alpha_\ell}]}(V_1 \otimes V_2 \otimes \cdots \otimes V_\ell, W_1 \otimes W_2 \otimes \cdots \otimes W_\ell). \end{aligned}$$

(In fact, when $\ell = 2$, the existence of such an isomorphism follows from Exercise 4.1.9(a); otherwise it follows by induction over ℓ using Exercise 4.1.9(a).) The existence of this isomorphism yields

$$\begin{aligned} &\dim_{\mathbb{C}}\left(\operatorname{Hom}_{\mathbb{C}[G_{\alpha_1} \times G_{\alpha_2} \times \cdots \times G_{\alpha_\ell}]}(V_1 \otimes V_2 \otimes \cdots \otimes V_\ell, W_1 \otimes W_2 \otimes \cdots \otimes W_\ell)\right) \\ &= \dim_{\mathbb{C}}\left(\operatorname{Hom}_{\mathbb{C}G_{\alpha_1}}(V_1, W_1) \otimes \operatorname{Hom}_{\mathbb{C}G_{\alpha_2}}(V_2, W_2) \otimes \cdots \otimes \operatorname{Hom}_{\mathbb{C}G_{\alpha_\ell}}(V_\ell, W_\ell)\right) \\ &= \prod_{i=1}^{\ell} \underbrace{\dim_{\mathbb{C}} \operatorname{Hom}_{\mathbb{C}G_{\alpha_i}}(V_i, W_i)}_{=\left(\chi_{V_i}, \chi_{W_i}\right)_{R\left(G_{\alpha_i}\right)} = \left(\chi_{V_i}, \chi_{W_i}\right)_A} = \prod_{i=1}^{\ell} \left(\chi_{V_i}, \chi_{W_i}\right)_A \\ &= \Bigg(\underbrace{\chi_{V_1} \otimes \chi_{V_2} \otimes \cdots \otimes \chi_{V_\ell}}_{\substack{=\chi_{V_1 \otimes V_2 \otimes \cdots \otimes V_\ell} = \chi_V \\ (\text{since } V_1 \otimes V_2 \otimes \cdots \otimes V_\ell = V)}}, \underbrace{\chi_{W_1} \otimes \chi_{W_2} \otimes \cdots \otimes \chi_{W_\ell}}_{\substack{=\chi_{W_1 \otimes W_2 \otimes \cdots \otimes W_\ell} = \chi_W \\ (\text{since } W_1 \otimes W_2 \otimes \cdots \otimes W_\ell = W)}}\Bigg)_{A^{\otimes \ell}} \\ &= \Bigg(\underbrace{\chi_V}_{=\varphi}, \underbrace{\chi_W}_{=\psi}\Bigg)_{A^{\otimes \ell}} = (\varphi, \psi)_{A^{\otimes \ell}}. \end{aligned}$$

Thus,

$$\begin{aligned} &(\varphi, \psi)_{A^{\otimes \ell}} \\ &= \dim_{\mathbb{C}}\left(\operatorname{Hom}_{\mathbb{C}[G_{\alpha_1} \times G_{\alpha_2} \times \cdots \times G_{\alpha_\ell}]}(V_1 \otimes V_2 \otimes \cdots \otimes V_\ell, W_1 \otimes W_2 \otimes \cdots \otimes W_\ell)\right) \\ &= \dim_{\mathbb{C}}\left(\operatorname{Hom}_{\mathbb{C}G_\alpha}(V_1 \otimes V_2 \otimes \cdots \otimes V_\ell, W_1 \otimes W_2 \otimes \cdots \otimes W_\ell)\right) \qquad (\text{since } G_{\alpha_1} \times G_{\alpha_2} \times \cdots \times G_{\alpha_\ell} = G_\alpha) \\ &= \Bigg(\underbrace{\chi_{V_1 \otimes V_2 \otimes \cdots \otimes V_\ell}}_{\substack{=\chi_V \\ (\text{since } V_1 \otimes V_2 \otimes \cdots \otimes V_\ell = V)}}, \underbrace{\chi_{W_1 \otimes W_2 \otimes \cdots \otimes W_\ell}}_{\substack{=\chi_W \\ (\text{since } W_1 \otimes W_2 \otimes \cdots \otimes W_\ell = W)}}\Bigg)_{R(G_\alpha)} = \Bigg(\underbrace{\chi_V}_{=\varphi}, \underbrace{\chi_W}_{=\psi}\Bigg)_{R(G_\alpha)} = (\varphi, \psi)_{R(G_\alpha)}. \end{aligned}$$

This proves (13.116.20).

[869]. Now,

$$\begin{aligned}
\left(\sum \operatorname{res}^n_\alpha \chi, \rho\right)_{A^{\otimes \ell}} &= \sum \left(\operatorname{res}^n_\alpha \chi, \rho\right)_{A^{\otimes \ell}} = \left(\operatorname{res}^n_\beta \chi, \rho\right)_{A^{\otimes \ell}} + \sum_{\alpha \neq \beta} \underbrace{\left(\operatorname{res}^n_\alpha \chi, \rho\right)_{A^{\otimes \ell}}}_{\substack{=0 \\ \text{(by (13.116.22))}}} = \left(\operatorname{res}^n_\beta \chi, \rho\right)_{A^{\otimes \ell}} \\
&= \left(\operatorname{res}^n_\beta \chi, \rho\right)_{R(G_\beta)} \qquad \left(\text{by (13.116.20), applied to } \varphi = \operatorname{res}^n_\beta \chi,\ \psi = \rho \text{ and } \alpha = \beta\right) \\
&= \left(\rho, \operatorname{res}^n_\beta \chi\right)_{R(G_\beta)} \\
&= \left(\operatorname{ind}^n_\beta \rho, \chi\right)_{R(G_n)} \qquad \left(\begin{array}{c} \text{since (13.116.16) (applied } \varphi = \rho,\ \psi = \chi \text{ and } \alpha = \beta) \\ \text{yields } \left(\operatorname{ind}^n_\beta \rho, \chi\right)_{R(G_n)} = \left(\rho, \operatorname{res}^n_\beta \chi\right)_{R(G_\beta)} \end{array}\right). \qquad (13.116.23)
\end{aligned}$$

But let us write the almost-composition β as $(\beta_1, \beta_2, ..., \beta_\ell)$. Then, $\rho \in \operatorname{Irr}(G_\beta)$ is an irreducible character of $G_\beta = G_{\beta_1} \times G_{\beta_2} \times \cdots \times G_{\beta_\ell}$, and thus has the form $\rho = \rho_1 \otimes \rho_2 \otimes \cdots \otimes \rho_\ell$, where each ρ_i is an irreducible character of G_{β_i}. Consider these ρ_i. We have

$$\begin{aligned}
m^{(\ell-1)} \underbrace{\rho}_{=\rho_1 \otimes \rho_2 \otimes \cdots \otimes \rho_\ell} &= m^{(\ell-1)} \left(\rho_1 \otimes \rho_2 \otimes \cdots \otimes \rho_\ell\right) = \rho_1 \rho_2 \cdots \rho_\ell = \operatorname{ind}^n_\beta \left(\underbrace{\rho_1 \otimes \rho_2 \otimes \cdots \otimes \rho_\ell}_{=\rho}\right) \\
&\qquad \left(\text{by Exercise 4.3.11(d), applied to } \beta \text{ and } \rho_i \text{ instead of } \alpha \text{ and } \chi_i\right) \\
&= \operatorname{ind}^n_\beta \rho.
\end{aligned}$$

Now,

$$\begin{aligned}
\left(\Delta^{(\ell-1)} \chi, \rho\right)_{A^{\otimes \ell}} &= \left(\chi, \underbrace{m^{(\ell-1)} \rho}_{=\operatorname{ind}^n_\beta \rho}\right)_A \qquad \left(\text{by (13.116.17), applied to } \varphi = \chi\right) \\
&= \left(\chi, \operatorname{ind}^n_\beta \rho\right)_A = \left(\chi, \operatorname{ind}^n_\beta \rho\right)_{R(G_n)} = \left(\operatorname{ind}^n_\beta \rho, \chi\right)_{R(G_n)} \\
&= \left(\sum \operatorname{res}^n_\alpha \chi, \rho\right)_{A^{\otimes \ell}} \qquad \left(\text{by (13.116.23)}\right).
\end{aligned}$$

Thus, (13.116.21) is proven. This completes our solution of Exercise 4.3.11(e).

13.117. **Solution to Exercise 4.4.3.** *Solution to Exercise 4.4.3.* Define the Iverson bracket notation as in Exercise 4.4.3(a).

(a) Let G be a finite group. For every subset P of G, the definition of $\underline{1}_P$ yields

$$\underline{1}_P (g) = [g \in P] \qquad \text{for every } g \in G. \tag{13.117.1}$$

Also, for every subset P of G, we have

$$|P| = \sum_{k \in G} [k \in P]. \tag{13.117.2}$$

Let $h \in G$ and $g \in G$. Then, (13.117.2) (applied to $P = Z_G(h)$) yields

$$|Z_G(h)| = \sum_{k \in G} \left[\underbrace{k \in Z_G(h)}_{\substack{\text{this is equivalent to } khk^{-1} = h \\ \text{(by the definition of } Z_G(h))}} \right] = \sum_{k \in G} \left[khk^{-1} = h\right].$$

[869] since

$$\left(\underbrace{\operatorname{res}^n_\alpha \chi}_{\in R(G_\alpha)}, \underbrace{\rho}_{\in \operatorname{Irr}(G_\beta) \subset R(G_\beta)}\right)_{A^{\otimes \ell}} \in \left(R(G_\alpha), R(G_\beta)\right)_{A^{\otimes \ell}} = 0 \qquad \left(\text{by (13.116.19)}\right)$$

Applying both sides of the identity $\alpha_{G,h} = |Z_G(h)| \underline{1}_{\operatorname{Conj}_G(h)}$ to g, we obtain

$$\alpha_{G,h}(g) = \underbrace{|Z_G(h)|}_{=\sum_{k\in G}[khk^{-1}=h]} \underbrace{\underline{1}_{\operatorname{Conj}_G(h)}(g)}_{\substack{=[g\in\operatorname{Conj}_G(h)] \\ \text{(by (13.117.1), applied} \\ \text{to } P=\operatorname{Conj}_G(h))}} = \left(\sum_{k\in G}\left[khk^{-1}=h\right]\right)\left[g\in\operatorname{Conj}_G(h)\right]. \tag{13.117.3}$$

We need to prove that $\alpha_{G,h}(g) = \sum_{k\in G}\left[khk^{-1}=g\right]$. We must be in one of the following two cases:

Case 1: We have $g \in \operatorname{Conj}_G(h)$.

Case 2: We don't have $g \in \operatorname{Conj}_G(h)$.

Let us first consider Case 1. In this case, we have $g \in \operatorname{Conj}_G(h)$. Thus, there exists a $p \in G$ satisfying $g = php^{-1}$. Consider this p. Since $g = php^{-1}$, we have $h = p^{-1}gp$. Hence, (13.117.3) becomes

$$\begin{aligned}
\alpha_{G,h}(g) &= \left(\sum_{k\in G}\left[khk^{-1}=h\right]\right)\underbrace{\left[g\in\operatorname{Conj}_G(h)\right]}_{\substack{=1 \\ \text{(since } g\in\operatorname{Conj}_G(h))}} = \sum_{k\in G}\left[k\underbrace{h}_{=p^{-1}gp}k^{-1}=h\right] \\
&= \sum_{k\in G}\left[\underbrace{kp^{-1}}_{=(pk^{-1})^{-1}}gpk^{-1}=h\right] = \sum_{k\in G}\left[\left(pk^{-1}\right)^{-1}g\left(kp^{-1}\right)^{-1}=h\right] \\
&= \sum_{k\in G}\left[\underbrace{k^{-1}gk=h}_{\substack{\text{this is equivalent to} \\ khk^{-1}=g}}\right] \qquad \left(\begin{array}{c}\text{here, we have substituted } k \text{ for } pk^{-1} \text{ in the sum} \\ \text{(since the map } G\to G,\ k\mapsto pk^{-1} \text{ is a bijection)}\end{array}\right) \\
&= \sum_{k\in G}\left[khk^{-1}=g\right].
\end{aligned}$$

Thus, $\alpha_{G,h}(g) = \sum_{k\in G}\left[khk^{-1}=g\right]$ is proven in Case 1.

Let us now consider Case 2. In this case, we don't have $g \in \operatorname{Conj}_G(h)$. In other words, g is not in the conjugacy class of h. In other words, g is not conjugate to h. In other words, there exists no $k \in G$ satisfying $khk^{-1} = g$. In other words, for every $k \in G$, we do not have $khk^{-1} = g$. Hence, for every $k \in G$, we have $\left[khk^{-1}=g\right] = 0$. Thus, $\sum_{k\in G}\underbrace{\left[khk^{-1}=g\right]}_{=0} = \sum_{k\in G} 0 = 0$. Comparing this with

$$\alpha_{G,h}(g) = \left(\sum_{k\in G}\left[khk^{-1}=h\right]\right)\underbrace{\left[g\in\operatorname{Conj}_G(h)\right]}_{\substack{=0 \\ \text{(since we don't have } g\in\operatorname{Conj}_G(h))}} = 0,$$

we obtain $\alpha_{G,h}(g) = \sum_{k\in G}\left[khk^{-1}=g\right]$. Thus, $\alpha_{G,h}(g) = \sum_{k\in G}\left[khk^{-1}=g\right]$ is proven in Case 2.

Thus, $\alpha_{G,h}(g) = \sum_{k\in G}\left[khk^{-1}=g\right]$ is proven in both Cases 1 and 2. Hence, Exercise 4.4.3(a) is solved.

(b) Let H be a subgroup of a finite group G. Let $h \in H$. Let $g \in G$. Then, the definition of $\operatorname{Ind}_H^G \alpha_{H,h}$ given in Exercise 4.1.1 yields

$$\left(\operatorname{Ind}_H^G \alpha_{H,h}\right)(g) = \frac{1}{|H|} \sum_{\substack{k\in G:\\ kgk^{-1}\in H}} \alpha_{H,h}\left(kgk^{-1}\right) = \frac{1}{|H|} \sum_{\substack{p\in G:\\ pgp^{-1}\in H}} \underbrace{\alpha_{H,h}\left(pgp^{-1}\right)}_{\substack{=\sum_{k\in H}\left[khk^{-1}=pgp^{-1}\right]\\ \text{(by Exercise 4.4.3(a), applied}\\ \text{to } H \text{ and } pgp^{-1} \text{ instead of } G \text{ and } g)}}$$

(here, we renamed the summation index k as p)

$$= \frac{1}{|H|} \underbrace{\sum_{\substack{p\in G:\\ pgp^{-1}\in H}} \sum_{k\in H}}_{=\sum_{k\in H} \sum_{\substack{p\in G:\\ pgp^{-1}\in H}}} \left[khk^{-1} = pgp^{-1}\right] = \frac{1}{|H|} \sum_{k\in H} \sum_{\substack{p\in G:\\ pgp^{-1}\in H}} \left[khk^{-1} = pgp^{-1}\right]. \tag{13.117.4}$$

But every $k \in H$ satisfies

$$\sum_{\substack{p\in G:\\ pgp^{-1}\in H}} \left[khk^{-1} = pgp^{-1}\right] = \sum_{p\in G} \left[khk^{-1} = pgp^{-1}\right] \tag{13.117.5}$$

[870]. Now, (13.117.4) becomes

$$\left(\operatorname{Ind}_H^G \alpha_{H,h}\right)(g) = \frac{1}{|H|} \sum_{k\in H} \underbrace{\sum_{\substack{p\in G:\\ pgp^{-1}\in H}} \left[khk^{-1} = pgp^{-1}\right]}_{\substack{=\sum_{p\in G}\left[khk^{-1}=pgp^{-1}\right]\\ \text{(by (13.117.5))}}} = \frac{1}{|H|} \sum_{k\in H} \sum_{p\in G} \left[\underbrace{khk^{-1} = pgp^{-1}}_{\substack{\text{this is equivalent to}\\ p^{-1}khk^{-1}p=g}}\right]$$

$$= \frac{1}{|H|} \sum_{k\in H} \sum_{p\in G} \left[p^{-1}kh \underbrace{k^{-1}p}_{=\left(p^{-1}k\right)^{-1}} = g\right] = \frac{1}{|H|} \sum_{k\in H} \sum_{p\in G} \left[p^{-1}kh\left(p^{-1}k\right)^{-1} = g\right]$$

$$= \frac{1}{|H|} \underbrace{\sum_{k\in H} \sum_{p\in G} \left[php^{-1} = g\right]}_{=|H| \sum_{p\in G}\left[php^{-1}=g\right]} \qquad \left(\begin{array}{c}\text{here, we substituted } p \text{ for } p^{-1}k \text{ in the inner sum}\\ \text{(since the map } G \to G,\ p \mapsto p^{-1}k \text{ is a bijection)}\end{array}\right)$$

$$= \frac{1}{|H|} |H| \sum_{p\in G} \left[php^{-1} = g\right] = \sum_{p\in G} \left[php^{-1} = g\right] = \sum_{k\in G} \left[khk^{-1} = g\right]$$

(here, we have substituted k for p in the sum)

$$= \alpha_{G,h}(g) \qquad \text{(by Exercise 4.4.3(a))}.$$

Let us now forget that we fixed g. We thus have proven that $\left(\operatorname{Ind}_H^G \alpha_{H,h}\right)(g) = \alpha_{G,h}(g)$ for every $g \in G$. Hence, $\operatorname{Ind}_H^G \alpha_{H,h} = \alpha_{G,h}$. This solves Exercise 4.4.3(b).

[870] *Proof of (13.117.5):* Let $k \in H$. Then,

$$\sum_{p\in G} \left[khk^{-1} = pgp^{-1}\right] = \sum_{\substack{p\in G:\\ pgp^{-1}\in H}} \left[khk^{-1} = pgp^{-1}\right] + \sum_{\substack{p\in G:\\ pgp^{-1}\notin H}} \underbrace{\left[khk^{-1} = pgp^{-1}\right]}_{\substack{=0\\ \text{(since we don't have } khk^{-1}=pgp^{-1}\\ \text{(since } khk^{-1}\in H \text{ (since } k\in H \text{ and } h\in H) \text{ and } pgp^{-1}\notin H))}}$$

$$= \sum_{\substack{p\in G:\\ pgp^{-1}\in H}} \left[khk^{-1} = pgp^{-1}\right] + \underbrace{\sum_{\substack{p\in G:\\ pgp^{-1}\notin H}} 0}_{=0} = \sum_{\substack{p\in G:\\ pgp^{-1}\in H}} \left[khk^{-1} = pgp^{-1}\right].$$

This proves (13.117.5).

(c) Let G_1 and G_2 be two finite groups. Let $h_1 \in G_1$ and $h_2 \in G_2$. Let $\mathbf{I}$ denote the canonical isomorphism $R_{\mathbb{C}}(G_1) \otimes R_{\mathbb{C}}(G_2) \to R_{\mathbb{C}}(G_1 \times G_2)$. We need to prove that

$$\mathbf{I}\left(\alpha_{G_1,h_1} \otimes \alpha_{G_2,h_2}\right) = \alpha_{G_1 \times G_2,(h_1,h_2)}. \tag{13.117.6}$$

Let $g \in G_1 \times G_2$. Let us write g in the form $g = (g_1, g_2)$. Then, Exercise 4.4.3(a) (applied to G_1, h_1 and g_1 instead of G, h and g) yields

$$\alpha_{G_1,h_1}(g_1) = \sum_{k \in G_1} \left[k h_1 k^{-1} = g_1\right] = \sum_{k_1 \in G_1} \left[k_1 h_1 k_1^{-1} = g_1\right]$$

(here, we have renamed the summation index k as k_1). Similarly,

$$\alpha_{G_2,h_2}(g_2) = \sum_{k_2 \in G_2} \left[k_2 h_2 k_2^{-2} = g_2\right].$$

Now,

$$\begin{aligned}
&\left(\mathbf{I}\left(\alpha_{G_1,h_1} \otimes \alpha_{G_2,h_2}\right)\right)\left(\underbrace{g}_{=(g_1,g_2)}\right) \\
&= \left(\mathbf{I}\left(\alpha_{G_1,h_1} \otimes \alpha_{G_2,h_2}\right)\right)\left((g_1,g_2)\right) = \underbrace{\alpha_{G_1,h_1}(g_1)}_{=\sum_{k_1 \in G_1}\left[k_1 h_1 k_1^{-1} = g_1\right]} \underbrace{\alpha_{G_2,h_2}(g_2)}_{=\sum_{k_2 \in G_2}\left[k_2 h_2 k_2^{-2} = g_2\right]} \qquad \text{(by the definition of } \mathbf{I}\text{)} \\
&= \left(\sum_{k_1 \in G_1} \left[k_1 h_1 k_1^{-1} = g_1\right]\right)\left(\sum_{k_2 \in G_2} \left[k_2 h_2 k_2^{-2} = g_2\right]\right) = \sum_{(k_1,k_2) \in G_1 \times G_2} \underbrace{\left[k_1 h_1 k_1^{-1} = g_1\right]\left[k_2 h_2 k_2^{-2} = g_2\right]}_{=\left[k_1 h_1 k_1^{-1} = g_1 \text{ and } k_2 h_2 k_2^{-2} = g_2\right]} \\
&= \sum_{(k_1,k_2) \in G_1 \times G_2} \left[\underbrace{k_1 h_1 k_1^{-1} = g_1 \text{ and } k_2 h_2 k_2^{-2} = g_2}_{\substack{\text{this is equivalent to} \\ \left(k_1 h_1 k_1^{-1}, k_2 h_2 k_2^{-1}\right) = (g_1,g_2)}}\right] = \sum_{(k_1,k_2) \in G_1 \times G_2} \left[\underbrace{\left(k_1 h_1 k_1^{-1}, k_2 h_2 k_2^{-1}\right)}_{=(k_1,k_2)(h_1,h_2)(k_1,k_2)^{-1}} = \underbrace{(g_1,g_2)}_{=g}\right] \\
&= \sum_{(k_1,k_2) \in G_1 \times G_2} \left[(k_1,k_2)(h_1,h_2)(k_1,k_2)^{-1} = g\right] = \sum_{k \in G_1 \times G_2} \left[k (h_1,h_2) k^{-1} = g\right]
\end{aligned}$$

(here, we have renamed the summation index (k_1, k_2) as k). Compared with

$$\alpha_{G_1 \times G_2,(h_1,h_2)}(g) = \sum_{k \in G_1 \times G_2} \left[k (h_1,h_2) k^{-1} = g\right]$$
$$\left(\text{by Exercise 4.4.3(a), applied to } G_1 \times G_2 \text{ and } (h_1,h_2) \text{ instead of } G \text{ and } h\right),$$

this yields $\left(\mathbf{I}\left(\alpha_{G_1,h_1} \otimes \alpha_{G_2,h_2}\right)\right)(g) = \alpha_{G_1 \times G_2,(h_1,h_2)}(g)$.

Now, let us forget that we fixed g. We thus have proven that $\left(\mathbf{I}\left(\alpha_{G_1,h_1} \otimes \alpha_{G_2,h_2}\right)\right)(g) = \alpha_{G_1 \times G_2,(h_1,h_2)}(g)$ for every $g \in G_1 \times G_2$. In other words, $\mathbf{I}\left(\alpha_{G_1,h_1} \otimes \alpha_{G_2,h_2}\right) = \alpha_{G_1 \times G_2,(h_1,h_2)}$. This proves (13.117.6). Exercise 4.4.3(c) is thus solved.

(d) For every partition λ, let us define $\widetilde{z}_\lambda$ as the size of the centralizer of a permutation in $\mathfrak{S}_{|\lambda|}$ having cycle type λ. Note that this does not depend on the choice of said permutation, since all permutations in $\mathfrak{S}_{|\lambda|}$ having cycle type λ are mutually conjugate (and thus their centralizers are of the same size). It is well-known that $\widetilde{z}_\lambda = z_\lambda$ (see Remark 2.5.16), but we shall avoid using this fact, as we can obtain an alternative proof of it from the following argument.

We shall use the same notations as in the proof of Theorem 4.4.1. In particular, $A_{\mathbb{C}} = \bigoplus_{n \geq 0} R_{\mathbb{C}}(\mathfrak{S}_n)$; this $\mathbb{C}$-vector space $A_{\mathbb{C}}$ becomes a $\mathbb{C}$-algebra as explained in the proof of Theorem 4.4.1. Its multiplication is given by $\operatorname{ind}_{i,j}^{i+j}$; more precisely, every $n \in \mathbb{N}$ and $m \in \mathbb{N}$, and every $\beta \in R_{\mathbb{C}}(\mathfrak{S}_n)$ and $\gamma \in R_{\mathbb{C}}(\mathfrak{S}_m)$ satisfy

$$\beta\gamma = \underbrace{\operatorname{ind}_{n,m}^{n+m}}_{=\operatorname{Ind}_{\mathfrak{S}_n \times \mathfrak{S}_m}^{\mathfrak{S}_{n+m}}} (\beta \otimes \gamma) = \operatorname{Ind}_{\mathfrak{S}_n \times \mathfrak{S}_m}^{\mathfrak{S}_{n+m}} (\beta \otimes \gamma). \tag{13.117.7}$$

We define a $\mathbb{C}$-linear map $\Phi : \Lambda_{\mathbb{C}} \to A_{\mathbb{C}}$ by setting

$$\Phi\left(p_\lambda\right) = \widetilde{z}_\lambda \underline{1}_\lambda \qquad \text{for every } \lambda \in \operatorname{Par}.$$

(This is well-defined, since $\left(p_\lambda\right)_{\lambda \in \operatorname{Par}}$ is a basis of the $\mathbb{C}$-vector space $\Lambda_{\mathbb{C}}$.) We notice that if λ is a partition of a nonnegative integer n, and if $g \in \mathfrak{S}_n$ is a permutation having cycle type λ, then

$$\Phi\left(p_\lambda\right) = \alpha_{\mathfrak{S}_n, g}. \tag{13.117.8}$$

[871] Also, if λ is a partition of a nonnegative integer n, then

$$\left|\left\{h \in \mathfrak{S}_n \mid h \text{ has cycle type } \lambda\right\}\right| = n!/\widetilde{z}_\lambda. \tag{13.117.9}$$

[872]

We shall now prove that Φ is a $\mathbb{C}$-algebra homomorphism. Since $\Phi(1) = 1$ is true[873], we only need to verify that $\Phi(uv) = \Phi(u)\Phi(v)$ for any $u \in \Lambda_{\mathbb{C}}$ and $v \in \Lambda_{\mathbb{C}}$. Let us prove this now. Fix $u \in \Lambda_{\mathbb{C}}$ and $v \in \Lambda_{\mathbb{C}}$. Since the equality $\Phi(uv) = \Phi(u)\Phi(v)$ is $\mathbb{C}$-linear in each of u and v, we can WLOG assume that u and v are elements of the basis $\left(p_\lambda\right)_{\lambda \in \operatorname{Par}}$ of the $\mathbb{C}$-vector space $\Lambda_{\mathbb{C}}$. Assume this, and set $u = p_\mu$ and $v = p_\nu$ for two partitions μ and ν. Let $n = |\mu|$ and $m = |\nu|$. Let g be a permutation in $\mathfrak{S}_n$ having cycle type μ. (Such

[871] *Proof of (13.117.8):* Let λ be a partition of a nonnegative integer n. Let $g \in \mathfrak{S}_n$ be a permutation having cycle type λ. Then, g is a permutation in $\mathfrak{S}_{|\lambda|}$ having cycle type λ. Thus, $\widetilde{z}_\lambda$ is the size of the centralizer of g (by the definition of $\widetilde{z}_\lambda$). In other words, $\widetilde{z}_\lambda = \left|Z_{\mathfrak{S}_n}(g)\right|$.

But any two permutations in $\mathfrak{S}_n$ having the same cycle type are mutually conjugate. Hence, if h is any permutation in $\mathfrak{S}_n$ having cycle type λ, then h and g are conjugate (since h and g are permutations in $\mathfrak{S}_n$ having the same cycle type). Conversely, if h is a permutation in $\mathfrak{S}_n$ such that h and g are conjugate, then h has cycle type λ (because h and g are conjugate permutations and thus have the same cycle type, but the cycle type of g is λ). Combining these two statements, we conclude that if h is a permutation in $\mathfrak{S}_n$, then h has cycle type λ if and only if h and g are conjugate. Hence,

$$\left\{h \in \mathfrak{S}_n \mid \underbrace{h \text{ has cycle type } \lambda}_{\substack{\text{this is equivalent to} \\ (h \text{ and } g \text{ are conjugate})}}\right\} = \left\{h \in \mathfrak{S}_n \mid \underbrace{h \text{ and } g \text{ are conjugate}}_{\text{this is equivalent to } h \in \operatorname{Conj}_{\mathfrak{S}_n}(g)}\right\} = \left\{h \in \mathfrak{S}_n \mid h \in \operatorname{Conj}_{\mathfrak{S}_n}(g)\right\}$$
$$= \operatorname{Conj}_{\mathfrak{S}_n}(g).$$

On the other hand, $\underline{1}_\lambda$ is defined as the characteristic function for the $\mathfrak{S}_n$-conjugacy class of permutations of cycle type λ. In other words, $\underline{1}_\lambda$ is the indicator function of the subset $\left\{h \in \mathfrak{S}_n \mid h \text{ has cycle type } \lambda\right\}$ of $\mathfrak{S}_n$. In other words, $\underline{1}_\lambda$ is the indicator function of the subset $\operatorname{Conj}_{\mathfrak{S}_n}(g)$ (since $\left\{h \in \mathfrak{S}_n \mid h \text{ has cycle type } \lambda\right\} = \operatorname{Conj}_{\mathfrak{S}_n}(g)$). In other words, $\underline{1}_\lambda = \underline{1}_{\operatorname{Conj}_{\mathfrak{S}_n}(g)}$.

Now, $\Phi\left(p_\lambda\right) = \underbrace{\widetilde{z}_\lambda}_{=\left|Z_{\mathfrak{S}_n}(g)\right|} \underbrace{\underline{1}_\lambda}_{=\underline{1}_{\operatorname{Conj}_{\mathfrak{S}_n}(g)}} = \left|Z_{\mathfrak{S}_n}(g)\right| \underline{1}_{\operatorname{Conj}_{\mathfrak{S}_n}(g)}$. Compared with $\alpha_{\mathfrak{S}_n, g} = \left|Z_{\mathfrak{S}_n}(g)\right| \underline{1}_{\operatorname{Conj}_{\mathfrak{S}_n}(g)}$ (by the definition of $\alpha_{\mathfrak{S}_n, g}$), this yields $\Phi\left(p_\lambda\right) = \alpha_{\mathfrak{S}_n, g}$. This proves (13.117.8).

[872] *Proof of (13.117.9):* Let λ be a partition of a nonnegative integer n. Fix a permutation $g \in \mathfrak{S}_n$ having cycle type λ. (Such a g clearly exists.) In the proof of (13.117.8), we have seen that $\widetilde{z}_\lambda = \left|Z_{\mathfrak{S}_n}(g)\right|$ and that $\left\{h \in \mathfrak{S}_n \mid h \text{ has cycle type } \lambda\right\} = \operatorname{Conj}_{\mathfrak{S}_n}(g)$.

Now, it is well-known that for every finite group G and every element $f \in G$, we have $\left|G / Z_G(f)\right| = \left|\operatorname{Conj}_G(f)\right|$ (in fact, there is a canonical bijection from the G-set $G / Z_G(f)$ to $\operatorname{Conj}_G(f)$, which sends the equivalence class $[\gamma] \in G / Z_G(f)$ of every $\gamma \in G$ to $\gamma f \gamma^{-1} \in \operatorname{Conj}_G(f)$). Applying this to $G = \mathfrak{S}_n$ and $f = g$, we obtain $\left|\mathfrak{S}_n / Z_{\mathfrak{S}_n}(g)\right| = \left|\operatorname{Conj}_{\mathfrak{S}_n}(g)\right|$. Thus, $\left|\operatorname{Conj}_{\mathfrak{S}_n}(g)\right| = \left|\mathfrak{S}_n / Z_{\mathfrak{S}_n}(g)\right| = \underbrace{\left|\mathfrak{S}_n\right|}_{=n!} / \underbrace{\left|Z_{\mathfrak{S}_n}(g)\right|}_{=\widetilde{z}_\lambda} = n!/\widetilde{z}_\lambda$, so that

$$n!/\widetilde{z}_\lambda = \left|\underbrace{\operatorname{Conj}_{\mathfrak{S}_n}(g)}_{=\left\{h \in \mathfrak{S}_n \mid h \text{ has cycle type } \lambda\right\}}\right| = \left|\left\{h \in \mathfrak{S}_n \mid h \text{ has cycle type } \lambda\right\}\right|.$$

This proves (13.117.9).

[873] This is because

$$\Phi\left(\underbrace{1}_{=p_\varnothing}\right) = \Phi\left(p_\varnothing\right) = \underbrace{\widetilde{z}_\varnothing}_{=1} \underbrace{\underline{1}_\varnothing}_{=1} \qquad \left(\text{by the definition of } \Phi\left(p_\varnothing\right)\right)$$
$$= 1.$$

a g clearly exists.) Let h be a permutation in $\mathfrak{S}_m$ having cycle type ν. (Such an h clearly exists.) Applying (13.117.8) to μ instead of λ, we obtain $\Phi(p_\mu) = \alpha_{\mathfrak{S}_n,g}$. Thus,

$$\Phi\left(\underbrace{u}_{=p_\mu}\right) = \Phi(p_\mu) = \alpha_{\mathfrak{S}_n,g}. \tag{13.117.10}$$

Applying (13.117.8) to ν, m and h instead of λ, n and g, we obtain $\Phi(p_\nu) = \alpha_{\mathfrak{S}_m,h}$. Thus,

$$\Phi\left(\underbrace{v}_{=p_\nu}\right) = \Phi(p_\nu) = \alpha_{\mathfrak{S}_m,h}. \tag{13.117.11}$$

The canonical isomorphism $R_{\mathbb{C}}(\mathfrak{S}_n) \otimes R_{\mathbb{C}}(\mathfrak{S}_m) \to R_{\mathbb{C}}(\mathfrak{S}_n \times \mathfrak{S}_m)$ sends $\alpha_{\mathfrak{S}_n,g} \otimes \alpha_{\mathfrak{S}_m,h}$ to $\alpha_{\mathfrak{S}_n \times \mathfrak{S}_m,(g,h)}$ (according to Exercise 4.4.3(c), applied to $G_1 = \mathfrak{S}_n$, $G_2 = \mathfrak{S}_m$, $h_1 = g$ and $h_2 = h$). Let us identify $R_{\mathbb{C}}(\mathfrak{S}_n) \otimes R_{\mathbb{C}}(\mathfrak{S}_m)$ with $R_{\mathbb{C}}(\mathfrak{S}_n \times \mathfrak{S}_m)$ along this isomorphism. Then, the statement we just made rewrites as follows:

$$\alpha_{\mathfrak{S}_n,g} \otimes \alpha_{\mathfrak{S}_m,h} = \alpha_{\mathfrak{S}_n \times \mathfrak{S}_m,(g,h)}. \tag{13.117.12}$$

Now, multiplying the equalities (13.117.10) and (13.117.11), we obtain

$$\begin{aligned}
\Phi(u)\Phi(v) &= \alpha_{\mathfrak{S}_n,g}\alpha_{\mathfrak{S}_m,h} \\
&= \operatorname{Ind}_{\mathfrak{S}_n \times \mathfrak{S}_m}^{\mathfrak{S}_{n+m}} \underbrace{(\alpha_{\mathfrak{S}_n,g} \otimes \alpha_{\mathfrak{S}_m,h})}_{\substack{=\alpha_{\mathfrak{S}_n \times \mathfrak{S}_m,(g,h)} \\ \text{(by (13.117.12))}}} \\
&\qquad \left(\text{by (13.117.7), applied to } \beta = \alpha_{\mathfrak{S}_n,g} \text{ and } \gamma = \alpha_{\mathfrak{S}_m,h}\right) \\
&= \operatorname{Ind}_{\mathfrak{S}_n \times \mathfrak{S}_m}^{\mathfrak{S}_{n+m}} \alpha_{\mathfrak{S}_n \times \mathfrak{S}_m,(g,h)} = \alpha_{\mathfrak{S}_{n+m},(g,h)} \qquad (13.117.13)
\end{aligned}$$

(by Exercise 4.4.3(b), applied to $\mathfrak{S}_{n+m}$, $\mathfrak{S}_n \times \mathfrak{S}_m$ and (g,h) instead of G, H and h).

Let λ be the partition whose parts are $\mu_1, \mu_2, \ldots, \mu_{\ell(\mu)}, \nu_1, \nu_2, \ldots, \nu_{\ell(\nu)}$. Then, $(\lambda_1, \lambda_2, \ldots, \lambda_{\ell(\lambda)})$ is a permutation of the list $(\mu_1, \mu_2, \ldots, \mu_{\ell(\mu)}, \nu_1, \nu_2, \ldots, \nu_{\ell(\nu)})$, and the partition λ has size

$$|\lambda| = \underbrace{\mu_1 + \mu_2 + \cdots + \mu_{\ell(\mu)}}_{=|\mu|=n} + \underbrace{\nu_1 + \nu_2 + \cdots + \nu_{\ell(\nu)}}_{=|\nu|=m} = n + m.$$

We also have $p_\lambda = p_\mu p_\nu$ [874]. Since $p_\mu = u$ and $p_\nu = v$, this rewrites as $p_\lambda = uv$.

Recall that the permutation g has cycle type μ. In other words, the cycles of g have lengths μ_1, μ_2, $\ldots$, $\mu_{\ell(\mu)}$. Similarly, the cycles of h have lengths ν_1, ν_2, $\ldots$, $\nu_{\ell(\nu)}$. Hence, the cycles of the permutation $(g,h) \in \mathfrak{S}_{n+m}$ (which, as we recall, sends every $i \in \{1, 2, \ldots, n+m\}$ to $\begin{cases} g(i), & \text{if } i \leq n; \\ n + h(i-n), & \text{if } i > n \end{cases}$) have lengths μ_1, μ_2, $\ldots$, $\mu_{\ell(\mu)}$, ν_1, ν_2, $\ldots$, $\nu_{\ell(\nu)}$ (in fact, on the subset $\{1, 2, \ldots, n\}$ of $\{1, 2, \ldots, n+m\}$, the permutation (g,h) acts as g and thus has cycles of lengths μ_1, μ_2, $\ldots$, $\mu_{\ell(\mu)}$, whereas on the complementary subset $\{n+1, n+2, \ldots, n+m\}$ of $\{1, 2, \ldots, n+m\}$, the permutation (g,h) acts as (a shifted version of) h and thus has cycles of lengths ν_1, ν_2, $\ldots$, $\nu_{\ell(\nu)}$). In other words, the cycles of the permutation $(g,h) \in \mathfrak{S}_{n+m}$ have lengths λ_1, λ_2, $\ldots$, $\lambda_{\ell(\lambda)}$ (since $(\lambda_1, \lambda_2, \ldots, \lambda_{\ell(\lambda)})$ is a permutation of the list $(\mu_1, \mu_2, \ldots, \mu_{\ell(\mu)}, \nu_1, \nu_2, \ldots, \nu_{\ell(\nu)})$). In other words, the permutation $(g,h) \in \mathfrak{S}_{n+m}$ has cycle type λ. Thus, (13.117.8) (applied to $n+m$ and (g,h) instead of n and g) yields $\Phi(p_\lambda) = \alpha_{\mathfrak{S}_{n+m},(g,h)}$. Compared with (13.117.13), this yields $\Phi(u)\Phi(v) = \Phi(p_\lambda)$. Since $p_\lambda = uv$, this rewrites as $\Phi(u)\Phi(v) = \Phi(uv)$. Thus,

[874] *Proof.* By the definition of p_μ, we have $p_\mu = p_{\mu_1} p_{\mu_2} \cdots p_{\mu_{\ell(\mu)}}$. Similarly, $p_\nu = p_{\nu_1} p_{\nu_2} \cdots p_{\nu_{\ell(\nu)}}$. Multiplying these two equalities, we obtain $p_\mu p_\nu = \left(p_{\mu_1} p_{\mu_2} \cdots p_{\mu_{\ell(\mu)}}\right)\left(p_{\nu_1} p_{\nu_2} \cdots p_{\nu_{\ell(\nu)}}\right)$.

But the definition of p_λ yields $p_\lambda = p_{\lambda_1} p_{\lambda_2} \cdots p_{\lambda_{\ell(\lambda)}}$. The product $p_{\lambda_1} p_{\lambda_2} \cdots p_{\lambda_{\ell(\lambda)}}$ has the same factors as the product $p_{\mu_1} p_{\mu_2} \cdots p_{\mu_{\ell(\mu)}} p_{\nu_1} p_{\nu_2} \cdots p_{\nu_{\ell(\nu)}}$ but possibly in a different order (since $(\lambda_1, \lambda_2, \ldots, \lambda_{\ell(\lambda)})$ is a permutation of the list $(\mu_1, \mu_2, \ldots, \mu_{\ell(\mu)}, \nu_1, \nu_2, \ldots, \nu_{\ell(\nu)})$). Hence, $p_{\lambda_1} p_{\lambda_2} \cdots p_{\lambda_{\ell(\lambda)}} = p_{\mu_1} p_{\mu_2} \cdots p_{\mu_{\ell(\mu)}} p_{\nu_1} p_{\nu_2} \cdots p_{\nu_{\ell(\nu)}}$, so that

$$p_\lambda = p_{\lambda_1} p_{\lambda_2} \cdots p_{\lambda_{\ell(\lambda)}} = p_{\mu_1} p_{\mu_2} \cdots p_{\mu_{\ell(\mu)}} p_{\nu_1} p_{\nu_2} \cdots p_{\nu_{\ell(\nu)}} = \left(p_{\mu_1} p_{\mu_2} \cdots p_{\mu_{\ell(\mu)}}\right)\left(p_{\nu_1} p_{\nu_2} \cdots p_{\nu_{\ell(\nu)}}\right) = p_\mu p_\nu,$$

qed.

$\Phi(uv) = \Phi(u)\Phi(v)$ is proven. As explained, this completes the proof of the fact that Φ is a $\mathbb{C}$-algebra homomorphism.

But we know that $\operatorname{ch} : A \to \Lambda$ is a $\mathbb{Z}$-Hopf algebra isomorphism, and thus the extension of ch to a $\mathbb{C}$-linear map $A_{\mathbb{C}} \to \Lambda_{\mathbb{C}}$ is a $\mathbb{C}$-Hopf algebra isomorphism. We shall denote this extension by $\operatorname{ch}_{\mathbb{C}}$.

We now shall show that $\operatorname{ch}_{\mathbb{C}} \circ \Phi = \operatorname{id}_{\Lambda_{\mathbb{C}}}$. Indeed, let n be a positive integer. Then, $\widetilde{z}_{(n)} = n$ [875]. Now, the definition of $\Phi\left(p_{(n)}\right)$ yields $\Phi\left(p_{(n)}\right) = \underbrace{\widetilde{z}_{(n)}}_{=n} \underline{1}_{(n)} = n\underline{1}_{(n)}$, so that

$$\left(\operatorname{ch}_{\mathbb{C}} \circ \Phi\right)(p_n) = \operatorname{ch}_{\mathbb{C}}\left(\Phi\left(\underbrace{p_n}_{=p_{(n)}}\right)\right) = \operatorname{ch}_{\mathbb{C}}\left(\underbrace{\Phi\left(p_{(n)}\right)}_{=n\underline{1}_{(n)}}\right) = \operatorname{ch}_{\mathbb{C}}\left(n\underline{1}_{(n)}\right) = n \underbrace{\operatorname{ch}_{\mathbb{C}}\left(\underline{1}_{(n)}\right)}_{\substack{=\frac{p_n}{n} \\ \text{(by a part of Theorem 4.4.1} \\ \text{that is already proven)}}}$$
$$= n \cdot \frac{p_n}{n} = p_n = \operatorname{id}_{\Lambda_{\mathbb{C}}}(p_n).$$

Now, let us forget that we fixed n. We thus have shown that $\left(\operatorname{ch}_{\mathbb{C}} \circ \Phi\right)(p_n) = \operatorname{id}_{\Lambda_{\mathbb{C}}}(p_n)$ for every positive integer n. Thus, the $\mathbb{C}$-algebra homomorphisms $\operatorname{ch}_{\mathbb{C}} \circ \Phi$ and $\operatorname{id}_{\Lambda_{\mathbb{C}}}$ are equal to each other on p_n for every positive integer n. Since $(p_n)_{n \geq 1}$ is a generating set of the $\mathbb{C}$-algebra $\Lambda_{\mathbb{C}}$, this yields that these homomorphisms are equal to each other on a generating set of the $\mathbb{C}$-algebra $\Lambda_{\mathbb{C}}$. Hence, these homomorphisms must be identical. That is, we have $\operatorname{ch}_{\mathbb{C}} \circ \Phi = \operatorname{id}_{\Lambda_{\mathbb{C}}}$.

Since $\operatorname{ch}_{\mathbb{C}}$ is an isomorphism, this yields that Φ is the inverse of $\operatorname{ch}_{\mathbb{C}}$. That is, $\Phi = \left(\operatorname{ch}_{\mathbb{C}}\right)^{-1}$.

Corollary 2.5.17(b) (applied to $\mathbf{k} = \mathbb{C}$) yields that $\{p_\lambda\}$ and $\left\{z_\lambda^{-1} p_\lambda\right\}$ are dual bases of $\Lambda_{\mathbb{C}}$ with respect to the Hall inner product on Λ. Thus,

$$\left(p_\lambda, z_\mu^{-1} p_\mu\right)_{\Lambda_{\mathbb{C}}} = \delta_{\lambda,\mu} \qquad \text{for any partitions } \lambda \text{ and } \mu. \tag{13.117.15}$$

[875] *Proof.* Let g denote the n-cycle $(1, 2, \ldots, n)$ in $\mathfrak{S}_n$. Then, g is a permutation in $\mathfrak{S}_n$ having cycle type (n). But $\widetilde{z}_{(n)}$ is defined as the size of the centralizer of a permutation in $\mathfrak{S}_{|(n)|}$ having cycle type (n). Hence, $\widetilde{z}_{(n)}$ is the size of the centralizer of g in $\mathfrak{S}_n$ (since g is a permutation in $\mathfrak{S}_n = \mathfrak{S}_{|(n)|}$ having cycle type (n)). In other words, $\widetilde{z}_{(n)} = \left|Z_{\mathfrak{S}_n}(g)\right|$.

It is clear that the subgroup $\langle g \rangle$ of $\mathfrak{S}_n$ generated by g satisfies $\langle g \rangle \subset Z_{\mathfrak{S}_n}(g)$ (since every power of g centralizes g). We shall now show that $\langle g \rangle = Z_{\mathfrak{S}_n}(g)$.

Indeed, let $z \in Z_{\mathfrak{S}_n}(g)$. Then, z must centralize g. That is, we have $zgz^{-1} = g$. Hence, $zg = gz$, so that the elements g and z of $\mathfrak{S}_n$ commute. Hence, these elements z and g generate a commutative subgroup of $\mathfrak{S}_n$. Denote this subgroup by T.

Now, every $i \in \{1, 2, \ldots, n\}$ satisfies

$$g^{i-1}(1) = i \tag{13.117.14}$$

(since g is the n-cycle $(1, 2, \ldots, n)$). Now, let $j = z(1)$. Then, $g^{j-1}(1) = j$ (by (13.117.14), applied to $i = j$). On the other hand, let $i \in \{1, 2, \ldots, n\}$. Then, g^{i-1} commutes with z (since g commutes with z); in other words, $g^{i-1}z = zg^{i-1}$. Now,

$$z\left(\underbrace{i}_{\substack{=g^{i-1}(1) \\ \text{(by (13.117.14))}}}\right) = z\left(g^{i-1}(1)\right) = \underbrace{\left(zg^{i-1}\right)}_{=g^{i-1}z}(1) = \left(g^{i-1}z\right)(1) = g^{i-1}\left(\underbrace{z(1)}_{=j=g^{j-1}(1)}\right)$$
$$= g^{i-1}\left(g^{j-1}(1)\right) = \underbrace{\left(g^{i-1}g^{j-1}\right)}_{=g^{(i-1)+(j-1)}=g^{j-1}g^{i-1}}(1) = \left(g^{j-1}g^{i-1}\right)(1) = g^{j-1}\left(\underbrace{g^{i-1}(1)}_{\substack{=i \\ \text{(by (13.117.14))}}}\right) = g^{j-1}(i).$$

Let us now forget that we fixed i. We thus have shown that $z(i) = g^{j-1}(i)$ for every $i \in \{1, 2, \ldots, n\}$. Hence, $z = g^{j-1} \in \langle g \rangle$.

Now, let us forget that we fixed z. We thus have proven that $z \in \langle g \rangle$ for every $z \in Z_{\mathfrak{S}_n}(g)$. Hence, $Z_{\mathfrak{S}_n}(g) \subset \langle g \rangle$. Combined with $\langle g \rangle \subset Z_{\mathfrak{S}_n}(g)$, this yields $\langle g \rangle = Z_{\mathfrak{S}_n}(g)$. Hence, $|\langle g \rangle| = \left|Z_{\mathfrak{S}_n}(g)\right|$. Compared with $\widetilde{z}_{(n)} = \left|Z_{\mathfrak{S}_n}(g)\right|$, this yields $\widetilde{z}_{(n)} = |\langle g \rangle| = (\text{the order of } g \text{ in } \mathfrak{S}_n) = n$ (since g is an n-cycle), qed.

Hence, every partition λ satisfies

$$\left(p_\lambda, \underbrace{p_\lambda}_{=z_\lambda z_\lambda^{-1} p_\lambda}\right)_{\Lambda_{\mathbb{C}}} = \left(p_\lambda, z_\lambda z_\lambda^{-1} p_\lambda\right)_{\Lambda_{\mathbb{C}}} = z_\lambda \underbrace{\left(p_\lambda, z_\lambda^{-1} p_\lambda\right)_{\Lambda_{\mathbb{C}}}}_{\substack{=\delta_{\lambda,\lambda} \\ \text{(by (13.117.15), applied} \\ \text{to } \mu=\lambda)}} = z_\lambda \underbrace{\delta_{\lambda,\lambda}}_{=1} = z_\lambda. \tag{13.117.16}$$

Our goal is to prove that $\operatorname{ch}_{\mathbb{C}}(\underline{1}_\lambda) = \dfrac{p_\lambda}{z_\lambda}$ for every partition λ.

Let λ be a partition. Let n be the size of λ. Then, $\lambda \in \operatorname{Par}_n \subset \operatorname{Par}$ and $\underline{1}_\lambda \in R_{\mathbb{C}}(\mathfrak{S}_n)$. We know that $\widetilde{z}_\lambda$ is a positive integer (since $\widetilde{z}_\lambda$ is defined as the size of a centralizer, and centralizers are subgroups). Hence, we can divide by $\widetilde{z}_\lambda$, and we have

$$\Phi\left(\widetilde{z}_\lambda^{-1} p_\lambda\right) = \widetilde{z}_\lambda^{-1} \underbrace{\Phi(p_\lambda)}_{\substack{=\widetilde{z}_\lambda \underline{1}_\lambda \\ \text{(by the definition of } \Phi(p_\lambda))}} = \widetilde{z}_\lambda^{-1} \widetilde{z}_\lambda \underline{1}_\lambda = \underline{1}_\lambda.$$

Since $\Phi = (\operatorname{ch}_{\mathbb{C}})^{-1}$, this rewrites as $(\operatorname{ch}_{\mathbb{C}})^{-1}\left(\widetilde{z}_\lambda^{-1} p_\lambda\right) = \underline{1}_\lambda$. Hence,

$$\operatorname{ch}_{\mathbb{C}}(\underline{1}_\lambda) = \widetilde{z}_\lambda^{-1} p_\lambda. \tag{13.117.17}$$

Recall that we want to prove that $\operatorname{ch}_{\mathbb{C}}(\underline{1}_\lambda) = \dfrac{p_\lambda}{z_\lambda}$. If we can show that $\widetilde{z}_\lambda = z_\lambda$, then (13.117.17) becomes $\operatorname{ch}_{\mathbb{C}}(\underline{1}_\lambda) = \left(\underbrace{\widetilde{z}_\lambda}_{=z_\lambda}\right)^{-1} p_\lambda = z_\lambda^{-1} p_\lambda = \dfrac{p_\lambda}{z_\lambda}$, and thus $\operatorname{ch}_{\mathbb{C}}(\underline{1}_\lambda) = \dfrac{p_\lambda}{z_\lambda}$ will be proven. Hence, all that remains to be done is proving $\widetilde{z}_\lambda = z_\lambda$.

But ch is a PSH-isomorphism, thus an isometry. Hence, $\operatorname{ch}_{\mathbb{C}}$ (being the extension of ch to a $\mathbb{C}$-linear map) must also be an isometry, i.e., we must have

$$(\operatorname{ch}_{\mathbb{C}} \beta, \operatorname{ch}_{\mathbb{C}} \gamma)_{\Lambda_{\mathbb{C}}} = (\beta, \gamma)_{A_{\mathbb{C}}} \qquad \text{for all } \beta \in A_{\mathbb{C}} \text{ and } \gamma \in A_{\mathbb{C}}.$$

Applying this to $\beta = \underline{1}_\lambda$ and $\gamma = \underline{1}_\lambda$, we obtain

$$\begin{aligned}
&(\operatorname{ch}_{\mathbb{C}}(\underline{1}_\lambda), \operatorname{ch}_{\mathbb{C}}(\underline{1}_\lambda))_{\Lambda_{\mathbb{C}}} \\
&= (\underline{1}_\lambda, \underline{1}_\lambda)_{A_{\mathbb{C}}} = \langle \underline{1}_\lambda, \underline{1}_\lambda \rangle_{\mathfrak{S}_n} \qquad \left(\begin{array}{c}\text{since } \underline{1}_\lambda \in R_{\mathbb{C}}(\mathfrak{S}_n)\text{, and since the bilinear form} \\ (\cdot,\cdot)_{A_{\mathbb{C}}} \text{ on } A_{\mathbb{C}} \text{ extends the bilinear form } \langle\cdot,\cdot\rangle_{\mathfrak{S}_n} \text{ on } R_{\mathbb{C}}(\mathfrak{S}_n)\end{array}\right) \\
&= \underbrace{\frac{1}{|\mathfrak{S}_n|}}_{=\frac{1}{n!}} \sum_{g \in \mathfrak{S}_n} \underbrace{\underline{1}_\lambda(g)}_{\substack{=[g \text{ has cycle type } \lambda] \\ \text{(by the definition of } \underline{1}_\lambda)}} \underbrace{\underline{1}_\lambda(g^{-1})}_{\substack{=[g^{-1} \text{ has cycle type } \lambda] \\ \text{(by the definition of } \underline{1}_\lambda)}} \qquad \left(\text{by the definition of the bilinear form } \langle\cdot,\cdot\rangle_{\mathfrak{S}_n}\right) \\
&= \frac{1}{n!} \sum_{g \in \mathfrak{S}_n} [g \text{ has cycle type } \lambda] \left[\underbrace{g^{-1} \text{ has cycle type } \lambda}_{\substack{\text{this is equivalent to} \\ (g \text{ has cycle type } \lambda) \\ \text{(since the cycle type of } g^{-1} \\ \text{equals the cycle type of } g)}}\right] = \frac{1}{n!} \sum_{g \in \mathfrak{S}_n} \underbrace{[g \text{ has cycle type } \lambda][g \text{ has cycle type } \lambda]}_{\substack{=[(g \text{ has cycle type } \lambda) \text{ and } (g \text{ has cycle type } \lambda)] \\ =[g \text{ has cycle type } \lambda]}} \\
&= \frac{1}{n!} \underbrace{\sum_{g \in \mathfrak{S}_n} [g \text{ has cycle type } \lambda]}_{\substack{=|\{h \in \mathfrak{S}_n \mid h \text{ has cycle type } \lambda\}| = n!/\widetilde{z}_\lambda \\ \text{(by (13.117.9))}}} = \frac{1}{n!} \cdot \frac{n!}{\widetilde{z}_\lambda} = \frac{1}{\widetilde{z}_\lambda}.
\end{aligned}$$

Hence,

$$\frac{1}{\widetilde{z}_\lambda} = \left(\underbrace{\operatorname{ch}_{\mathbb{C}}\left(\underline{1}_\lambda\right)}_{\substack{=\widetilde{z}_\lambda^{-1} p_\lambda \\ \text{(by (13.117.17))}}} , \underbrace{\operatorname{ch}_{\mathbb{C}}\left(\underline{1}_\lambda\right)}_{\substack{=\widetilde{z}_\lambda^{-1} p_\lambda \\ \text{(by (13.117.17))}}} \right)_{\Lambda_{\mathbb{C}}} = \left(\widetilde{z}_\lambda^{-1} p_\lambda, \widetilde{z}_\lambda^{-1} p_\lambda\right)_{\Lambda_{\mathbb{C}}} = \left(\widetilde{z}_\lambda^{-1}\right)^2 \left(p_\lambda, p_\lambda\right)_{\Lambda_{\mathbb{C}}}.$$

Multiplying this equality with $\widetilde{z}_\lambda^2$, we obtain

$$\widetilde{z}_\lambda = \left(p_\lambda, p_\lambda\right)_{\Lambda_{\mathbb{C}}} = z_\lambda \qquad \text{(by (13.117.16))}.$$

Thus, $\widetilde{z}_\lambda = z_\lambda$ is proven. As we have explained, this concludes the proof of $\operatorname{ch}_{\mathbb{C}}\left(\underline{1}_\lambda\right) = \frac{p_\lambda}{z_\lambda}$, and thus Exercise 4.4.3(d) is solved.

(e) Let λ be a partition. Let us work with the notations of the solution of Exercise 4.4.3(d) above. In the latter solution, we have shown that $\widetilde{z}_\lambda = z_\lambda$. Thus,

$$\begin{aligned} z_\lambda = \widetilde{z}_\lambda &= \left(\text{the size of the centralizer of a permutation in } \mathfrak{S}_{|\lambda|} \text{ having cycle type } \lambda\right) \\ &\qquad \left(\text{by the definition of } \widetilde{z}_\lambda\right) \\ &= \left(\text{the size of the centralizer in } \mathfrak{S}_n \text{ of a permutation having cycle type } \lambda, \text{ where } n = |\lambda|\right). \end{aligned}$$

This proves Remark 2.5.16. Thus, Exercise 4.4.3(e) is solved.

(f) Let G and H be two finite groups. Let $\rho : H \to G$ be a group homomorphism. Let $y \in H$. We shall show that $\operatorname{Ind}_\rho \alpha_{H,y} = \alpha_{G,\rho(y)}$.

Let $u \in G$. Then, the definition of $\operatorname{Ind}_\rho \alpha_{H,y}$ yields

$$\begin{aligned} \left(\operatorname{Ind}_\rho \alpha_{H,y}\right)(u) &= \frac{1}{|H|} \sum_{\substack{(h,k) \in H \times G; \\ k\rho(h)k^{-1} = u}} \alpha_{H,y}(h) = \frac{1}{|H|} \sum_{\substack{(h,x) \in H \times G; \\ x\rho(h)x^{-1} = u}} \underbrace{\alpha_{H,y}(h)}_{\substack{= \sum_{k \in H} \left[kyk^{-1} = h\right] \\ \text{(by Exercise 4.4.3(a), applied to} \\ H,\ y \text{ and } h \text{ instead of } G,\ h \text{ and } g)}} \\ &\qquad \left(\text{here, we renamed the summation index } (h,k) \text{ as } (h,x)\right) \\ &= \frac{1}{|H|} \underbrace{\sum_{\substack{(h,x) \in H \times G; \\ x\rho(h)x^{-1} = u}}}_{= \sum_{x \in G} \sum_{\substack{h \in H; \\ x\rho(h)x^{-1} = u}}} \sum_{k \in H} \left[kyk^{-1} = h\right] \\ &= \frac{1}{|H|} \sum_{x \in G} \underbrace{\sum_{\substack{h \in H; \\ x\rho(h)x^{-1} = u}} \sum_{k \in H}}_{= \sum_{k \in H} \sum_{\substack{h \in H; \\ x\rho(h)x^{-1} = u}}} \left[kyk^{-1} = h\right] \\ &= \frac{1}{|H|} \sum_{x \in G} \sum_{k \in H} \sum_{\substack{h \in H; \\ x\rho(h)x^{-1} = u}} \left[kyk^{-1} = h\right]. \end{aligned} \tag{13.117.18}$$

But every $x \in G$ and $k \in H$ satisfy

$$\sum_{\substack{h \in H; \\ x\rho(h)x^{-1} = u}} \left[kyk^{-1} = h\right] = \left[x\rho(k)\rho(y)\left(x\rho(k)\right)^{-1} = u\right] \tag{13.117.19}$$

[876]. Hence, (13.117.18) becomes

$$\begin{aligned}
\left(\operatorname{Ind}_\rho \alpha_{H,y}\right)(u) &= \frac{1}{|H|} \underbrace{\sum_{x\in G}\sum_{k\in H}}_{=\sum_{k\in H}\sum_{x\in G}} \underbrace{\sum_{\substack{h\in H;\\ x\rho(h)x^{-1}=u}} \left[kyk^{-1}=h\right]}_{\substack{=\left[x\rho(k)\rho(y)(x\rho(k))^{-1}=u\right]\\ \text{(by (13.117.19))}}} \\
&= \frac{1}{|H|}\sum_{k\in H} \underbrace{\sum_{x\in G}\left[x\rho(k)\rho(y)(x\rho(k))^{-1}=u\right]}_{\substack{=\sum_{x\in G}\left[x\rho(y)x^{-1}=u\right]\\ \text{(here, we have substituted } x \text{ for } x\rho(k) \text{ in the sum}\\ \text{(because the map } G\to G,\ x\mapsto x\rho(k) \text{ is a bijection}\\ \text{(since } G \text{ is a group, and since } \rho(k)\in G)))}} = \frac{1}{|H|} \underbrace{\sum_{k\in H}\sum_{x\in G}\left[x\rho(y)x^{-1}=u\right]}_{=|H|\cdot\sum_{x\in G}\left[x\rho(y)x^{-1}=u\right]} \\
&= \frac{1}{|H|}|H|\cdot\sum_{x\in G}\left[x\rho(y)x^{-1}=u\right] = \sum_{x\in G}\left[x\rho(y)x^{-1}=u\right] \\
&= \sum_{k\in G}\left[k\rho(y)k^{-1}=u\right] \qquad \text{(here, we renamed the summation index } x \text{ as } k).
\end{aligned}$$

Compared with

$$\alpha_{G,\rho(y)}(u) = \sum_{k\in G}\left[k\rho(y)k^{-1}=u\right] \qquad \text{(by Exercise 4.4.3(a), applied to } \rho(y) \text{ and } u \text{ instead of } h \text{ and } g),$$

[876] *Proof of (13.117.19):* Let $x\in G$ and $k\in H$. We have

$$\begin{aligned}
&\sum_{h\in H}\left[x\rho(h)x^{-1}=u\right]\left[kyk^{-1}=h\right] \\
&= \sum_{\substack{h\in H;\\ x\rho(h)x^{-1}=u}} \underbrace{\left[x\rho(h)x^{-1}=u\right]}_{\substack{=1\\ \text{(since we have } x\rho(h)x^{-1}=u)}} \left[kyk^{-1}=h\right] + \sum_{\substack{h\in H;\\ x\rho(h)x^{-1}\neq u}} \underbrace{\left[x\rho(h)x^{-1}=u\right]}_{\substack{=0\\ \text{(since we don't have } x\rho(h)x^{-1}=u\\ \text{(since } x\rho(h)x^{-1}\neq u))}} \left[kyk^{-1}=h\right] \\
&= \sum_{\substack{h\in H;\\ x\rho(h)x^{-1}=u}} \left[kyk^{-1}=h\right] + \underbrace{\sum_{\substack{h\in H;\\ x\rho(h)x^{-1}\neq u}} 0\left[kyk^{-1}=h\right]}_{=0} = \sum_{\substack{h\in H;\\ x\rho(h)x^{-1}=u}} \left[kyk^{-1}=h\right],
\end{aligned}$$

so that

$$\begin{aligned}
\sum_{\substack{h\in H;\\ x\rho(h)x^{-1}=u}} \left[kyk^{-1}=h\right] &= \sum_{h\in H}\left[x\rho(h)x^{-1}=u\right]\left[kyk^{-1}=h\right] \\
&= \sum_{\substack{h\in H;\\ h=kyk^{-1}}}\left[x\rho(h)x^{-1}=u\right] \underbrace{\left[kyk^{-1}=h\right]}_{\substack{=1\\ \text{(since we have } kyk^{-1}=h\\ \text{(since } h=kyk^{-1}))}} + \sum_{\substack{h\in H;\\ h\neq kyk^{-1}}}\left[x\rho(h)x^{-1}=u\right] \underbrace{\left[kyk^{-1}=h\right]}_{\substack{=0\\ \text{(since we don't have } kyk^{-1}=h\\ \text{(since } kyk^{-1}\neq h \text{ (since } h\neq kyk^{-1})))}} \\
&= \sum_{\substack{h\in H;\\ h=kyk^{-1}}}\left[x\rho(h)x^{-1}=u\right] + \underbrace{\sum_{\substack{h\in H;\\ h\neq kyk^{-1}}}\left[x\rho(h)x^{-1}=u\right]0}_{=0} = \sum_{\substack{h\in H;\\ h=kyk^{-1}}}\left[x\rho(h)x^{-1}=u\right] \\
&= \left[x \underbrace{\rho\left(kyk^{-1}\right)}_{\substack{=\rho(k)\rho(y)(\rho(k))^{-1}\\ \text{(since } \rho \text{ is a group}\\ \text{homomorphism)}}} x^{-1}=u\right] \qquad \left(\text{since } kyk^{-1}\in H \text{ (since } k\in H \text{ and } y\in H)\right) \\
&= \left[x\rho(k)\rho(y)\underbrace{(\rho(k))^{-1}x^{-1}}_{=(x\rho(k))^{-1}}=u\right] = \left[x\rho(k)\rho(y)(x\rho(k))^{-1}=u\right].
\end{aligned}$$

This proves (13.117.19).

this yields $\left(\operatorname{Ind}_\rho \alpha_{H,y}\right)(u) = \alpha_{G,\rho(y)}(u)$.

Let us now forget that we fixed u. We thus have shown that $\left(\operatorname{Ind}_\rho \alpha_{H,y}\right)(u) = \alpha_{G,\rho(y)}(u)$ for every $u \in G$. In other words, $\operatorname{Ind}_\rho \alpha_{H,y} = \alpha_{G,\rho(y)}$.

Let us now forget that we fixed y. We thus have shown that $\operatorname{Ind}_\rho \alpha_{H,y} = \alpha_{G,\rho(y)}$ for every $y \in H$. Renaming y as h in this statement, we obtain that $\operatorname{Ind}_\rho \alpha_{H,h} = \alpha_{G,\rho(h)}$ for every $h \in H$. This solves Exercise 4.4.3(f).

13.118. **Solution to Exercise 4.4.4.** *Solution to Exercise 4.4.4.*

Step 1: Study of inner tensor products.

The well-definedness of the inner tensor product is clear (since the inclusion map $G \to G \times G,\ g \mapsto (g, g)$ is a group homomorphism). We notice that if G is a group and U_1 and U_2 are two $\mathbb{C}G$-modules, then the character $\chi_{U_1 \otimes U_2}$ of the inner tensor product $U_1 \otimes U_2$ of U_1 and U_2 is given by

$$\chi_{U_1 \otimes U_2}(g) = \chi_{U_1}(g)\,\chi_{U_2}(g) \qquad \text{for all } g \in G. \tag{13.118.1}$$

[877]

Step 2: The involutions $\widetilde{\omega}_n$.

Now let $n \geq 0$. For every $f \in R_{\mathbb{C}}(\mathfrak{S}_n)$, it is easy to see that the map $\mathfrak{S}_n \to \mathbb{C}$ which sends every $g \in \mathfrak{S}_n$ to $\operatorname{sgn}(g) f(g)$ is a class function (because both $\operatorname{sgn}(g)$ and $f(g)$ are uniquely determined by the conjugacy class of g). This class function $\mathfrak{S}_n \to \mathbb{C}$ is denoted by $\operatorname{sgn}_{\mathfrak{S}_n} * f$ and belongs to $R_{\mathbb{C}}(\mathfrak{S}_n)$ (being a class function). We can thus define a map $\widetilde{\omega}_n : R_{\mathbb{C}}(\mathfrak{S}_n) \to R_{\mathbb{C}}(\mathfrak{S}_n)$ as follows:

$$\widetilde{\omega}_n(f) = \operatorname{sgn}_{\mathfrak{S}_n} * f \qquad \text{for all } f \in R_{\mathbb{C}}(\mathfrak{S}_n).$$

Consider this map $\widetilde{\omega}_n$. It is $\mathbb{C}$-linear (obviously) and an involution[878]. Hence, $\widetilde{\omega}_n$ is precisely the involution on class functions $f : \mathfrak{S}_n \to \mathbb{C}$ sending $f \mapsto \operatorname{sgn}_{\mathfrak{S}_n} * f$.

Now, let V be any finite-dimensional $\mathbb{C}\mathfrak{S}_n$-module. Then, every $g \in \mathfrak{S}_n$ satisfies

$$\underbrace{\left(\widetilde{\omega}_n(\chi_V)\right)}_{\substack{=\operatorname{sgn}_{\mathfrak{S}_n} * \chi_V \\ \text{(by the definition of } \widetilde{\omega}_n)}}(g) = \left(\operatorname{sgn}_{\mathfrak{S}_n} * \chi_V\right)(g) = \operatorname{sgn}(g)\,\chi_V(g)$$

and

$$\chi_{\operatorname{sgn}_{\mathfrak{S}_n} \otimes V}(g) = \underbrace{\chi_{\operatorname{sgn}_{\mathfrak{S}_n}}(g)}_{=\operatorname{sgn}(g)}\chi_V(g) \qquad (\text{by } (13.118.1))$$
$$= \operatorname{sgn}(g)\,\chi_V(g).$$

Hence, every $g \in \mathfrak{S}_n$ satisfies $\left(\widetilde{\omega}_n(\chi_V)\right)(g) = \operatorname{sgn}(g)\,\chi_V(g) = \chi_{\operatorname{sgn}_{\mathfrak{S}_n} \otimes V}(g)$. In other words, $\widetilde{\omega}_n(\chi_V) = \chi_{\operatorname{sgn}_{\mathfrak{S}_n} \otimes V}$.

Forget that we fixed V. We thus have proven that $\widetilde{\omega}_n(\chi_V) = \chi_{\operatorname{sgn}_{\mathfrak{S}_n} \otimes V}$ for every finite-dimensional $\mathbb{C}\mathfrak{S}_n$-module V. In particular, for every irreducible $\mathbb{C}\mathfrak{S}_n$-module V, we have $\widetilde{\omega}_n(\chi_V) = \chi_{\operatorname{sgn}_{\mathfrak{S}_n} \otimes V} \in R(\mathfrak{S}_n)$. Since the χ_V span $R(\mathfrak{S}_n)$ as a $\mathbb{Z}$-module as V ranges through (a set of representatives of the isomorphism classes of) the irreducible $\mathbb{C}\mathfrak{S}_n$-modules V, this entails that the involution $\widetilde{\omega}_n$ preserves the $\mathbb{Z}$-lattice $R(\mathfrak{S}_n)$. Since $\widetilde{\omega}_n$ is the involution on class functions $f : \mathfrak{S}_n \to \mathbb{C}$ sending $f \mapsto \operatorname{sgn}_{\mathfrak{S}_n} * f$, this rewrites as follows:

[877] This is because every $g \in G$ satisfies

$$\chi_{U_1 \otimes U_2}(g) = \operatorname{trace} \underbrace{\left(g : U_1 \otimes U_2 \to U_1 \otimes U_2\right)}_{\substack{=(g,g) : U_1 \otimes U_2 \to U_1 \otimes U_2 \\ \text{(by the definition of the inner tensor product)}}}$$
$$= \operatorname{trace} \underbrace{\left((g,g) : U_1 \otimes U_2 \to U_1 \otimes U_2\right)}_{=(g : U_1 \to U_1) \otimes (g : U_2 \to U_2)} = \operatorname{trace}\left((g : U_1 \to U_1) \otimes (g : U_2 \to U_2)\right)$$
$$= \underbrace{\operatorname{trace}(g : U_1 \to U_1)}_{=\chi_{U_1}(g)} \cdot \underbrace{\operatorname{trace}(g : U_2 \to U_2)}_{=\chi_{U_2}(g)} = \chi_{U_1}(g) \cdot \chi_{U_2}(g).$$

[878] This is because multiplying a scalar by $\operatorname{sgn}(g)$ twice (for fixed $g \in \mathfrak{S}_n$) does nothing (since $(\operatorname{sgn}(g))^2 = 1$).

The involution on class functions $f : \mathfrak{S}_n \to \mathbb{C}$ sending $f \mapsto \operatorname{sgn}_{\mathfrak{S}_n} * f$ preserves the $\mathbb{Z}$-lattice $R(\mathfrak{S}_n)$. This proves one claim of Theorem 4.4.1(b).

Recall that the involution $\widetilde{\omega}_n$ preserves the $\mathbb{Z}$-lattice $R(\mathfrak{S}_n)$. Thus, $\widetilde{\omega}_n$ restricts to a $\mathbb{Z}$-linear involution $R(\mathfrak{S}_n) \to R(\mathfrak{S}_n)$. Denote this involution by $\widetilde{\omega}'_n$. It clearly satisfies

$$\widetilde{\omega}'_n(f) = \widetilde{\omega}_n(f) \qquad (\text{by the definition of } \widetilde{\omega}'_n) \tag{13.118.2}$$
$$= \operatorname{sgn}_{\mathfrak{S}_n} * f \qquad \text{for all } f \in R(\mathfrak{S}_n). \tag{13.118.3}$$

Step 3: The involution $\widetilde{\omega}_{\mathbb{Z}}$.

Now, forget that we fixed n. We thus have constructed a $\mathbb{Z}$-linear involution $\widetilde{\omega}'_n : R(\mathfrak{S}_n) \to R(\mathfrak{S}_n)$ for every $n \geq 0$. The direct sum of these involutions over all $n \geq 0$ is a graded $\mathbb{Z}$-linear involution $\bigoplus_{n\geq 0} \widetilde{\omega}'_n : \bigoplus_{n\geq 0} R(\mathfrak{S}_n) \to \bigoplus_{n\geq 0} R(\mathfrak{S}_n)$. Denote this involution $\bigoplus_{n\geq 0} \widetilde{\omega}'_n$ by $\widetilde{\omega}_{\mathbb{Z}}$. Then, $\widetilde{\omega}_{\mathbb{Z}}$ is a graded $\mathbb{Z}$-linear involution $A(\mathfrak{S}) \to A(\mathfrak{S})$ (since $\bigoplus_{n\geq 0} R(\mathfrak{S}_n) = A(\mathfrak{S})$), and is precisely the involution on $A(\mathfrak{S})$ defined in Theorem 4.4.1(b).

We have $\widetilde{\omega}_{\mathbb{Z}} = \bigoplus_{n\geq 0} \widetilde{\omega}'_n$. Hence, for every $n \in \mathbb{N}$ and $f \in R_{\mathbb{C}}(\mathfrak{S}_n)$, we have

$$\widetilde{\omega}_{\mathbb{Z}}(f) = \widetilde{\omega}'_n(f) = \widetilde{\omega}_n(f) \qquad (\text{since } \widetilde{\omega}'_n \text{ is defined as a restriction of } \widetilde{\omega}_n) \tag{13.118.4}$$
$$= \operatorname{sgn}_{\mathfrak{S}_n} * f. \tag{13.118.5}$$

For every $n \in \mathbb{N}$ and every finite-dimensional $\mathbb{C}\mathfrak{S}_n$-module V, we have

$$\widetilde{\omega}_{\mathbb{Z}}(\chi_V) = \widetilde{\omega}_n(\chi_V) \qquad (\text{by (13.118.4), applied to } f = \chi_V)$$
$$= \chi_{\operatorname{sgn}_{\mathfrak{S}_n} \otimes V}. \tag{13.118.6}$$

In other words, the involution $\widetilde{\omega}_{\mathbb{Z}}$ sends χ_V to $\chi_{\operatorname{sgn}_{\mathfrak{S}_n} \otimes V}$ for every $n \in \mathbb{N}$ and every finite-dimensional $\mathbb{C}\mathfrak{S}_n$-module V. In other words, the involution on $A(\mathfrak{S})$ defined in Theorem 4.4.1(b) sends χ_V to $\chi_{\operatorname{sgn}_{\mathfrak{S}_n} \otimes V}$ for every $n \in \mathbb{N}$ and every finite-dimensional $\mathbb{C}\mathfrak{S}_n$-module V (since $\widetilde{\omega}_{\mathbb{Z}}$ is precisely the involution on $A(\mathfrak{S})$ defined in Theorem 4.4.1(b)). This proves one of the claims of Exercise 4.4.4.

Step 4: Properties of $\widetilde{\omega}_{\mathbb{Z}}$.

We are now going to prove that $\operatorname{ch} \circ \widetilde{\omega}_{\mathbb{Z}} = \omega \circ \operatorname{ch}$ as maps $A(\mathfrak{S}) \to \Lambda$.

Indeed, let λ be a partition. Let $n = |\lambda|$. Then, $\lambda \in \operatorname{Par}_n$. Theorem 4.4.1(a) yields $\operatorname{ch}\left(\operatorname{Ind}_{\mathfrak{S}_\lambda}^{\mathfrak{S}_n} \underline{1}_{\mathfrak{S}_\lambda}\right) = h_\lambda$ and $\operatorname{ch}\left(\operatorname{Ind}_{\mathfrak{S}_\lambda}^{\mathfrak{S}_n} \operatorname{sgn}_{\mathfrak{S}_\lambda}\right) = e_\lambda$. Since $\operatorname{ch}\left(\operatorname{Ind}_{\mathfrak{S}_\lambda}^{\mathfrak{S}_n} \underline{1}_{\mathfrak{S}_\lambda}\right) = h_\lambda$, we have $\operatorname{ch}^{-1}(h_\lambda) = \operatorname{Ind}_{\mathfrak{S}_\lambda}^{\mathfrak{S}_n} \underline{1}_{\mathfrak{S}_\lambda}$. Since $\operatorname{ch}\left(\operatorname{Ind}_{\mathfrak{S}_\lambda}^{\mathfrak{S}_n} \operatorname{sgn}_{\mathfrak{S}_\lambda}\right) = e_\lambda$, we have $\operatorname{ch}^{-1}(e_\lambda) = \operatorname{Ind}_{\mathfrak{S}_\lambda}^{\mathfrak{S}_n} \operatorname{sgn}_{\mathfrak{S}_\lambda}$.

On the other hand, $\omega(h_\lambda) = e_\lambda$ [879].

But it is easy to see that $\widetilde{\omega}_{\mathbb{Z}}\left(\operatorname{Ind}_{\mathfrak{S}_\lambda}^{\mathfrak{S}_n} \underline{1}_{\mathfrak{S}_\lambda}\right) = \operatorname{Ind}_{\mathfrak{S}_\lambda}^{\mathfrak{S}_n} \operatorname{sgn}_{\mathfrak{S}_\lambda}$. [880]

[879] This follows from the fact that ω is an algebra homomorphism and that $\omega(h_m) = e_m$ for every $m \in \mathbb{N}$.

[880] *Proof.* Let $g \in \mathfrak{S}_n$. We have $\widetilde{\omega}_{\mathbb{Z}}\left(\operatorname{Ind}_{\mathfrak{S}_\lambda}^{\mathfrak{S}_n} \underline{1}_{\mathfrak{S}_\lambda}\right) = \operatorname{sgn}_{\mathfrak{S}_n} * \left(\operatorname{Ind}_{\mathfrak{S}_\lambda}^{\mathfrak{S}_n} \underline{1}_{\mathfrak{S}_\lambda}\right)$ (by (13.118.4), applied to $f = \operatorname{Ind}_{\mathfrak{S}_\lambda}^{\mathfrak{S}_n} \underline{1}_{\mathfrak{S}_\lambda}$). Hence,

$$\left(\widetilde{\omega}_{\mathbb{Z}}\left(\operatorname{Ind}_{\mathfrak{S}_\lambda}^{\mathfrak{S}_n} \underline{1}_{\mathfrak{S}_\lambda}\right)\right)(g) = \left(\operatorname{sgn}_{\mathfrak{S}_n} * \left(\operatorname{Ind}_{\mathfrak{S}_\lambda}^{\mathfrak{S}_n} \underline{1}_{\mathfrak{S}_\lambda}\right)\right)(g) = \operatorname{sgn}(g) \cdot \left(\operatorname{Ind}_{\mathfrak{S}_\lambda}^{\mathfrak{S}_n} \underline{1}_{\mathfrak{S}_\lambda}\right)(g)$$

(by the definition of $\operatorname{sgn}_{\mathfrak{S}_n} * \left(\operatorname{Ind}_{\mathfrak{S}_\lambda}^{\mathfrak{S}_n} \underline{1}_{\mathfrak{S}_\lambda}\right)$). Since (by the definition of $\operatorname{Ind}_{\mathfrak{S}_\lambda}^{\mathfrak{S}_n} \underline{1}_{\mathfrak{S}_\lambda}$) we have

$$\left(\operatorname{Ind}_{\mathfrak{S}_\lambda}^{\mathfrak{S}_n} \underline{1}_{\mathfrak{S}_\lambda}\right)(g) = \frac{1}{|\mathfrak{S}_\lambda|} \sum_{\substack{k \in \mathfrak{S}_n: \\ kgk^{-1} \in \mathfrak{S}_\lambda}} \underbrace{\underline{1}_{\mathfrak{S}_\lambda}(kgk^{-1})}_{=1} = \frac{1}{|\mathfrak{S}_\lambda|} \sum_{\substack{k \in \mathfrak{S}_n: \\ kgk^{-1} \in \mathfrak{S}_\lambda}} 1,$$

this rewrites as

$$\left(\widetilde{\omega}_{\mathbb{Z}}\left(\operatorname{Ind}_{\mathfrak{S}_\lambda}^{\mathfrak{S}_n} \underline{1}_{\mathfrak{S}_\lambda}\right)\right)(g) = \operatorname{sgn}(g) \cdot \underbrace{\left(\operatorname{Ind}_{\mathfrak{S}_\lambda}^{\mathfrak{S}_n} \underline{1}_{\mathfrak{S}_\lambda}\right)(g)}_{= \frac{1}{|\mathfrak{S}_\lambda|} \sum_{\substack{k \in \mathfrak{S}_n: \\ kgk^{-1} \in \mathfrak{S}_\lambda}} 1} = \operatorname{sgn}(g) \cdot \frac{1}{|\mathfrak{S}_\lambda|} \sum_{\substack{k \in \mathfrak{S}_n: \\ kgk^{-1} \in \mathfrak{S}_\lambda}} 1.$$

Now, comparing

$$\left(\widetilde{\omega}_{\mathbb{Z}} \circ \operatorname{ch}^{-1}\right)(h_\lambda) = \widetilde{\omega}_{\mathbb{Z}}\left(\underbrace{\operatorname{ch}^{-1}(h_\lambda)}_{=\operatorname{Ind}_{\mathfrak{S}_\lambda}^{\mathfrak{S}_n} \underline{1}_{\mathfrak{S}_\lambda}}\right) = \widetilde{\omega}_{\mathbb{Z}}\left(\operatorname{Ind}_{\mathfrak{S}_\lambda}^{\mathfrak{S}_n} \underline{1}_{\mathfrak{S}_\lambda}\right) = \operatorname{Ind}_{\mathfrak{S}_\lambda}^{\mathfrak{S}_n} \operatorname{sgn}_{\mathfrak{S}_\lambda}$$

with

$$\left(\operatorname{ch}^{-1} \circ \omega\right)(h_\lambda) = \operatorname{ch}^{-1}\left(\underbrace{\omega(h_\lambda)}_{=e_\lambda}\right) = \operatorname{ch}^{-1}(e_\lambda) = \operatorname{Ind}_{\mathfrak{S}_\lambda}^{\mathfrak{S}_n} \operatorname{sgn}_{\mathfrak{S}_\lambda},$$

we obtain $\left(\widetilde{\omega}_{\mathbb{Z}} \circ \operatorname{ch}^{-1}\right)(h_\lambda) = \left(\operatorname{ch}^{-1} \circ \omega\right)(h_\lambda)$.

So we have shown that $\left(\widetilde{\omega}_{\mathbb{Z}} \circ \operatorname{ch}^{-1}\right)(h_\lambda) = \left(\operatorname{ch}^{-1} \circ \omega\right)(h_\lambda)$ for every partition λ. Thus, $\widetilde{\omega}_{\mathbb{Z}} \circ \operatorname{ch}^{-1} = \operatorname{ch}^{-1} \circ \omega$ (since the h_λ form a $\mathbb{Z}$-module basis of Λ). Hence, $\operatorname{ch}^{-1} \circ \omega = \widetilde{\omega}_{\mathbb{Z}} \circ \operatorname{ch}^{-1}$ so that $\operatorname{ch}^{-1} \circ \omega \circ \operatorname{ch} = \widetilde{\omega}_{\mathbb{Z}}$ and $\omega \circ \operatorname{ch} = \operatorname{ch} \circ \widetilde{\omega}_{\mathbb{Z}}$.

We thus have $\operatorname{ch} \circ \widetilde{\omega}_{\mathbb{Z}} = \omega \circ \operatorname{ch}$ as maps $A(\mathfrak{S}) \to \Lambda$. In other words, the map $\widetilde{\omega}_{\mathbb{Z}}$ corresponds under ch to the involution ω on Λ. In other words, the involution on $A(\mathfrak{S})$ defined in Theorem 4.4.1(b) corresponds under ch to the involution ω on Λ (since $\widetilde{\omega}_{\mathbb{Z}}$ is precisely the involution on $A(\mathfrak{S})$ defined in Theorem 4.4.1(b)). This completes the proof of Theorem 4.4.1(b).

Step 5: The structure-preserving properties of $\widetilde{\omega}_{\mathbb{Z}}$.

Now, it only remains to show that the involution on $A(\mathfrak{S})$ defined in Theorem 4.4.1(b) is a nontrivial PSH-automorphism of $A(\mathfrak{S})$.

Recall that ch is a PSH-isomorphism. Hence, its inverse ch^{-1} also is a PSH-isomorphism.

Proposition 2.4.3(f) shows that $\omega : \Lambda \to \Lambda$ is a Hopf automorphism. Moreover, the map ω is clearly graded. Let Σ denote the PSH-basis $\{s_\lambda \mid \lambda \in \operatorname{Par}\}$ of Λ. Then, ω restricts to a bijection $\Sigma \to \Sigma$ [881]. As a consequence, $\omega(\Sigma) \subset \Sigma$, so that $\omega(\mathbb{N}\Sigma) = \mathbb{N}\underbrace{\omega(\Sigma)}_{\subset\Sigma} \subset \mathbb{N}\Sigma$. Hence, ω is a PSH-morphism $\Lambda \to \Lambda$ (since ω is a graded Hopf algebra morphism), and thus a PSH-isomorphism $\Lambda \to \Lambda$ (since ω is an isomorphism and restricts to a bijection $\Sigma \to \Sigma$).

Now, $\operatorname{ch}^{-1} \circ \omega \circ \operatorname{ch}$ is the composition of three PSH-isomorphisms (since ch^{-1}, ω and ch are PSH-isomorphisms), therefore a PSH-isomorphism itself. In other words, $\widetilde{\omega}_{\mathbb{Z}}$ is a PSH-isomorphism (because $\widetilde{\omega}_{\mathbb{Z}} = \operatorname{ch}^{-1} \circ \omega \circ \operatorname{ch}$). Thus, $\widetilde{\omega}_{\mathbb{Z}}$ is a PSH-automorphism.

Since $\omega(h_2) = e_2 \neq h_2$, we have $\omega \neq \operatorname{id}$ and therefore $\widetilde{\omega}_{\mathbb{Z}} \neq \operatorname{id}$ (since $\widetilde{\omega}_{\mathbb{Z}} = \operatorname{ch}^{-1} \circ \omega \circ \operatorname{ch}$). In other words, $\widetilde{\omega}_{\mathbb{Z}}$ is nontrivial.

So we know that $\widetilde{\omega}_{\mathbb{Z}}$ is a nontrivial PSH-automorphism of $A(\mathfrak{S})$. In other words, the involution on $A(\mathfrak{S})$ defined in Theorem 4.4.1(b) is a nontrivial PSH-automorphism of $A(\mathfrak{S})$ (since $\widetilde{\omega}_{\mathbb{Z}}$ is precisely the involution on $A(\mathfrak{S})$ defined in Theorem 4.4.1(b)). This completes the solution of Exercise 4.4.4.

Remark. There are some alternative ways to solve parts of this exercise.

For example, in order to prove that $\widetilde{\omega}_{\mathbb{Z}} \circ \operatorname{ch}^{-1} = \operatorname{ch}^{-1} \circ \omega$, we showed that $\left(\widetilde{\omega}_{\mathbb{Z}} \circ \operatorname{ch}^{-1}\right)(h_\lambda) = \left(\operatorname{ch}^{-1} \circ \omega\right)(h_\lambda)$ for every partition λ. Instead of doing this, it is possible to prove that $\widetilde{\omega}_{\mathbb{Z}} \circ \operatorname{ch}^{-1} = \operatorname{ch}^{-1} \circ \omega$ by showing that

Compared with

$$\begin{aligned}\left(\operatorname{Ind}_{\mathfrak{S}_\lambda}^{\mathfrak{S}_n} \operatorname{sgn}_{\mathfrak{S}_\lambda}\right)(g) &= \frac{1}{|\mathfrak{S}_\lambda|} \sum_{\substack{k \in \mathfrak{S}_n : \\ kgk^{-1} \in \mathfrak{S}_\lambda}} \underbrace{\operatorname{sgn}_{\mathfrak{S}_\lambda}\left(kgk^{-1}\right)}_{=\operatorname{sgn}\left(kgk^{-1}\right)} \qquad \left(\text{by the definition of } \operatorname{Ind}_{\mathfrak{S}_\lambda}^{\mathfrak{S}_n} \operatorname{sgn}_{\mathfrak{S}_\lambda}\right) \\ &= \frac{1}{|\mathfrak{S}_\lambda|} \sum_{\substack{k \in \mathfrak{S}_n : \\ kgk^{-1} \in \mathfrak{S}_\lambda}} \underbrace{\operatorname{sgn}\left(kgk^{-1}\right)}_{\substack{=\operatorname{sgn}(k) \cdot \operatorname{sgn}(g) \cdot (\operatorname{sgn}(k))^{-1} \\ =\operatorname{sgn}(g) = \operatorname{sgn}(g) \cdot 1}} = \frac{1}{|\mathfrak{S}_\lambda|} \sum_{\substack{k \in \mathfrak{S}_n : \\ kgk^{-1} \in \mathfrak{S}_\lambda}} \operatorname{sgn}(g) \cdot 1 = \operatorname{sgn}(g) \cdot \frac{1}{|\mathfrak{S}_\lambda|} \sum_{\substack{k \in \mathfrak{S}_n : \\ kgk^{-1} \in \mathfrak{S}_\lambda}} 1,\end{aligned}$$

this yields $\left(\widetilde{\omega}_{\mathbb{Z}}\left(\operatorname{Ind}_{\mathfrak{S}_\lambda}^{\mathfrak{S}_n} \underline{1}_{\mathfrak{S}_\lambda}\right)\right)(g) = \left(\operatorname{Ind}_{\mathfrak{S}_\lambda}^{\mathfrak{S}_n} \operatorname{sgn}_{\mathfrak{S}_\lambda}\right)(g)$.

Now, forget that we fixed g. We thus have shown that $\left(\widetilde{\omega}_{\mathbb{Z}}\left(\operatorname{Ind}_{\mathfrak{S}_\lambda}^{\mathfrak{S}_n} \underline{1}_{\mathfrak{S}_\lambda}\right)\right)(g) = \left(\operatorname{Ind}_{\mathfrak{S}_\lambda}^{\mathfrak{S}_n} \operatorname{sgn}_{\mathfrak{S}_\lambda}\right)(g)$ for every $g \in \mathfrak{S}_n$. In other words, $\widetilde{\omega}_{\mathbb{Z}}\left(\operatorname{Ind}_{\mathfrak{S}_\lambda}^{\mathfrak{S}_n} \underline{1}_{\mathfrak{S}_\lambda}\right) = \operatorname{Ind}_{\mathfrak{S}_\lambda}^{\mathfrak{S}_n} \operatorname{sgn}_{\mathfrak{S}_\lambda}$, qed.

[881] This follows from Lemma 13.99.3.

$\left(\widetilde{\omega}_{\mathbb{Z}} \circ \operatorname{ch}^{-1}\right)(p_\lambda) = \left(\operatorname{ch}^{-1} \circ \omega\right)(p_\lambda)$ for every partition λ. This, however, requires a little technicality (the p_λ do not span Λ as a $\mathbb{Z}$-module, only as a $\mathbb{Q}$-module).

We showed that $\widetilde{\omega}_{\mathbb{Z}} \circ \operatorname{ch}^{-1} = \operatorname{ch}^{-1} \circ \omega$ and used this to conclude that $\widetilde{\omega}_{\mathbb{Z}}$ is a PSH-automorphism of $A(\mathfrak{S})$. An alternative argument proceeds the other way round: The map $\widetilde{\omega}_{\mathbb{Z}}$ is $\mathbb{Z}$-linear and graded and is easily seen to be self-adjoint with respect to the inner product on $A(\mathfrak{S})$. It also is a coalgebra morphism, as $\operatorname{Res}^{\mathfrak{S}_n}_{\mathfrak{S}_i \times \mathfrak{S}_j} \omega_n(f) = (\omega_i \otimes \omega_j)\left(\operatorname{Res}^{\mathfrak{S}_n}_{\mathfrak{S}_i \times \mathfrak{S}_j} f\right)$ for all $n = i + j$ and all $f \in R(\mathfrak{S}_n)$. Hence, this map $\widetilde{\omega}_{\mathbb{Z}}$ also is an algebra morphism (since it is self-adjoint, and $A(\mathfrak{S})$ is self-dual), hence a bialgebra morphism and thus a Hopf morphism (by Corollary 1.4.27). It also sends irreducible characters to irreducible characters (by (13.118.6)), and thus restricts to a bijection $\{\chi^\lambda\} \to \{\chi^\lambda\}$. Hence, $\widetilde{\omega}_{\mathbb{Z}}$ is a PSH-automorphism of $A(\mathfrak{S})$. Every $n \geq 0$ satisfies

$$\begin{aligned}
\left(\widetilde{\omega}_{\mathbb{Z}} \circ \operatorname{ch}^{-1}\right)(h_n) &= \widetilde{\omega}_{\mathbb{Z}}\left(\underbrace{\operatorname{ch}^{-1}(h_n)}_{=\underline{1}_{\mathfrak{S}_n}}\right) = \widetilde{\omega}_{\mathbb{Z}}\left(\underline{1}_{\mathfrak{S}_n}\right) = \omega_n\left(\underline{1}_{\mathfrak{S}_n}\right) = \operatorname{sgn}_{\mathfrak{S}_n} = \operatorname{ch}^{-1}(e_n) \\
&= \operatorname{ch}^{-1}(\omega(h_n)) \qquad \text{(since Proposition 2.4.3(b) yields } e_n = \omega(h_n)) \\
&= \left(\operatorname{ch}^{-1} \circ \omega\right)(h_n).
\end{aligned}$$

Since $\widetilde{\omega}_{\mathbb{Z}} \circ \operatorname{ch}^{-1}$ and $\operatorname{ch}^{-1} \circ \omega$ are algebra morphisms whereas the h_n generate Λ, this yields $\widetilde{\omega}_{\mathbb{Z}} \circ \operatorname{ch}^{-1} = \operatorname{ch}^{-1} \circ \omega$.

13.119. **Solution to Exercise 4.4.5.** *Solution to Exercise 4.4.5.* It is known that two permutations in $\mathfrak{S}_n$ have the same cycle type if and only if they are conjugate. In other words, if g and h are two permutations in $\mathfrak{S}_n$, then we have the following logical equivalence:

$$(g \text{ and } h \text{ have the same cycle type}) \iff (g \text{ and } h \text{ are conjugate}). \tag{13.119.1}$$

(a) Let $f \in R_{\mathbb{C}}(\mathfrak{S}_n)$.

If $\sigma \in \mathfrak{S}_n$, then the cycle type of σ is a partition of n. In other words, we have $\operatorname{type}\sigma \in \operatorname{Par}_n$ for each $\sigma \in \mathfrak{S}_n$ (since $\operatorname{type}\sigma$ denotes the cycle type of σ, while Par_n denotes the set of all partitions of n).

For each partition λ, we define a positive integer z_λ as in Proposition 2.5.15.

Recall the following fact from finite group theory:

> *Claim 1:* Let G be any finite group. Let $g \in G$ be any element. Let C_g denote the conjugacy class of g. Let Z_g denote the centralizer of g. Then,

$$|C_g| = \frac{|G|}{|Z_g|}. \tag{13.119.2}$$

Applying Claim 1 to the symmetric group $\mathfrak{S}_n$, we quickly arrive at the following:

> *Claim 2:* For each $\lambda \in \operatorname{Par}_n$, we have
> $$(\text{the number of all } \sigma \in \mathfrak{S}_n \text{ satisfying } \operatorname{type}\sigma = \lambda) = \frac{n!}{z_\lambda}.$$

[*Proof of Claim 2:* Let $\lambda \in \operatorname{Par}_n$. Thus, λ is a partition of n. Hence, there exists a permutation in $\mathfrak{S}_n$ that has cycle type λ. [882] Choose such a permutation, and denote it by g.

Let C_g denote the conjugacy class of g in $\mathfrak{S}_n$. Let Z_g denote the centralizer of g in $\mathfrak{S}_n$.

We have $|\lambda| = n$ (since λ is a partition of n). Hence, from Remark 2.5.16, we know that z_λ is the size of the $\mathfrak{S}_n$-centralizer subgroup for a permutation having cycle type λ. Thus, z_λ is the size of the $\mathfrak{S}_n$-centralizer subgroup of g (since $g \in \mathfrak{S}_n$ is a permutation having cycle type λ). In other words, $z_\lambda = |Z_g|$ (since Z_g is the $\mathfrak{S}_n$-centralizer subgroup of g).

[882] Indeed, we can obtain such a permutation as follows: Let $\ell = \ell(\lambda)$, so that $\lambda_1 + \lambda_2 + \cdots + \lambda_\ell = |\lambda| = n$. Partition the n-element set $\{1, 2, \ldots, n\}$ into ℓ disjoint subsets $K_1, K_2, \ldots, K_\ell$ of sizes $\lambda_1, \lambda_2, \ldots, \lambda_\ell$, respectively. For each $i \in \{1, 2, \ldots, \ell\}$, pick an arbitrary permutation in $\mathfrak{S}_n$ that fixes each element of $\{1, 2, \ldots, n\} \setminus K_i$ while consisting of a single λ_i-cycle on the set K_i. (Thus, c_i cycles through the λ_i elements of K_i in some order while leaving all remaining elements of $\{1, 2, \ldots, n\}$ unchanged.) Then, the composition $c_1 c_2 \cdots c_\ell$ of these cycles is a permutation in $\mathfrak{S}_n$ that has cycle type λ.

On the other hand, the definition of C_g yields

$$\begin{aligned} C_g &= (\text{the conjugacy class of } g) \\ &= \left\{ h \in \mathfrak{S}_n \mid \underbrace{h \text{ is conjugate to } g}_{\substack{\iff (g \text{ and } h \text{ are conjugate}) \\ \iff (g \text{ and } h \text{ have the same cycle type}) \\ (\text{by the equivalence (13.119.1)})}} \right\} = \left\{ h \in \mathfrak{S}_n \mid \underbrace{g \text{ and } h \text{ have the same cycle type}}_{\substack{\iff (h \text{ has cycle type } \lambda) \\ (\text{since } g \text{ has cycle type } \lambda)}} \right\} \\ &= \left\{ h \in \mathfrak{S}_n \mid \underbrace{h \text{ has cycle type } \lambda}_{\iff (\operatorname{type} h = \lambda)} \right\} = \{ h \in \mathfrak{S}_n \mid \operatorname{type} h = \lambda \} = \{ \sigma \in \mathfrak{S}_n \mid \operatorname{type} \sigma = \lambda \} \end{aligned}$$

(here, we have renamed the index h as σ). Thus,

$$|C_g| = |\{ \sigma \in \mathfrak{S}_n \mid \operatorname{type} \sigma = \lambda \}| = (\text{the number of all } \sigma \in \mathfrak{S}_n \text{ satisfying } \operatorname{type} \sigma = \lambda).$$

Comparing this with

$$\begin{aligned} |C_g| &= \frac{|\mathfrak{S}_n|}{|Z_g|} && (\text{by Claim 1, applied to } G = \mathfrak{S}_n) \\ &= \frac{n!}{z_\lambda} && (\text{since } |\mathfrak{S}_n| = n! \text{ and } |Z_g| = z_\lambda), \end{aligned}$$

we obtain

$$(\text{the number of all } \sigma \in \mathfrak{S}_n \text{ satisfying } \operatorname{type} \sigma = \lambda) = \frac{n!}{z_\lambda}.$$

This proves Claim 2.]

For each partition λ of n, there exists a permutation in $\mathfrak{S}_n$ that has cycle type λ. Choose such a permutation, and denote it by g_λ. Every $h \in \mathfrak{S}_n$ satisfies

$$f(h) = f(g_{\operatorname{type} h}). \tag{13.119.3}$$

[*Proof of* (13.119.3)*:* Let $h \in \mathfrak{S}_n$. Hence, $\operatorname{type} h \in \operatorname{Par}_n$ (since $\operatorname{type} \sigma \in \operatorname{Par}_n$ for each $\sigma \in \mathfrak{S}_n$). Therefore, the permutation $g_{\operatorname{type} h} \in \mathfrak{S}_n$ is well-defined.

The definition of $\operatorname{type} h$ shows that the permutation h has cycle type $\operatorname{type} h$. On the other hand, the permutation $g_{\operatorname{type} h}$ also has cycle type $\operatorname{type} h$ (by the definition of $g_{\operatorname{type} h}$). Hence, the two permutations $g_{\operatorname{type} h}$ and h in $\mathfrak{S}_n$ have the same cycle type (namely, $\operatorname{type} h$). But (13.119.1) (applied to $g_{\operatorname{type} h}$ instead of g) shows that we have the equivalence

$$(g_{\operatorname{type} h} \text{ and } h \text{ have the same cycle type}) \iff (g_{\operatorname{type} h} \text{ and } h \text{ are conjugate}).$$

Hence, $g_{\operatorname{type} h}$ and h are conjugate (since $g_{\operatorname{type} h}$ and h have the same cycle type).

But we have $f \in R_{\mathbb{C}}(\mathfrak{S}_n)$. In other words, f is a class function of $\mathfrak{S}_n$ (since $R_{\mathbb{C}}(\mathfrak{S}_n)$ is the space of all class functions $\mathfrak{S}_n \to \mathbb{C}$ of $\mathfrak{S}_n$). In other words, f is a map $\mathfrak{S}_n \to \mathbb{C}$ that is constant on $\mathfrak{S}_n$-conjugacy classes (because this is what it means to be a class function of $\mathfrak{S}_n$). Thus, in particular, f is constant on $\mathfrak{S}_n$-conjugacy classes. In other words, if x and y are two conjugate elements of $\mathfrak{S}_n$, then $f(x) = f(y)$. Applying this to $x = g_{\operatorname{type} h}$ and $y = h$, we obtain $f(g_{\operatorname{type} h}) = f(h)$ (since the elements $g_{\operatorname{type} h}$ and h of $\mathfrak{S}_n$ are conjugate). This proves (13.119.3).]

Recall that $\operatorname{type}\sigma \in \operatorname{Par}_n$ for each $\sigma \in \mathfrak{S}_n$. Thus, we can split the sum $\sum_{\sigma\in\mathfrak{S}_n} f(\sigma)\, p_{\operatorname{type}\sigma}$ according to the value of $\operatorname{type}\sigma$. We thus obtain

$$\begin{aligned}
\underbrace{\sum_{\sigma\in\mathfrak{S}_n}}_{\substack{=\sum_{\lambda\in\operatorname{Par}_n}\ \sum_{\substack{\sigma\in\mathfrak{S}_n;\\ \operatorname{type}\sigma=\lambda}}\\ \text{(since } \operatorname{type}\sigma\in\operatorname{Par}_n\\ \text{for each } \sigma\in\mathfrak{S}_n)}} f(\sigma)\, p_{\operatorname{type}\sigma} &= \sum_{\lambda\in\operatorname{Par}_n}\ \sum_{\substack{\sigma\in\mathfrak{S}_n;\\ \operatorname{type}\sigma=\lambda}} \underbrace{f(\sigma)}_{\substack{=f(g_{\operatorname{type}\sigma})\\ \text{(by (13.119.3), applied to } h=\sigma)}} p_{\operatorname{type}\sigma} \\
&= \sum_{\lambda\in\operatorname{Par}_n}\ \sum_{\substack{\sigma\in\mathfrak{S}_n;\\ \operatorname{type}\sigma=\lambda}} \underbrace{f(g_{\operatorname{type}\sigma})}_{\substack{=f(g_\lambda)\\ \text{(since } \operatorname{type}\sigma=\lambda)}} \underbrace{p_{\operatorname{type}\sigma}}_{\substack{=p_\lambda\\ \text{(since } \operatorname{type}\sigma=\lambda)}} \\
&= \sum_{\lambda\in\operatorname{Par}_n} \underbrace{\sum_{\substack{\sigma\in\mathfrak{S}_n;\\ \operatorname{type}\sigma=\lambda}} f(g_\lambda)\, p_\lambda}_{=(\text{the number of all } \sigma\in\mathfrak{S}_n \text{ satisfying } \operatorname{type}\sigma=\lambda)\cdot f(g_\lambda) p_\lambda} \\
&= \sum_{\lambda\in\operatorname{Par}_n} \underbrace{(\text{the number of all } \sigma \in \mathfrak{S}_n \text{ satisfying } \operatorname{type}\sigma = \lambda)}_{\substack{=\frac{n!}{z_\lambda}\\ \text{(by Claim 2)}}} \cdot f(g_\lambda)\, p_\lambda \\
&= \sum_{\lambda\in\operatorname{Par}_n} \frac{n!}{z_\lambda} \cdot f(g_\lambda)\, p_\lambda .
\end{aligned}$$

Multiplying both sides of this equality by $\dfrac{1}{n!}$, we obtain

$$\begin{aligned}
\frac{1}{n!} \sum_{\sigma\in\mathfrak{S}_n} f(\sigma)\, p_{\operatorname{type}\sigma} = \frac{1}{n!} \cdot \sum_{\lambda\in\operatorname{Par}_n} \frac{n!}{z_\lambda} \cdot f(g_\lambda)\, p_\lambda &= \sum_{\lambda\in\operatorname{Par}_n} \underbrace{\frac{1}{n!} \cdot \frac{n!}{z_\lambda} \cdot f(g_\lambda)\, p_\lambda}_{=f(g_\lambda)\cdot \frac{p_\lambda}{z_\lambda}} \\
&= \sum_{\lambda\in\operatorname{Par}_n} f(g_\lambda) \cdot \frac{p_\lambda}{z_\lambda}. \qquad (13.119.4)
\end{aligned}$$

We shall now show that

$$f = \sum_{\lambda\in\operatorname{Par}_n} f(g_\lambda)\, \underline{1}_\lambda. \qquad (13.119.5)$$

[*Proof of* (13.119.5)*:* Let $h \in \mathfrak{S}_n$. Then, $\operatorname{type} h$ is the cycle type of h (by the definition of $\operatorname{type} h$). In other words, h has cycle type $\operatorname{type} h$.

Recall that $\operatorname{type}\sigma \in \operatorname{Par}_n$ for each $\sigma \in \mathfrak{S}_n$. Applying this to $\sigma = h$, we obtain $\operatorname{type} h \in \operatorname{Par}_n$.

For each partition $\lambda \in \operatorname{Par}_n$, we have

$$\underline{1}_\lambda(h) = \begin{cases} 1, & \text{if } h \text{ has cycle type } \lambda; \\ 0, & \text{otherwise} \end{cases} \qquad (13.119.6)$$

(since $\underline{1}_\lambda$ was defined as the characteristic function for the $\mathfrak{S}_n$-conjugacy class of permutations of cycle type λ). Applying this to $\lambda = \operatorname{type} h$, we obtain

$$\underline{1}_{\operatorname{type} h}(h) = \begin{cases} 1, & \text{if } h \text{ has cycle type } \operatorname{type} h; \\ 0, & \text{otherwise} \end{cases} = 1 \qquad (13.119.7)$$

(since h has cycle type $\operatorname{type} h$). On the other hand, if $\lambda \in \operatorname{Par}_n$ is a partition such that $\lambda \neq \operatorname{type} h$, then h does not have cycle type λ [883], and therefore we have

$$\begin{aligned} \underline{1}_\lambda(h) &= \begin{cases} 1, & \text{if } h \text{ has cycle type } \lambda; \\ 0, & \text{otherwise} \end{cases} \qquad (\text{by } (13.119.6)) \\ &= 0 \qquad (\text{since } h \text{ does not have cycle type } \lambda). \end{aligned} \tag{13.119.8}$$

Now,

$$\begin{aligned} \left(\sum_{\lambda \in \operatorname{Par}_n} f(g_\lambda) \underline{1}_\lambda\right)(h) &= \sum_{\lambda \in \operatorname{Par}_n} f(g_\lambda) \underline{1}_\lambda(h) \\ &= f\left(g_{\operatorname{type} h}\right) \underbrace{\underline{1}_{\operatorname{type} h}(h)}_{\substack{=1 \\ (\text{by } (13.119.7))}} + \sum_{\substack{\lambda \in \operatorname{Par}_n; \\ \lambda \neq \operatorname{type} h}} f(g_\lambda) \underbrace{\underline{1}_\lambda(h)}_{\substack{=0 \\ (\text{by } (13.119.8))}} \\ &\qquad \left(\begin{array}{c} \text{here, we have split off the addend for } \lambda = \operatorname{type} h \text{ from the sum,} \\ \text{since } \operatorname{type} h \in \operatorname{Par}_n \end{array}\right) \\ &= f\left(g_{\operatorname{type} h}\right) + \underbrace{\sum_{\substack{\lambda \in \operatorname{Par}_n; \\ \lambda \neq \operatorname{type} h}} f(g_\lambda) 0}_{=0} = f\left(g_{\operatorname{type} h}\right) = f(h) \qquad (\text{by } (13.119.3)). \end{aligned}$$

Forget that we fixed h. We thus have shown that $\left(\sum_{\lambda \in \operatorname{Par}_n} f(g_\lambda) \underline{1}_\lambda\right)(h) = f(h)$ for each $h \in \mathfrak{S}_n$. In other words, $\sum_{\lambda \in \operatorname{Par}_n} f(g_\lambda) \underline{1}_\lambda = f$ (since both $\sum_{\lambda \in \operatorname{Par}_n} f(g_\lambda) \underline{1}_\lambda$ and f are maps from $\mathfrak{S}_n$ to $\mathbb{C}$). This proves (13.119.5).]

One of the claims of Theorem 4.4.1(a) is the formula

$$\operatorname{ch}\left(\underline{1}_\lambda\right) = \frac{p_\lambda}{z_\lambda} \qquad \text{for every } \lambda \in \operatorname{Par}_n. \tag{13.119.9}$$

Now, applying the map $\operatorname{ch} : A_{\mathbb{C}} \to \Lambda_{\mathbb{C}}$ to both sides of the equality (13.119.5), we find

$$\begin{aligned} \operatorname{ch}(f) &= \operatorname{ch}\left(\sum_{\lambda \in \operatorname{Par}_n} f(g_\lambda) \underline{1}_\lambda\right) = \sum_{\lambda \in \operatorname{Par}_n} f(g_\lambda) \underbrace{\operatorname{ch}\left(\underline{1}_\lambda\right)}_{\substack{= \frac{p_\lambda}{z_\lambda} \\ (\text{by } (13.119.9))}} \\ &\qquad (\text{since the map } \operatorname{ch} : A_{\mathbb{C}} \to \Lambda_{\mathbb{C}} \text{ is } \mathbb{C}\text{-linear}) \\ &= \sum_{\lambda \in \operatorname{Par}_n} f(g_\lambda) \cdot \frac{p_\lambda}{z_\lambda} = \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_n} f(\sigma) p_{\operatorname{type} \sigma} \end{aligned}$$

(by (13.119.4)). This solves Exercise 4.4.5(a).

(b) It is easy to see that every $k \in \mathfrak{S}_n$ and $\sigma \in \mathfrak{S}_n$ satisfy

$$p_{\operatorname{type} \sigma} = p_{\operatorname{type}(k\sigma k^{-1})}. \tag{13.119.10}$$

[*Proof of* (13.119.10)*:* Let $k \in \mathfrak{S}_n$ and $\sigma \in \mathfrak{S}_n$. Then, (13.119.1) (applied to $g = \sigma$ and $h = k\sigma k^{-1}$) yields the equivalence

$$\left(\sigma \text{ and } k\sigma k^{-1} \text{ have the same cycle type}\right) \iff \left(\sigma \text{ and } k\sigma k^{-1} \text{ are conjugate}\right).$$

Hence, σ and $k\sigma k^{-1}$ have the same cycle type (since σ and $k\sigma k^{-1}$ are conjugate). In other words, $\operatorname{type} \sigma = \operatorname{type}\left(k\sigma k^{-1}\right)$. Hence, $p_{\operatorname{type} \sigma} = p_{\operatorname{type}(k\sigma k^{-1})}$. This proves (13.119.10).]

[883] *Proof.* Assume the contrary. Thus, h has cycle type λ. In other words, the cycle type of h is λ. In other words, $\operatorname{type} h$ is λ (since $\operatorname{type} h$ is the cycle type of h). In other words, $\operatorname{type} h = \lambda$. But this contradicts $\lambda \neq \operatorname{type} h$. This contradiction shows that our assumption was false. Qed.

Let $f \in R_{\mathbb{C}}(H)$. Applying (4.1.4) to $G = \mathfrak{S}_n$, we obtain

$$\left(\operatorname{Ind}_H^{\mathfrak{S}_n} f\right)(g) = \frac{1}{|H|} \sum_{\substack{k \in \mathfrak{S}_n: \\ kgk^{-1} \in H}} f\left(kgk^{-1}\right) \tag{13.119.11}$$

for all $g \in \mathfrak{S}_n$. Furthermore, Exercise 4.1.1(a) (applied to $G = \mathfrak{S}_n$) shows that $\operatorname{Ind}_H^{\mathfrak{S}_n} f$ is a class function on $\mathfrak{S}_n$, hence belongs to $R_{\mathbb{C}}(\mathfrak{S}_n)$. Hence, Exercise 4.4.5(a) (applied to $\operatorname{Ind}_H^{\mathfrak{S}_n} f$ instead of f) yields

$$\begin{aligned}
\operatorname{ch}\left(\operatorname{Ind}_H^{\mathfrak{S}_n} f\right) &= \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_n} \underbrace{\left(\operatorname{Ind}_H^{\mathfrak{S}_n} f\right)(\sigma)}_{\substack{= \frac{1}{|H|} \sum_{\substack{k \in \mathfrak{S}_n: \\ k\sigma k^{-1} \in H}} f\left(k\sigma k^{-1}\right) \\ \text{(by (13.119.11), applied to } g = \sigma)}} p_{\operatorname{type}\sigma} \\
&= \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_n} \frac{1}{|H|} \sum_{\substack{k \in \mathfrak{S}_n: \\ k\sigma k^{-1} \in H}} f\left(k\sigma k^{-1}\right) \underbrace{p_{\operatorname{type}\sigma}}_{\substack{= p_{\operatorname{type}(k\sigma k^{-1})} \\ \text{(by (13.119.10))}}} \\
&= \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_n} \frac{1}{|H|} \sum_{\substack{k \in \mathfrak{S}_n: \\ k\sigma k^{-1} \in H}} f\left(k\sigma k^{-1}\right) p_{\operatorname{type}(k\sigma k^{-1})} \\
&= \frac{1}{n!} \underbrace{\sum_{\sigma \in \mathfrak{S}_n} \sum_{\substack{k \in \mathfrak{S}_n: \\ k\sigma k^{-1} \in H}}}_{= \sum_{k \in \mathfrak{S}_n} \sum_{\substack{\sigma \in \mathfrak{S}_n: \\ k\sigma k^{-1} \in H}}} \frac{1}{|H|} f\left(k\sigma k^{-1}\right) p_{\operatorname{type}(k\sigma k^{-1})} \\
&= \frac{1}{n!} \sum_{k \in \mathfrak{S}_n} \underbrace{\sum_{\substack{\sigma \in \mathfrak{S}_n: \\ k\sigma k^{-1} \in H}} \frac{1}{|H|} f\left(k\sigma k^{-1}\right) p_{\operatorname{type}(k\sigma k^{-1})}}_{\substack{= \sum_{\substack{h \in \mathfrak{S}_n: \\ h \in H}} \frac{1}{|H|} f(h) p_{\operatorname{type}h} \\ \text{(here, we have substituted } h \text{ for } k\sigma k^{-1} \text{ in the sum,} \\ \text{since the map } \mathfrak{S}_n \to \mathfrak{S}_n,\ \sigma \mapsto k\sigma k^{-1} \text{ is a bijection)}}} \\
&= \frac{1}{n!} \sum_{k \in \mathfrak{S}_n} \underbrace{\sum_{\substack{h \in \mathfrak{S}_n: \\ h \in H}}}_{\substack{= \sum_{h \in H} \\ \text{(since } H \subset \mathfrak{S}_n)}} \frac{1}{|H|} f(h) p_{\operatorname{type}h} = \frac{1}{n!} \underbrace{\sum_{k \in \mathfrak{S}_n} \sum_{h \in H} \frac{1}{|H|} f(h) p_{\operatorname{type}h}}_{= |\mathfrak{S}_n| \cdot \sum_{h \in H} \frac{1}{|H|} f(h) p_{\operatorname{type}h}} \\
&= \frac{1}{n!} \cdot \underbrace{|\mathfrak{S}_n|}_{= n!} \cdot \sum_{h \in H} \frac{1}{|H|} f(h) p_{\operatorname{type}h} = \underbrace{\frac{1}{n!} \cdot n!}_{= 1} \cdot \sum_{h \in H} \frac{1}{|H|} f(h) p_{\operatorname{type}h} \\
&= \sum_{h \in H} \frac{1}{|H|} f(h) p_{\operatorname{type}h} = \frac{1}{|H|} \sum_{h \in H} f(h) p_{\operatorname{type}h}.
\end{aligned}$$

This solves Exercise 4.4.5(b).

13.120. **Solution to Exercise 4.4.6.** *Solution to Exercise 4.4.6.* We need an auxiliary observation first. If G is a finite group, then a class function $\chi \in R_{\mathbb{C}}(G)$ of G will be called *integral* if every $g \in G$ satisfies $\chi(g) \in \mathbb{Z}$. Then, if G and H are two groups and if ϕ and ψ are two integral class functions of G and H, respectively, then

$$\phi \otimes \psi \text{ is an integral class function of } G \times H. \tag{13.120.1}$$

(This is a consequence of the fact that $(\phi \otimes \psi)(g,h) = \phi(g)\psi(h)$ for all $(g,h) \in G \times H$.) Moreover, if H is a subgroup of a group G and if χ is an integral class function of H, then

$$\operatorname{Ind}_H^G \chi \text{ is an integral class function of } G. \tag{13.120.2}$$

(*Proof of* (13.120.2)*:* Let J be a system of coset representatives for $H \backslash G$, so that $G = \bigsqcup_{j \in J} Hj$. Then, Exercise 4.1.1(a) (applied to $f = \chi$) shows that $\operatorname{Ind}_H^G \chi$ is a class function on G. Also, Exercise 4.1.1(b) (applied to $f = \chi$) yields

$$\left(\operatorname{Ind}_H^G \chi\right)(g) = \sum_{\substack{j \in J: \\ jgj^{-1} \in H}} \underbrace{\chi\left(jgj^{-1}\right)}_{\substack{\in \mathbb{Z} \\ \text{(since } \chi \text{ is an integral} \\ \text{class function)}}} \in \mathbb{Z} \qquad \text{for all } g \in G.$$

In other words, $\operatorname{Ind}_H^G \chi$ is an integral class function on G. This proves (13.120.2).)

Now let $A = A(\mathfrak{S})$. For every $a \in A$, every $n \geq 0$ and every $g \in \mathfrak{S}_n$, we define $a(g)$ to mean the value at g of the n-th homogeneous component of a. Let $\widetilde{A}$ denote the subset of A formed by all $a \in A$ such that:

$$(\text{for every } n \geq 0 \text{ and every } g \in \mathfrak{S}_n, \text{ we have } a(g) \in \mathbb{Z}). \tag{13.120.3}$$

It is clear that $\widetilde{A}$ is a graded $\mathbb{Z}$-submodule of A and contains $1 = \underline{1}_{\mathfrak{S}_0}$.

We are now going to show that $\widetilde{A}$ is closed under multiplication. In order to do so, it is clearly enough to prove that if $a \in \widetilde{A}$ is homogeneous of degree n and $b \in \widetilde{A}$ is homogeneous of degree m, then $ab \in \widetilde{A}$. But this is now easy: Since $a \in \widetilde{A}$, we know that a is an integral class function on $\mathfrak{S}_n$. Similarly, b is an integral class function on $\mathfrak{S}_m$. Hence, (13.120.1) shows that $a \otimes b$ is an integral class function on $\mathfrak{S}_n \times \mathfrak{S}_m$. Thus, $\operatorname{Ind}_{\mathfrak{S}_n \times \mathfrak{S}_m}^{\mathfrak{S}_{n+m}} (a \otimes b)$ is an integral class function on $\mathfrak{S}_{n+m}$ (by (13.120.2)). In other words, ab is an integral class function on $\mathfrak{S}_{n+m}$ (since $ab = \operatorname{ind}_{n,m}^{n+m}(a \otimes b) = \operatorname{Ind}_{\mathfrak{S}_n \times \mathfrak{S}_m}^{\mathfrak{S}_{n+m}}(a \otimes b)$). In other words, $ab \in \widetilde{A}$. This proves that $\widetilde{A}$ is closed under multiplication. Hence, $\widetilde{A}$ is a subring of A (since $\widetilde{A}$ is a $\mathbb{Z}$-submodule of A and contains $1 = \underline{1}_{\mathfrak{S}_0}$). Thus, $\operatorname{ch}\left(\widetilde{A}\right)$ (where ch is defined as in Theorem 4.4.1) is a subring of Λ. For every $n \geq 1$, we have $\underline{1}_{\mathfrak{S}_n} \in \widetilde{A}$ (by the definitions), so that $\operatorname{ch}\left(\underline{1}_{\mathfrak{S}_n}\right) \in \operatorname{ch}\left(\widetilde{A}\right)$. Since $\operatorname{ch}\left(\underline{1}_{\mathfrak{S}_n}\right) = h_n$, this rewrites as $h_n \in \operatorname{ch}\left(\widetilde{A}\right)$. Thus, the subring $\operatorname{ch}\left(\widetilde{A}\right)$ of Λ contains h_n for all $n \geq 1$. Hence, $\operatorname{ch}\left(\widetilde{A}\right) = \Lambda$ (because the h_n generate Λ as a ring). Therefore, $\widetilde{A} = A$ (since ch is an isomorphism). If we recall how $\widetilde{A}$ was defined, we thus see that every $a \in A$ satisfies (13.120.3).

Now, fix $n \geq 0$ and $g \in \mathfrak{S}_n$ and a finite-dimensional $\mathbb{C}\mathfrak{S}_n$-module V. Then, $\chi_V \in R(\mathfrak{S}_n) \subset A$. Thus, (13.120.3) (which holds for every $a \in A$, as we now know) yields that $\chi_V(g) \in \mathbb{Z}$, which solves part (a) of the problem.

Note that a different way to solve part (a) would be by showing that Λ is contained in the $\mathbb{Z}$-submodule of $\Lambda_{\mathbb{Q}}$ spanned by the $\dfrac{p_\lambda}{z_\lambda}$ for all partitions λ. This yields (by taking preimages under ch) that A is contained in the $\mathbb{Z}$-submodule of $R_{\mathbb{C}}(\mathfrak{S})$ generated by the $\underline{1}_\lambda$. This way, we wouldn't have to show that $\widetilde{A}$ is closed under multiplication; instead, we would obtain $\widetilde{A} = A$ by noticing that all $\underline{1}_\lambda$ are integral class functions.

(b) This can be done using the Noether-Deuring theorem (in fact, it is easy to show that there are two $\mathbb{C}\mathfrak{S}_n$-modules U and U' defined over $\mathbb{Q}$ satisfying $U' \oplus V \cong U$, and then the Noether-Deuring theorem allows to "pull back" V to a $\mathbb{Q}\mathfrak{S}_n$-module as well, showing that V is also defined over $\mathbb{Q}$). We are omitting this argument because it is somewhat technical and not very enlightening. (The "right" approach, in my opinion, is to construct the required $\mathbb{Q}\mathfrak{S}_n$-module W explicitly; this is done, e.g., in [73, §7, between Proposition 1 and Lemma 3] or in [115, Corollaire 2.2.26]. Of course, this does not have much to do with what we are doing in our notes.)

13.121. **Solution to Exercise 4.4.8.** *Solution to Exercise 4.4.8.* (a) Let G be any group. If U_1 and U_2 are two $\mathbb{C}G$-modules, then the character $\chi_{U_1 \boxtimes U_2}$ of the inner tensor product $U_1 \boxtimes U_2$ of U_1 and U_2 is given by

$$\chi_{U_1 \boxtimes U_2}(g) = \chi_{U_1}(g) \chi_{U_2}(g) \qquad \text{for all } g \in G. \tag{13.121.1}$$

(This is just a restatement of (13.118.1) using our notation $U_1 \boxtimes U_2$ for what, in (13.118.1), was called $U_1 \otimes U_2$.)

Define a map $* : R_{\mathbb{C}}(G) \times R_{\mathbb{C}}(G) \to R_{\mathbb{C}}(G)$, which will be written in infix notation (that is, we will write $a * b$ instead of $*(a, b)$), by setting

$$(a * b)(g) = a(g) b(g) \qquad \text{for any } a \in R_{\mathbb{C}}(G), b \in R_{\mathbb{C}}(G) \text{ and } g \in G.$$

(This notation $a*b$ generalizes the notation $\operatorname{sgn}_{\mathfrak{S}_n} *f$ used in Theorem 4.4.1.) The map $*$ is clearly $\mathbb{C}$-bilinear.

Notice that

$$\chi_{U_1} * \chi_{U_2} = \chi_{U_1 \boxtimes U_2} \qquad \text{for any two } \mathbb{C}G\text{-modules } U_1 \text{ and } U_2 \tag{13.121.2}$$

[884]

Let $R_{\mathbb{Q}}(G)$ denote the set of class functions $G \to \mathbb{Q}$. This is clearly a subset of $R_{\mathbb{C}}(G)$. In general, it is not true that $R(G) \subset R_{\mathbb{Q}}(G)$, but we will see that this holds for $G = \mathfrak{S}_n$ for any $n \in \mathbb{N}$.

It is clear that $a * b \in R_{\mathbb{Q}}(G)$ for any $a \in R_{\mathbb{Q}}(G)$ and $b \in R_{\mathbb{Q}}(G)$.

Now, forget that we fixed G. We have thus introduced a map $* : R_{\mathbb{C}}(G) \times R_{\mathbb{C}}(G) \to R_{\mathbb{C}}(G)$ for every group G, and we have proved some properties of this map.

Let now $n \in \mathbb{N}$. Exercise 4.4.6(a) yields that $\chi_V(g) \in \mathbb{Z} \subset \mathbb{Q}$ for every $g \in \mathfrak{S}_n$ and every finite-dimensional $\mathbb{C}\mathfrak{S}_n$-module V. Hence, χ_V is a map from $\mathfrak{S}_n$ to $\mathbb{Q}$ for every finite-dimensional $\mathbb{C}\mathfrak{S}_n$-module V. In other words, $\chi_V \in R_{\mathbb{Q}}(\mathfrak{S}_n)$ for every finite-dimensional $\mathbb{C}\mathfrak{S}_n$-module V. Thus, $R(\mathfrak{S}_n) \subset R_{\mathbb{Q}}(\mathfrak{S}_n)$ (since $R(\mathfrak{S}_n)$ is the $\mathbb{Z}$-module generated by the χ_V for V ranging over the irreducible $\mathbb{C}\mathfrak{S}_n$-modules).

Every element of $R_{\mathbb{Q}}(\mathfrak{S}_n)$ is a class function $\mathfrak{S}_n \to \mathbb{Q}$, and thus a $\mathbb{Q}$-linear combination of the $\underline{1}_\lambda$ for $\lambda \in \operatorname{Par}_n$ (because these functions $\underline{1}_\lambda$ are the indicator functions for the conjugacy classes of $\mathfrak{S}_n$). Thus, the $\mathbb{Q}$-module $R_{\mathbb{Q}}(\mathfrak{S}_n)$ is spanned by the $\underline{1}_\lambda$ for $\lambda \in \operatorname{Par}_n$.

As in Theorem 4.4.1, we extend the PSH-isomorphism $\operatorname{ch} : A \to \Lambda$ to a $\mathbb{C}$-Hopf algebra isomorphism $A_{\mathbb{C}} \to \Lambda_{\mathbb{C}}$; we shall denote the latter isomorphism by ch as well. Now, let us notice that

$$\operatorname{ch}(R_{\mathbb{Q}}(\mathfrak{S}_n)) \subset \Lambda_{\mathbb{Q}}. \tag{13.121.3}$$

[885] Furthermore,

$$\operatorname{ch}(a * b) = (\operatorname{ch} a) * (\operatorname{ch} b) \qquad \text{for any } a \in R_{\mathbb{Q}}(\mathfrak{S}_n) \text{ and } b \in R_{\mathbb{Q}}(\mathfrak{S}_n). \tag{13.121.4}$$

[886]

[884] *Proof of (13.121.2):* Let U_1 and U_2 be two $\mathbb{C}G$-modules. Then, every $g \in G$ satisfies

$$\begin{aligned}(\chi_{U_1} * \chi_{U_2})(g) &= \chi_{U_1}(g) \chi_{U_2}(g) && (\text{by the definition of } *)\\ &= \chi_{U_1 \boxtimes U_2}(g) && (\text{by (13.121.1)}).\end{aligned}$$

In other words, $\chi_{U_1} * \chi_{U_2} = \chi_{U_1 \boxtimes U_2}$. This proves (13.121.2).

[885] *Proof of (13.121.3):* It is clearly enough to show that $\operatorname{ch}(\underline{1}_\lambda) \in \Lambda_{\mathbb{Q}}$ for every $\lambda \in \operatorname{Par}_n$ (because the $\mathbb{Q}$-module $R_{\mathbb{Q}}(\mathfrak{S}_n)$ is spanned by the $\underline{1}_\lambda$ for $\lambda \in \operatorname{Par}_n$). But this follows from the fact that $\operatorname{ch}(\underline{1}_\lambda) = \frac{p_\lambda}{z_\lambda}$ for every $\lambda \in \operatorname{Par}_n$ (this is part of Theorem 4.4.1(a)). Thus, (13.121.3) is proven.

[886] *Proof of (13.121.4):* Let $a \in R_{\mathbb{Q}}(\mathfrak{S}_n)$ and $b \in R_{\mathbb{Q}}(\mathfrak{S}_n)$. The equality (13.121.4) is clearly $\mathbb{Q}$-linear in each of a and b. Hence, in proving this equality, we can WLOG assume that a and b are two of the functions $\underline{1}_\lambda$ for $\lambda \in \operatorname{Par}_n$ (because the $\mathbb{Q}$-module $R_{\mathbb{Q}}(\mathfrak{S}_n)$ is spanned by the $\underline{1}_\lambda$ for $\lambda \in \operatorname{Par}_n$). In other words, we can WLOG assume that $a = \underline{1}_\lambda$ and $b = \underline{1}_\mu$ for some $\lambda \in \operatorname{Par}_n$ and $\mu \in \operatorname{Par}_n$. Assume this, and consider these λ and μ.

Recall that $\underline{1}_\lambda$ is the indicator function for the set of all permutations in $\mathfrak{S}_n$ having cycle type λ, and that $\underline{1}_\mu$ is the indicator function for the set of all permutations in $\mathfrak{S}_n$ having cycle type μ. Hence, every $g \in \mathfrak{S}_n$ satisfies

$$\underline{1}_\lambda(g) \cdot \underline{1}_\mu(g) = \delta_{\lambda,\mu} \cdot \underline{1}_\lambda(g)$$

(because both sides of this equality vanish if g does not have cycle type λ, and also vanish if $\lambda \neq \mu$, but in the remaining case are both equal to 1). Thus, every $g \in \mathfrak{S}_n$ satisfies

$$\begin{aligned}\left(\underline{1}_\lambda * \underline{1}_\mu\right)(g) &= \underline{1}_\lambda(g) \cdot \underline{1}_\mu(g) && \left(\text{by the definition of } \underline{1}_\lambda * \underline{1}_\mu\right)\\ &= \delta_{\lambda,\mu} \cdot \underline{1}_\lambda(g) = (\delta_{\lambda,\mu}\underline{1}_\lambda)(g).\end{aligned}$$

Now, let U_1 and U_2 be two $\mathbb{C}\mathfrak{S}_n$-modules. Then, $\chi_{U_1} \in R(\mathfrak{S}_n) \subset R_{\mathbb{Q}}(\mathfrak{S}_n)$ and similarly $\chi_{U_2} \in R_{\mathbb{Q}}(\mathfrak{S}_n)$, so that (13.121.4) (applied to $a = \chi_{U_1}$ and $b = \chi_{U_2}$) yields $\operatorname{ch}(\chi_{U_1} * \chi_{U_2}) = \operatorname{ch}(\chi_{U_1}) * \operatorname{ch}(\chi_{U_2})$. Due to (13.121.2), this rewrites as $\operatorname{ch}(\chi_{U_1 \boxtimes U_2}) = \operatorname{ch}(\chi_{U_1}) * \operatorname{ch}(\chi_{U_2})$. This solves Exercise 4.4.8(a).

(c) Let us first recall that $p_\lambda * p_\mu = \delta_{\lambda,\mu} z_\lambda p_\lambda$ for any two partitions λ and μ. Thus, $p_\lambda * p_\mu = 0$ whenever $\lambda \neq \mu$. This yields, in particular, that $p_\lambda * p_\mu = 0$ whenever $|\lambda| \neq |\mu|$. In other words, for any two distinct integers n and m, we have $p_\lambda * p_\mu = 0$ for every $\lambda \in \operatorname{Par}_n$ and every $\mu \in \operatorname{Par}_m$. This yields that, for any two distinct integers n and m, we have $a * b = 0$ for every $a \in \Lambda_n$ and $b \in \Lambda_m$ (because a is a $\mathbb{Q}$-linear combination of the p_λ with $\lambda \in \operatorname{Par}_n$, whereas b is a $\mathbb{Q}$-linear combination of the p_μ with $\mu \in \operatorname{Par}_m$). In particular, for any two distinct integers n and m, we have

$$s_\mu * s_\nu = 0 \qquad \text{for any } \nu \in \operatorname{Par}_n \text{ and } \mu \in \operatorname{Par}_m$$

(because $s_\mu \in \Lambda_m$ and $s_\nu \in \Lambda_n$). In other words,

$$s_\mu * s_\nu = 0 \qquad \text{for any partitions } \mu \text{ and } \nu \text{ satisfying } |\mu| \neq |\nu| . \tag{13.121.5}$$

Now, let μ and ν be two partitions. We need to prove that $s_\mu * s_\nu \in \sum_{\lambda \in \operatorname{Par}} \mathbb{N} s_\lambda$. This is obvious (because of (13.121.5)) in the case when $|\mu| \neq |\nu|$, so we can WLOG assume that $|\mu| = |\nu|$. Assume this, and let $n = |\mu| = |\nu|$. Consider the two irreducible characters χ^μ and χ^ν of $\mathbb{C}\mathfrak{S}_n$ defined as in Theorem 4.4.1(a). Then, $\chi^\mu = \chi_{U_1}$ and $\chi^\nu = \chi_{U_2}$ for two $\mathbb{C}\mathfrak{S}_n$-modules U_1 and U_2. Consider these U_1 and U_2. We have $\operatorname{ch}\left(\underbrace{\chi_{U_1}}_{=\chi^\mu}\right) = \operatorname{ch}(\chi^\mu) = s_\mu$ (by Theorem 4.4.1(a)) and similarly $\operatorname{ch}(\chi_{U_2}) = s_\nu$. But $U_1 \boxtimes U_2$ (being a $\mathbb{C}\mathfrak{S}_n$-module) must be a direct sum of finitely many irreducible $\mathbb{C}\mathfrak{S}_n$-modules, and thus the character $\chi_{U_1 \boxtimes U_2}$ is the sum of finitely many irreducible characters of $\mathbb{C}\mathfrak{S}_n$. Since the irreducible characters of $\mathbb{C}\mathfrak{S}_n$ are the χ^λ for $\lambda \in \operatorname{Par}_n$, this shows that $\chi_{U_1 \boxtimes U_2}$ is the sum of finitely many χ^λ. In other words, $\chi_{U_1 \boxtimes U_2} \in \sum_{\lambda \in \operatorname{Par}} \mathbb{N} \chi^\lambda$. Applying the map ch to both sides of this relation, we obtain

$$\begin{aligned} \operatorname{ch}(\chi_{U_1 \boxtimes U_2}) \in \operatorname{ch}\left(\sum_{\lambda \in \operatorname{Par}} \mathbb{N} \chi^\lambda\right) &= \sum_{\lambda \in \operatorname{Par}} \mathbb{N} \underbrace{\operatorname{ch}\left(\chi^\lambda\right)}_{\substack{=s_\lambda \\ \text{(by Theorem 4.4.1(a))}}} \qquad \text{(since ch is } \mathbb{Z}\text{-linear)} \\ &= \sum_{\lambda \in \operatorname{Par}} \mathbb{N} s_\lambda . \end{aligned}$$

Hence, $\underline{1}_\lambda * \underline{1}_\mu = \delta_{\lambda,\mu} \underline{1}_\lambda$. Applying the map ch to this equality, we obtain

$$\operatorname{ch}\left(\underline{1}_\lambda * \underline{1}_\mu\right) = \operatorname{ch}\left(\delta_{\lambda,\mu} \underline{1}_\lambda\right) = \delta_{\lambda,\mu} \underbrace{\operatorname{ch}\left(\underline{1}_\lambda\right)}_{\substack{=\frac{p_\lambda}{z_\lambda} \\ \text{(by Theorem 4.4.1(a))}}} = \delta_{\lambda,\mu} \frac{p_\lambda}{z_\lambda} = \delta_{\lambda,\mu} z_\lambda^{-1} p_\lambda .$$

Compared with

$$\begin{aligned} &\underbrace{\left(\operatorname{ch}\left(\underline{1}_\lambda\right)\right)}_{\substack{=\frac{p_\lambda}{z_\lambda} \\ \text{(by Theorem 4.4.1(a))}}} * \underbrace{\left(\operatorname{ch}\left(\underline{1}_\mu\right)\right)}_{\substack{=\frac{p_\mu}{z_\mu} \\ \text{(by Theorem 4.4.1(a))}}} \\ &= \frac{p_\lambda}{z_\lambda} * \frac{p_\mu}{z_\mu} = z_\lambda^{-1} z_\mu^{-1} \underbrace{p_\lambda * p_\mu}_{\substack{=\delta_{\lambda,\mu} z_\lambda p_\lambda \\ \text{(by the definition of } *)}} = z_\lambda^{-1} z_\mu^{-1} \delta_{\lambda,\mu} z_\lambda p_\lambda = \underbrace{\delta_{\lambda,\mu} z_\mu^{-1}}_{\substack{=\delta_{\lambda,\mu} z_\lambda^{-1} \\ \text{(because the two sides of this} \\ \text{equality are equal if } \lambda=\mu \text{, and} \\ \text{vanish otherwise)}}} p_\lambda = \delta_{\lambda,\mu} z_\lambda^{-1} p_\lambda , \end{aligned}$$

this yields $\operatorname{ch}\left(\underline{1}_\lambda * \underline{1}_\mu\right) = \left(\operatorname{ch}\left(\underline{1}_\lambda\right)\right) * \left(\operatorname{ch}\left(\underline{1}_\mu\right)\right)$. This rewrites as $\operatorname{ch}(a * b) = (\operatorname{ch} a) * (\operatorname{ch} b)$ (since $a = \underline{1}_\lambda$ and $b = \underline{1}_\mu$). Thus, (13.121.4) is proven.

Since

$$\begin{aligned}\operatorname{ch}\left(\chi_{U_1 \boxtimes U_2}\right) &= \underbrace{\operatorname{ch}\left(\chi_{U_1}\right)}_{=s_\mu} * \underbrace{\operatorname{ch}\left(\chi_{U_2}\right)}_{=s_\nu} \qquad \text{(by Exercise 4.4.8(a))} \\ &= s_\mu * s_\nu,\end{aligned}$$

this rewrites as $s_\mu * s_\nu \in \sum_{\lambda \in \operatorname{Par}} \mathbb{N} s_\lambda$. This solves Exercise 4.4.8(c).

(b) *Alternative solution of Exercise 2.9.4(h).* We need to prove that $f * g \in \Lambda$ for any $f \in \Lambda$ and $g \in \Lambda$. Since the binary operation $*$ is $\mathbb{Z}$-bilinear, it is clearly enough to prove that $s_\mu * s_\nu \in \Lambda$ for any partitions μ and ν (since $(s_\lambda)_{\lambda \in \operatorname{Par}}$ is a $\mathbb{Z}$-basis of Λ). But this follows from the fact that

$$\begin{aligned}s_\mu * s_\nu &\in \sum_{\lambda \in \operatorname{Par}} \mathbb{N} s_\lambda \qquad \text{(by Exercise 4.4.8(c))} \\ &\subset \Lambda\end{aligned}$$

for any partitions μ and ν. Thus, Exercise 2.9.4 (h) is solved again.

13.122. **Solution to Exercise 4.4.9.** *Solution to Exercise 4.4.9.* Let us introduce a fundamental construction. If X is any set, then we let $\mathfrak{S}_X$ denote the symmetric group on X (that is, the group of all permutations of X). For any two sets X and Y and any permutations $\sigma \in \mathfrak{S}_X$ and $\tau \in \mathfrak{S}_Y$, we define a permutation $\sigma \times \tau$ of $X \times Y$ by setting

$$(\sigma \times \tau)((x, y)) = (\sigma(x), \tau(y)) \qquad \text{for every } x \in X \text{ and } y \in Y.$$

Thus, for any two sets X and Y, we can define a map

$$\begin{aligned}\operatorname{cross} : \mathfrak{S}_X \times \mathfrak{S}_Y &\to \mathfrak{S}_{X \times Y}, \\ (\sigma, \tau) &\mapsto \sigma \times \tau.\end{aligned}$$

We notice a few properties of this map (whose simple proof we leave to the reader):

Lemma 13.122.1. *Let X and Y be two sets.*

(a) *The map* cross $: \mathfrak{S}_X \times \mathfrak{S}_Y \to \mathfrak{S}_{X \times Y}$ *is a group homomorphism.*
(b) *If the sets X and Y are nonempty, then the map* cross $: \mathfrak{S}_X \times \mathfrak{S}_Y \to \mathfrak{S}_{X \times Y}$ *is injective.*

The following deeper fact provides the first hint of a combinatorial interpretation of $\boxdot$:

Lemma 13.122.2. *Let X and Y be two finite sets. Let $\sigma \in \mathfrak{S}_X$ and $\tau \in \mathfrak{S}_Y$ be two permutations. Let λ, μ and κ be the cycle types of the permutations $\sigma \in \mathfrak{S}_X$, $\tau \in \mathfrak{S}_Y$ and $\sigma \times \tau \in \mathfrak{S}_{X \times Y}$. Then,*

$$p_\lambda \boxdot p_\mu = p_\kappa.$$

Proof of Lemma 13.122.2. The partition λ is the cycle type of the permutation σ. In other words, $\lambda_1, \lambda_2, \ldots, \lambda_{\ell(\lambda)}$ are the lengths of the cycles of the permutation σ. Let $C_1, C_2, \ldots, C_{\ell(\lambda)}$ denote these cycles, labelled in such a way that each C_i has length λ_i. Clearly, $\left(C_1, C_2, \ldots, C_{\ell(\lambda)}\right)$ is a set partition of the set X; thus, $X = \bigsqcup_{i=1}^{\ell(\lambda)} C_i$.

The partition μ is the cycle type of the permutation τ. In other words, $\mu_1, \mu_2, \ldots, \mu_{\ell(\mu)}$ are the lengths of the cycles of the permutation τ. Let $D_1, D_2, \ldots, D_{\ell(\mu)}$ denote these cycles, labelled in such a way that each D_j has length μ_j. Clearly, $\left(D_1, D_2, \ldots, D_{\ell(\mu)}\right)$ is a set partition of the set Y; thus, $Y = \bigsqcup_{j=1}^{\ell(\mu)} D_j$.

Since $X = \bigsqcup_{i=1}^{\ell(\lambda)} C_i$ and $Y = \bigsqcup_{j=1}^{\ell(\mu)} D_j$, we have

$$X \times Y = \left(\bigsqcup_{i=1}^{\ell(\lambda)} C_i\right) \times \left(\bigsqcup_{j=1}^{\ell(\mu)} D_j\right) = \bigsqcup_{i=1}^{\ell(\lambda)} \bigsqcup_{j=1}^{\ell(\mu)} (C_i \times D_j).$$

In particular, the sets $C_i \times D_j$ for (i, j) ranging over all $(i, j) \in \{1, 2, \ldots, \ell(\lambda)\} \times \{1, 2, \ldots, \ell(\mu)\}$ are disjoint.

We now make the following claim:

> *Claim A:* Let $i \in \{1, 2, \ldots, \ell(\lambda)\}$ and $j \in \{1, 2, \ldots, \ell(\mu)\}$. Then, the subset $C_i \times D_j$ of $X \times Y$ is the union of $\gcd(\lambda_i, \mu_j)$ disjoint cycles of the permutation $\sigma \times \tau$, and each of these cycles has length $\operatorname{lcm}(\lambda_i, \mu_j)$.

Proof of Claim A: We know that C_i is a cycle of σ; thus, $\sigma(C_i) \subset C_i$. Similarly, $\tau(D_j) \subset D_j$. Now, $(\sigma \times \tau)(C_i \times D_j) = \underbrace{\sigma(C_i)}_{\subset C_i} \times \underbrace{\tau(D_j)}_{\subset D_j} \subset C_i \times D_j$. Hence, the permutation $\sigma \times \tau$ restricts to a permutation of the subset $C_i \times D_j$ of $X \times Y$.

Let us first show that, for every $h \in C_i \times D_j$, the cycle of $\sigma \times \tau$ containing h is a subset of $C_i \times D_j$ and has length $\operatorname{lcm}(\lambda_i, \mu_j)$.

Indeed, fix $h \in C_i \times D_j$. Then, all of the elements $h, (\sigma \times \tau)(h), (\sigma \times \tau)^2(h), (\sigma \times \tau)^3(h), \ldots$ belong to $C_i \times D_j$ [887]. Hence, the cycle of $\sigma \times \tau$ containing h is a subset of $C_i \times D_j$ (because this cycle consists of these very elements $h, (\sigma \times \tau)(h), (\sigma \times \tau)^2(h), (\sigma \times \tau)^3(h), \ldots$). The length of this cycle is the smallest positive integer N such that $(\sigma \times \tau)^N(h) = h$. We shall now prove that this smallest positive integer is $\operatorname{lcm}(\lambda_i, \mu_j)$.

Indeed, let us write h in the form $h = (c, d)$ for some $c \in C_i$ and $d \in D_j$. Then, the element c belongs to a cycle of σ which has length λ_i (namely, C_i). Hence, the sequence $c, \sigma(c), \sigma^2(c), \sigma^3(c), \ldots$ repeats every λ_i elements (and not more frequently). Thus, for any $N \in \mathbb{N}$, we have the following equivalence of statements:

$$\left(\sigma^N(c) = c\right) \iff (\lambda_i \mid N).$$

Similarly, for any $N \in \mathbb{N}$, we have the following equivalence of statements:

$$\left(\tau^N(d) = d\right) \iff (\mu_j \mid N).$$

Now, for any $N \in \mathbb{N}$, we have the following equivalence of statements:

$$\begin{aligned}\left((\sigma \times \tau)^N(h) = h\right) &\iff \left(\underbrace{(\sigma \times \tau)^N((c,d))}_{=(\sigma^N(c), \tau^N(d))} = (c,d)\right) \qquad (\text{since } h = (c,d)) \\ &\iff \left(\left(\sigma^N(c), \tau^N(d)\right) = (c,d)\right) \iff \left(\underbrace{\left(\sigma^N(c) = c\right)}_{\iff (\lambda_i \mid N)} \text{ and } \underbrace{\left(\tau^N(d) = d\right)}_{\iff (\mu_j \mid N)}\right) \\ &\iff (\lambda_i \mid N \text{ and } \mu_j \mid N) \iff (\operatorname{lcm}(\lambda_i, \mu_j) \mid N).\end{aligned}$$

Thus, the smallest positive integer N such that $(\sigma \times \tau)^N(h) = h$ is $\operatorname{lcm}(\lambda_i, \mu_j)$. In other words, the length of the cycle of $\sigma \times \tau$ containing h is $\operatorname{lcm}(\lambda_i, \mu_j)$ (since the length of the cycle of $\sigma \times \tau$ containing h is the smallest positive integer N such that $(\sigma \times \tau)^N(h) = h$).

Now, let us forget that we fixed h. We thus have proven that, for every $h \in C_i \times D_j$,

$$\text{the cycle of } \sigma \times \tau \text{ containing } h \text{ is a subset of } C_i \times D_j \text{ and has length } \operatorname{lcm}(\lambda_i, \mu_j). \tag{13.122.1}$$

These cycles (for h ranging over all $C_i \times D_j$) clearly cover the set $C_i \times D_j$ (because each $h \in C_i \times D_j$ is contained in its corresponding cycle). Thus, $C_i \times D_j$ is the union of several cycles of the permutation $\sigma \times \tau$, and each of these cycles has length $\operatorname{lcm}(\lambda_i, \mu_j)$. Since any two cycles of $\sigma \times \tau$ are either disjoint or identical, we can get rid of redundant cycles in this union, and thus obtain the following conclusion: $C_i \times D_j$ is the union of several disjoint cycles of the permutation $\sigma \times \tau$, and each of these cycles has length $\operatorname{lcm}(\lambda_i, \mu_j)$. In order to complete the proof of Claim A, it thus remains only to show that the number of these cycles is $\gcd(\lambda_i, \mu_j)$. But this is easy: These cycles are all disjoint, and cover a set of size $\lambda_i \mu_j$ (in fact, they cover the set $C_i \times D_j$, which has size $|C_i \times D_j| = \underbrace{|C_i|}_{=\lambda_i} \cdot \underbrace{|D_j|}_{=\mu_j} = \lambda_i \mu_j$); since they have length $\operatorname{lcm}(\lambda_i, \mu_j)$ each, their number must be $\dfrac{\lambda_i \mu_j}{\operatorname{lcm}(\lambda_i, \mu_j)} = \gcd(\lambda_i, \mu_j)$. Thus, the proof of Claim A is complete.

We shall now continue our proof of Lemma 13.122.2.

[887] since the permutation $\sigma \times \tau$ restricts to a permutation of the subset $C_i \times D_j$ of $X \times Y$

The partition κ is the cycle type of the permutation $\sigma \times \tau$. In other words, $\kappa_1, \kappa_2, \ldots, \kappa_{\ell(\kappa)}$ are the lengths of the cycles of the permutation $\sigma \times \tau$. Hence,

$$\prod_{u=1}^{\ell(\kappa)} p_{\kappa_u} = \prod_{E \text{ is a cycle of } \sigma\times\tau} p_{|E|},$$

where $|E|$ denotes the length of any cycle E. But every cycle E of $\sigma \times \tau$ must satisfy $E \subset C_i \times D_j$ (where we regard E as a set) for precisely one $(i,j) \in \{1,2,\ldots,\ell(\lambda)\} \times \{1,2,\ldots,\ell(\mu)\}$ [888]. Hence,

$$\prod_{E \text{ is a cycle of } \sigma\times\tau} p_{|E|} = \prod_{i=1}^{\ell(\lambda)} \prod_{j=1}^{\ell(\mu)} \prod_{\substack{E \text{ is a cycle of } \sigma\times\tau;\\ E\subset C_i\times D_j}} p_{|E|}.$$

But every $(i,j) \in \{1,2,\ldots,\ell(\lambda)\} \times \{1,2,\ldots,\ell(\mu)\}$ satisfies

$$\prod_{\substack{E \text{ is a cycle of } \sigma\times\tau;\\ E\subset C_i\times D_j}} p_{|E|} = p_{\operatorname{lcm}(\lambda_i,\mu_j)}^{\gcd(\lambda_i,\mu_j)}$$

[889]. Now, the definition of p_κ yields

$$p_\kappa = p_{\kappa_1} p_{\kappa_2} \cdots p_{\kappa_{\ell(\kappa)}} = \prod_{u=1}^{\ell(\kappa)} p_{\kappa_u} = \prod_{E \text{ is a cycle of } \sigma\times\tau} p_{|E|} = \prod_{i=1}^{\ell(\lambda)} \prod_{j=1}^{\ell(\mu)} \underbrace{\prod_{\substack{E \text{ is a cycle of } \sigma\times\tau;\\ E\subset C_i\times D_j}} p_{|E|}}_{=p_{\operatorname{lcm}(\lambda_i,\mu_j)}^{\gcd(\lambda_i,\mu_j)}}$$

$$= \prod_{i=1}^{\ell(\lambda)} \prod_{j=1}^{\ell(\mu)} p_{\operatorname{lcm}(\lambda_i,\mu_j)}^{\gcd(\lambda_i,\mu_j)} = p_\lambda \boxdot p_\mu \qquad \left(\text{since } p_\lambda \boxdot p_\mu \text{ is defined to be } \prod_{i=1}^{\ell(\lambda)} \prod_{j=1}^{\ell(\mu)} p_{\operatorname{lcm}(\lambda_i,\mu_j)}^{\gcd(\lambda_i,\mu_j)}\right).$$

This proves Lemma 13.122.2. □

(a) We already know that the operation $\boxdot$ is $\mathbb{Q}$-bilinear. Thus, we only need to prove that the binary operation $\boxdot$ is commutative and associative and has unity p_1.

We shall only prove the associativity of $\boxdot$ (since the other two properties can be proven similarly). In other words, we shall prove that $(u \boxdot v) \boxdot w = u \boxdot (v \boxdot w)$ for any three elements u, v and w of $\Lambda_{\mathbb{Q}}$.

So let u, v and w be three elements of $\Lambda_{\mathbb{Q}}$. We need to prove the equality $(u \boxdot v) \boxdot w = u \boxdot (v \boxdot w)$. Since this equality is $\mathbb{Q}$-linear in each of u, v and w (this is because the operation $\boxdot$ is $\mathbb{Q}$-bilinear), we can WLOG assume that u, v and w belong to the basis $(p_\lambda)_{\lambda\in\mathrm{Par}}$ of the $\mathbb{Q}$-vector space $\Lambda_{\mathbb{Q}}$. Assume this. Then, there exist three partitions λ, μ and ν such that $u = p_\lambda$, $v = p_\mu$ and $w = p_\nu$. Consider these λ, μ and ν.

There exist a finite set A and a permutation $\alpha \in \mathfrak{S}_A$ such that λ is the cycle type of α. Consider these A and α.

[888] *Proof.* Let E be a cycle of $\sigma \times \tau$. Then, E is nonempty, so that there exists an element h of E. Fix such an h. We have $h \in E \subset X \times Y = \bigsqcup_{i=1}^{\ell(\lambda)} \bigsqcup_{j=1}^{\ell(\mu)} (C_i \times D_j)$. Hence, there exists some $(i,j) \in \{1,2,\ldots,\ell(\lambda)\} \times \{1,2,\ldots,\ell(\mu)\}$ such that $h \in C_i \times D_j$. Consider this (i,j). Then, E is the cycle of $\sigma \times \tau$ containing h. Hence, (13.122.1) shows that E is a subset of $C_i \times D_j$ and has length $\operatorname{lcm}(\lambda_i, \mu_j)$. So we know that $E \subset C_i \times D_j$.

So we have found a pair $(i,j) \in \{1,2,\ldots,\ell(\lambda)\} \times \{1,2,\ldots,\ell(\mu)\}$ such that $E \subset C_i \times D_j$. There clearly cannot be two distinct such pairs (i,j) (since the sets $C_i \times D_j$ for (i,j) ranging over all $(i,j) \in \{1,2,\ldots,\ell(\lambda)\} \times \{1,2,\ldots,\ell(\mu)\}$ are disjoint), and so this pair (i,j) is unique.

[889] *Proof.* Let $(i,j) \in \{1,2,\ldots,\ell(\lambda)\} \times \{1,2,\ldots,\ell(\mu)\}$. Claim A yields that the subset $C_i \times D_j$ of $X \times Y$ is the union of $\gcd(\lambda_i, \mu_j)$ disjoint cycles of the permutation $\sigma \times \tau$, and each of these cycles has length $\operatorname{lcm}(\lambda_i, \mu_j)$. Obviously, these $\gcd(\lambda_i, \mu_j)$ disjoint cycles are exactly all the cycles E of $\sigma \times \tau$ which satisfy $E \subset C_i \times D_j$. Thus, there are precisely $\gcd(\lambda_i, \mu_j)$ cycles E of $\sigma \times \tau$ which satisfy $E \subset C_i \times D_j$, and each of these cycles E has length $\operatorname{lcm}(\lambda_i, \mu_j)$. Hence,

$$\prod_{\substack{E \text{ is a cycle of } \sigma\times\tau;\\ E\subset C_i\times D_j}} \underbrace{p_{|E|}}_{\substack{=p_{\operatorname{lcm}(\lambda_i,\mu_j)}\\ \text{(since } E \text{ has length } \operatorname{lcm}(\lambda_i,\mu_j),\\ \text{that is, we have } |E|=\operatorname{lcm}(\lambda_i,\mu_j))}} = \prod_{\substack{E \text{ is a cycle of } \sigma\times\tau;\\ E\subset C_i\times D_j}} p_{\operatorname{lcm}(\lambda_i,\mu_j)} = p_{\operatorname{lcm}(\lambda_i,\mu_j)}^{\gcd(\lambda_i,\mu_j)}$$

(since there are precisely $\gcd(\lambda_i, \mu_j)$ cycles E of $\sigma \times \tau$ which satisfy $E \subset C_i \times D_j$), qed.

There exist a finite set B and a permutation $\beta \in \mathfrak{S}_B$ such that μ is the cycle type of β. Consider these B and β.

There exist a finite set C and a permutation $\gamma \in \mathfrak{S}_C$ such that ν is the cycle type of γ. Consider these C and γ.

For every permutation π of any finite set X, we let $\operatorname{type}\pi$ denote the cycle type of π. Lemma 13.122.2 (applied to $X = A$, $Y = B$, $\sigma = \alpha$, $\tau = \beta$ and $\kappa = \operatorname{type}(\alpha \times \beta)$) yields $p_\lambda \boxdot p_\mu = p_{\operatorname{type}(\alpha\times\beta)}$. Lemma 13.122.2 (applied to $A \times B$, C, $\alpha \times \beta$, γ, $\operatorname{type}(\alpha \times \beta)$, ν and $\operatorname{type}((\alpha \times \beta) \times \gamma)$ instead of X, Y, σ, τ, λ, μ and κ) yields $p_{\operatorname{type}(\alpha\times\beta)} \boxdot p_\nu = p_{\operatorname{type}((\alpha\times\beta)\times\gamma)}$. Thus,

$$\left(\underbrace{u}_{=p_\lambda} \boxdot \underbrace{v}_{=p_\mu}\right) \boxdot \underbrace{w}_{=p_\nu} = \underbrace{(p_\lambda \boxdot p_\mu)}_{=p_{\operatorname{type}(\alpha\times\beta)}} \boxdot p_\nu = p_{\operatorname{type}(\alpha\times\beta)} \boxdot p_\nu = p_{\operatorname{type}((\alpha\times\beta)\times\gamma)}.$$

A similar argument shows that $u \boxdot (v \boxdot w) = p_{\operatorname{type}(\alpha\times(\beta\times\gamma))}$.

Let us now say that if U and V are two finite sets, and if $\sigma \in \mathfrak{S}_U$ and $\tau \in \mathfrak{S}_V$ are two permutations, then the permutations σ and τ are *isomorphic* if and only if there exists a bijection $\varphi : U \to V$ such that $\varphi \circ \sigma = \tau \circ \varphi$. The intuition behind this meaning of "isomorphism" is that two permutations are isomorphic if one of them becomes the other after a relabelling of its ground set. It is clear that two isomorphic permutations of finite sets must have the same cycle type.

But the permutations $(\alpha \times \beta) \times \gamma \in \mathfrak{S}_{(A\times B)\times C}$ and $\alpha \times (\beta \times \gamma) \in \mathfrak{S}_{A\times(B\times C)}$ are isomorphic (as witnessed by the bijection $\varphi : (A \times B) \times C \to A \times (B \times C)$ sending every $((a, b), c) \in (A \times B) \times C$ to $(a, (b, c))$). Hence, $\operatorname{type}((\alpha \times \beta) \times \gamma) = \operatorname{type}(\alpha \times (\beta \times \gamma))$ (since two isomorphic permutations of finite sets must have the same cycle type). Thus,

$$\begin{aligned}(u \boxdot v) \boxdot w &= p_{\operatorname{type}((\alpha\times\beta)\times\gamma)} = p_{\operatorname{type}(\alpha\times(\beta\times\gamma))} && (\text{since } \operatorname{type}((\alpha \times \beta) \times \gamma) = \operatorname{type}(\alpha \times (\beta \times \gamma)))\\ &= u \boxdot (v \boxdot w).\end{aligned}$$

This finishes the proof of the equality $(u \boxdot v) \boxdot w = u \boxdot (v \boxdot w)$, and thus Exercise 4.4.9(a) is solved.

(b) Let $f \in \Lambda_{\mathbb{Q}}$. We need to prove the equality $1 \boxdot f = \epsilon_1(f) 1$. Since this equality is $\mathbb{Q}$-linear in f (because the operation $\boxdot$ is $\mathbb{Q}$-bilinear and the map ϵ_1 is $\mathbb{Q}$-linear), we can WLOG assume that f belongs to the basis $(p_\lambda)_{\lambda\in\operatorname{Par}}$ of the $\mathbb{Q}$-vector space $\Lambda_{\mathbb{Q}}$. Assume this. Then, there exists a partition λ such that $f = p_\lambda$. Consider this λ. The definition of $p_\varnothing \boxdot p_\lambda$ yields

$$\begin{aligned}p_\varnothing \boxdot p_\lambda &= \prod_{i=1}^{\ell(\varnothing)} \prod_{j=1}^{\ell(\lambda)} p_{\operatorname{lcm}(\varnothing_i,\lambda_j)}^{\gcd(\varnothing_i,\lambda_j)} = \prod_{i=1}^{0} \prod_{j=1}^{\ell(\lambda)} p_{\operatorname{lcm}(\varnothing_i,\lambda_j)}^{\gcd(\varnothing_i,\lambda_j)} && (\text{since } \ell(\varnothing) = 0)\\ &= (\text{empty product}) = 1.\end{aligned}$$

Hence, $\underbrace{1}_{=p_\varnothing} \boxdot \underbrace{f}_{=p_\lambda} = p_\varnothing \boxdot p_\lambda = 1$.

On the other hand, $\epsilon_1(p_\lambda) = 1$ [890]. Thus, $\epsilon_1\left(\underbrace{f}_{=p_\lambda}\right) = \epsilon_1(p_\lambda) = 1$. Hence, $1 \boxdot f = 1 = \epsilon_1(f) = \epsilon_1(f) 1$. This solves Exercise 4.4.9(b).

(c) For every finite group G and every $h \in G$, we define a class function $\alpha_{G,h} \in R_{\mathbb{C}}(G)$ as in Exercise 4.4.3.

[890] *Proof.* We have $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_{\ell(\lambda)})$. Thus, the definition of p_λ yields $p_\lambda = p_{\lambda_1} p_{\lambda_2} \cdots p_{\lambda_{\ell(\lambda)}} = \prod_{i=1}^{\ell(\lambda)} p_{\lambda_i}$. Applying the map ϵ_1 to both sides of this equality, we conclude

$$\begin{aligned}\epsilon_1(p_\lambda) &= \epsilon_1\left(\prod_{i=1}^{\ell(\lambda)} p_{\lambda_i}\right) = \prod_{i=1}^{\ell(\lambda)} \underbrace{\epsilon_1(p_{\lambda_i})}_{\substack{=1\\ (\text{by the definition}\\ \text{of } \epsilon_1)}} && (\text{since } \epsilon_1 \text{ is a } \mathbb{Q}\text{-algebra homomorphism})\\ &= \prod_{i=1}^{\ell(\lambda)} 1 = 1,\end{aligned}$$

qed.

We next recall the notation U^τ defined in Definition 4.3.3. We state a few properties of this notation:

Lemma 13.122.3. *Let K and H be two finite groups. Let $\tau : K \to H$ be a group homomorphism.*

(a) *For every $f \in R_{\mathbb{C}}(H)$, the map $f \circ \tau : K \to \mathbb{C}$ belongs to $R_{\mathbb{C}}(K)$. We can thus define a $\mathbb{C}$-linear map $\tau^* : R_{\mathbb{C}}(H) \to R_{\mathbb{C}}(K)$ by*

$$(\tau^*(f) = f \circ \tau \qquad \text{for every } f \in R_{\mathbb{C}}(H)).$$

(b) *For every (finite-dimensional) $\mathbb{C}H$-module U, we have $\chi_{U^\tau} = \tau^*(\chi_U)$.*

(c) *We have $\tau^*(R(H)) \subset R(K)$.*

(d) *Assume that $\tau : K \to H$ is a group isomorphism. Then, $\tau^*\left(\alpha_{H,\tau(g)}\right) = \alpha_{K,g}$ for every $g \in K$.*

The proof of Lemma 13.122.3 is straightforward and left to the reader. (We will not use its part (c).)

Finally, here come two more simple facts whose proofs we leave to the reader:

Lemma 13.122.4. *Let G and H be two groups. Let $\Omega : G \to H$ be an injective group homomorphism. Define a map $\overline{\Omega} : G \to \Omega(G)$ by*

$$\left(\overline{\Omega}(g) = \Omega(g) \qquad \text{for every } g \in G\right).$$

Then, $\overline{\Omega}$ is a well-defined group isomorphism.

Lemma 13.122.5. *Let U and V be two finite sets. Let $\varphi : U \to V$ be a bijection. Define a map $\varphi^* : \mathfrak{S}_U \to \mathfrak{S}_V$ by*

$$\left(\varphi^*(\pi) = \varphi \circ \pi \circ \varphi^{-1} \qquad \text{for every } \pi \in \mathfrak{S}_U\right).$$

(a) *This map φ^* is well-defined and a group isomorphism.*

(b) *Let $\pi \in \mathfrak{S}_U$. Then, the cycle type of $\varphi^*(\pi)$ equals the cycle type of π.*

Recall that $\boxdot$ is a $\mathbb{Q}$-bilinear map $\Lambda_{\mathbb{Q}} \times \Lambda_{\mathbb{Q}} \to \Lambda_{\mathbb{Q}}$. Let us extend $\boxdot$ to a $\mathbb{C}$-bilinear map $\Lambda_{\mathbb{C}} \times \Lambda_{\mathbb{C}} \to \Lambda_{\mathbb{C}}$; we will still denote this extended map by $\boxdot$.

Recall that $\operatorname{ch} : A \to \Lambda$ is a $\mathbb{Z}$-Hopf algebra isomorphism. Hence, the extension of ch to a $\mathbb{C}$-linear map $A_{\mathbb{C}} \to \Lambda_{\mathbb{C}}$ is a $\mathbb{C}$-Hopf algebra isomorphism. We shall denote this extension by $\operatorname{ch}_{\mathbb{C}}$.

We are going to repeatedly use the following fact (which is easy to obtain from the solution of Exercise 4.4.3): If λ is a partition of a nonnegative integer n, and if $g \in \mathfrak{S}_n$ is a permutation having cycle type λ, then

$$\operatorname{ch}_{\mathbb{C}}\left(\alpha_{\mathfrak{S}_n,g}\right) = p_\lambda. \tag{13.122.2}$$

[891]

Theorem 4.4.1(a) yields $\operatorname{ch}\left(\chi^\lambda\right) = s_\lambda$ for every partition λ (where χ^λ is defined as in Theorem 4.4.1(a)). Since $\operatorname{ch}_{\mathbb{C}}$ is the extension of ch to a $\mathbb{C}$-linear map $A_{\mathbb{C}} \to \Lambda_{\mathbb{C}}$, we have

$$\operatorname{ch}_{\mathbb{C}}\left(\chi^\lambda\right) = \operatorname{ch}\left(\chi^\lambda\right) = s_\lambda \qquad \text{for every partition } \lambda. \tag{13.122.3}$$

Let us now come back to solving Exercise 4.4.9(c). Let μ and ν be two partitions. We need to show that $s_\mu \boxdot s_\nu \in \sum_{\lambda \in \operatorname{Par}} \mathbb{N} s_\lambda$.

Let $m = |\mu|$ and $n = |\nu|$. Then, $\mu \in \operatorname{Par}_m$ and $\nu \in \operatorname{Par}_n$. If $\min\{m,n\} = 0$, then $s_\mu \boxdot s_\nu \in \sum_{\lambda \in \operatorname{Par}} \mathbb{N} s_\lambda$ is easily seen to hold[892]. Hence, for the rest of this proof, we can WLOG assume that we don't have $\min\{m,n\} = 0$. Assume this. Thus, $\min\{m,n\} \geq 1$. Hence, $m \geq 1$ and $n \geq 1$.

[891] *Proof of (13.122.2):* Let λ be a partition of a nonnegative integer n. Let $g \in \mathfrak{S}_n$ be a permutation having cycle type λ. Define a map $\Phi : \Lambda_{\mathbb{C}} \to A_{\mathbb{C}}$ as in the solution of Exercise 4.4.3(d). Then, (13.117.8) yields $\Phi(p_\lambda) = \alpha_{\mathfrak{S}_n,g}$, so that $\alpha_{\mathfrak{S}_n,g} = \Phi(p_\lambda)$. But $\operatorname{ch}_{\mathbb{C}} \circ \Phi = \operatorname{id}_{\Lambda_{\mathbb{C}}}$ (this was shown in the solution of Exercise 4.4.3(d)). Now, $\operatorname{ch}_{\mathbb{C}}\left(\underbrace{\alpha_{\mathfrak{S}_n,g}}_{=\Phi(p_\lambda)}\right) = \operatorname{ch}_{\mathbb{C}}(\Phi(p_\lambda)) = \underbrace{(\operatorname{ch}_{\mathbb{C}} \circ \Phi)}_{=\operatorname{id}_{\Lambda_{\mathbb{C}}}}(p_\lambda) = \operatorname{id}_{\Lambda_{\mathbb{C}}}(p_\lambda) = p_\lambda$. This proves (13.122.2).

[892] *Proof.* Assume that $\min\{m,n\} = 0$.

Recall that $\Lambda_{\mathbb{Q}}$, equipped with the binary operation $\boxdot$, becomes a commutative $\mathbb{Q}$-algebra with unity p_1 (according to Exercise 4.4.9(a)). Thus, the operation $\boxdot$ is commutative. Hence, $s_\mu \boxdot s_\nu = s_\nu \boxdot s_\mu$.

We notice that

$$\text{(13.122.4)} \qquad \operatorname{ch}_{\mathbb{C}}(\chi_P) \in \sum_{\lambda \in \operatorname{Par}} \mathbb{N} s_\lambda \qquad \text{for every finite-dimensional } \mathbb{C}\mathfrak{S}_{nm}\text{-module } P.$$

[893]

We have $\mathfrak{S}_n = \mathfrak{S}_{\{1,2,\ldots,n\}}$ (by the definition of $\mathfrak{S}_n$) and $\mathfrak{S}_m = \mathfrak{S}_{\{1,2,\ldots,m\}}$ (by the definition of $\mathfrak{S}_m$) and $\mathfrak{S}_{nm} = \mathfrak{S}_{\{1,2,\ldots,nm\}}$ (by the definition of $\mathfrak{S}_{nm}$). Recall that we have defined a map $\operatorname{cross} : \mathfrak{S}_X \times \mathfrak{S}_Y \to \mathfrak{S}_{X\times Y}$ for any two sets X and Y. Now, set $X = \{1,2,\ldots,n\}$ and $Y = \{1,2,\ldots,m\}$, and consider the map $\operatorname{cross} : \mathfrak{S}_X \times \mathfrak{S}_Y \to \mathfrak{S}_{X\times Y}$. In the following, when we speak of cross, we will always mean this map.

We have $\left|\underbrace{X}_{=\{1,2,\ldots,n\}}\right| = |\{1,2,\ldots,n\}| = n$ and $\left|\underbrace{Y}_{=\{1,2,\ldots,m\}}\right| = |\{1,2,\ldots,m\}| = m$. Also, since $X = \{1,2,\ldots,n\}$, we have $\mathfrak{S}_X = \mathfrak{S}_{\{1,2,\ldots,n\}} = \mathfrak{S}_n$. Similarly, $\mathfrak{S}_Y = \mathfrak{S}_m$.

The set X is nonempty (since $|X| = n \geq 1$), and the set Y is nonempty (similarly). Hence, Lemma 13.122.1(b) yields that the map $\operatorname{cross} : \mathfrak{S}_X \times \mathfrak{S}_Y \to \mathfrak{S}_{X\times Y}$ is injective. Lemma 13.122.1(a) yields that the map $\operatorname{cross} : \mathfrak{S}_X \times \mathfrak{S}_Y \to \mathfrak{S}_{X\times Y}$ is a group homomorphism. Thus, the map cross is a group homomorphism from $\mathfrak{S}_X \times \mathfrak{S}_Y$ to $\mathfrak{S}_{X\times Y}$. Since $\mathfrak{S}_X = \mathfrak{S}_n$ and $\mathfrak{S}_Y = \mathfrak{S}_m$, this rewrites as follows: The map cross is a group homomorphism from $\mathfrak{S}_n \times \mathfrak{S}_m$ to $\mathfrak{S}_{X\times Y}$.

The set $X \times Y$ has cardinality $|X \times Y| = \underbrace{|X|}_{=n} \cdot \underbrace{|Y|}_{=m} = nm$, and thus is in bijection with the set $\{1,2,\ldots,nm\}$. In other words, there exists a bijection $\varphi : X \times Y \to \{1,2,\ldots,nm\}$. Fix such a bijection φ. Define a map $\varphi^* : \mathfrak{S}_{X\times Y} \to \mathfrak{S}_{\{1,2,\ldots,nm\}}$ as in Lemma 13.122.5 (applied to $U = X \times Y$ and $V = \{1,2,\ldots,nm\}$). Then, Lemma 13.122.5(a) (applied to $U = X \times Y$ and $V = \{1,2,\ldots,nm\}$) yields that this map φ^* is well-defined and a group isomorphism. In particular, φ^* is an injective group homomorphism.

We can WLOG assume that $m \leq n$ (since otherwise, we can just interchange μ and m with ν and n (because $s_\mu \boxdot s_\nu = s_\nu \boxdot s_\mu$)). Assume this. Then, $\min\{m,n\} = m$, so that $m = \min\{m,n\} = 0$. Hence, $\mu \in \operatorname{Par}_m = \operatorname{Par}_0$ (since $m = 0$), so that $\mu = \varnothing$ and thus $s_\mu = s_\varnothing = 1$.

Let us use the notations of Exercise 4.4.9(b). We have $\underbrace{s_\mu}_{=1} \boxdot s_\nu = 1 \boxdot s_\nu = \epsilon_1(s_\nu)\,1$ (by Exercise 4.4.9(b), applied to $f = s_\nu$). But $\epsilon_1(s_\nu) = s_\nu(1)$ (by Exercise 2.9.4(i), applied to $f = s_\nu$) and $s_\nu(1) \in \mathbb{N}$ (since s_ν is a sum of monomials). Hence, $s_\mu \boxdot s_\nu = \underbrace{\epsilon_1(s_\nu)}_{\in\mathbb{N}}\underbrace{1}_{=s_\varnothing} \in \mathbb{N}s_\varnothing \subset \sum_{\lambda\in\operatorname{Par}} \mathbb{N}s_\lambda$, qed.

[893] *Proof of (13.122.4):* Let P be a finite-dimensional $\mathbb{C}\mathfrak{S}_{nm}$-module. Then, P must be a direct sum of finitely many simple $\mathbb{C}\mathfrak{S}_{nm}$-modules. In other words, there exist some simple $\mathbb{C}\mathfrak{S}_{nm}$-modules $V_1, V_2, \ldots, V_j$ such that $P \cong V_1 \oplus V_2 \oplus \cdots \oplus V_j$ as $\mathbb{C}\mathfrak{S}_{nm}$-modules. Consider these $V_1, V_2, \ldots, V_j$. We shall now show that $\chi_{V_i} \in \sum_{\lambda\in\operatorname{Par}} \mathbb{N}\chi^\lambda$ for every $i \in \{1,2,\ldots,j\}$.

Indeed, let $i \in \{1,2,\ldots,j\}$. Then, V_i is a simple $\mathbb{C}\mathfrak{S}_{nm}$-module. Hence, χ_{V_i} is an irreducible character of $\mathbb{C}\mathfrak{S}_{nm}$. But (a part of) Theorem 4.4.1(a) (applied to nm instead of n) says that all irreducible characters of $\mathfrak{S}_{nm}$ have the form χ^λ with $\lambda \in \operatorname{Par}_{nm}$. Thus, χ_{V_i} has the form $\chi_{V_i} = \chi^\lambda$ for some $\lambda \in \operatorname{Par}_{nm}$. In other words, $\chi_{V_i} \in \left\{\chi^\lambda \mid \lambda \in \underbrace{\operatorname{Par}_{nm}}_{\subset\operatorname{Par}}\right\} \subset \left\{\chi^\lambda \mid \lambda \in \operatorname{Par}\right\} \subset \sum_{\lambda\in\operatorname{Par}} \mathbb{N}\chi^\lambda$.

Now, let us forget that we fixed i. We thus have shown that $\chi_{V_i} \in \sum_{\lambda\in\operatorname{Par}} \mathbb{N}\chi^\lambda$ for every $i \in \{1,2,\ldots,j\}$. Thus, $\sum_{i=1}^j \underbrace{\chi_{V_i}}_{\in\sum_{\lambda\in\operatorname{Par}}\mathbb{N}\chi^\lambda} \in \sum_{i=1}^j \sum_{\lambda\in\operatorname{Par}} \mathbb{N}\chi^\lambda \subset \sum_{\lambda\in\operatorname{Par}} \mathbb{N}\chi^\lambda$ (since $\sum_{\lambda\in\operatorname{Par}} \mathbb{N}\chi^\lambda$ is closed under addition).

But isomorphic $\mathbb{C}\mathfrak{S}_{nm}$-modules have equal characters. Hence, since $P \cong V_1 \oplus V_2 \oplus \cdots \oplus V_j$, we have

$$\chi_P = \chi_{V_1\oplus V_2\oplus\cdots\oplus V_j} = \chi_{V_1} + \chi_{V_2} + \cdots + \chi_{V_j} = \sum_{i=1}^j \chi_{V_i} \in \sum_{\lambda\in\operatorname{Par}} \mathbb{N}\chi^\lambda.$$

Applying the map $\operatorname{ch}_{\mathbb{C}}$ to both sides of this relation, we obtain

$$\begin{aligned} \operatorname{ch}_{\mathbb{C}}(\chi_P) &\in \operatorname{ch}_{\mathbb{C}}\left(\sum_{\lambda\in\operatorname{Par}} \mathbb{N}\chi^\lambda\right) \subset \sum_{\lambda\in\operatorname{Par}} \mathbb{N} \underbrace{\operatorname{ch}_{\mathbb{C}}\left(\chi^\lambda\right)}_{\substack{=s_\lambda \\ \text{(by (13.122.3))}}} \qquad \text{(since the map } \operatorname{ch}_{\mathbb{C}} \text{ is } \mathbb{Z}\text{-linear)} \\ &= \sum_{\lambda\in\operatorname{Par}} \mathbb{N}s_\lambda. \end{aligned}$$

This proves (13.122.4).

Also, φ^* is a group homomorphism from $\mathfrak{S}_{X\times Y}$ to $\mathfrak{S}_{\{1,2,\ldots,nm\}}$, therefore a group homomorphism from $\mathfrak{S}_{X\times Y}$ to $\mathfrak{S}_{nm}$ (since $\mathfrak{S}_{nm} = \mathfrak{S}_{\{1,2,\ldots,nm\}}$).

Hence, $\varphi^* \circ \operatorname{cross}$ is a group homomorphism from $\mathfrak{S}_n \times \mathfrak{S}_m$ to $\mathfrak{S}_{nm}$ (since φ^* is a group homomorphism from $\mathfrak{S}_{X\times Y}$ to $\mathfrak{S}_{nm}$, and since cross is a group homomorphism from $\mathfrak{S}_n \times \mathfrak{S}_m$ to $\mathfrak{S}_{X\times Y}$). Also, $\varphi^* \circ \operatorname{cross}$ is injective (since φ^* and cross are injective). Let $\Omega = \varphi^* \circ \operatorname{cross}$. Then, Ω is an injective group homomorphism from $\mathfrak{S}_n \times \mathfrak{S}_m$ to $\mathfrak{S}_{nm}$ (since $\varphi^* \circ \operatorname{cross}$ is an injective group homomorphism from $\mathfrak{S}_n \times \mathfrak{S}_m$ to $\mathfrak{S}_{nm}$). We can define a map $\overline{\Omega} : \mathfrak{S}_n \times \mathfrak{S}_m \to \Omega\left(\mathfrak{S}_n \times \mathfrak{S}_m\right)$ by

$$\left(\overline{\Omega}\left(g\right) = \Omega\left(g\right) \qquad \text{for every } g \in \mathfrak{S}_n \times \mathfrak{S}_m\right)$$

(since $\Omega\left(g\right) \in \Omega\left(\mathfrak{S}_n \times \mathfrak{S}_m\right)$ for every $g \in \mathfrak{S}_n \times \mathfrak{S}_m$). Consider this $\overline{\Omega}$. Lemma 13.122.4 (applied to $\mathfrak{S}_n \times \mathfrak{S}_m$ and $\mathfrak{S}_{nm}$ instead of G and H) yields that $\overline{\Omega}$ is a well-defined group isomorphism. Hence, the inverse $\overline{\Omega}^{-1}$ of $\overline{\Omega}$ is well-defined and also a group isomorphism. Let τ denote this inverse $\overline{\Omega}^{-1}$. Thus,

$$\tau = \overline{\Omega}^{-1} \text{ is a group isomorphism } \Omega\left(\mathfrak{S}_n \times \mathfrak{S}_m\right) \to \mathfrak{S}_n \times \mathfrak{S}_m.$$

Hence, a $\mathbb{C}$-linear map $\tau^* : R_{\mathbb{C}}\left(\mathfrak{S}_n \times \mathfrak{S}_m\right) \to R_{\mathbb{C}}\left(\Omega\left(\mathfrak{S}_n \times \mathfrak{S}_m\right)\right)$ is defined (according to Lemma 13.122.3(a), applied to $\Omega\left(\mathfrak{S}_n \times \mathfrak{S}_m\right)$ and $\mathfrak{S}_n \times \mathfrak{S}_m$ instead of K and H).

The following commutative diagram illustrates the group homomorphisms we have just introduced:

$$\begin{tikzcd}
\mathfrak{S}_n \times \mathfrak{S}_m \arrow[r, hook, "\text{cross}"] \arrow[rr, bend left=40, "\Omega"] \arrow[dr, "\overline{\Omega}", "\cong"'] & \mathfrak{S}_{X\times Y} \arrow[r, "\varphi^*", "\cong"'] & \mathfrak{S}_{nm} \\
& \Omega\left(\mathfrak{S}_n \times \mathfrak{S}_m\right) \arrow[ur, hook, "\text{inclusion}"'] \arrow[ul, bend left=60, "\tau", "\cong"'] &
\end{tikzcd}$$

(where the cycle formed by the $\overline{\Omega}$ and τ arrows is not a mistake: these two maps are mutually inverse!).

In the following, for every $k \in \mathbb{N}$, we let $\left(\Lambda_{\mathbb{C}}\right)_k$ denote the k-th homogeneous component of the graded $\mathbb{C}$-algebra $\Lambda_{\mathbb{C}}$. Note that $\left(p_\lambda\right)_{\lambda \in \operatorname{Par}_k}$ is a basis of this $\mathbb{C}$-vector space $\left(\Lambda_{\mathbb{C}}\right)_k$.

Let us now claim that every $a \in \left(\Lambda_{\mathbb{C}}\right)_n$ and $b \in \left(\Lambda_{\mathbb{C}}\right)_m$ satisfy

$$a \boxdot b = \operatorname{ch}_{\mathbb{C}}\left(\operatorname{Ind}_{\Omega\left(\mathfrak{S}_n\times\mathfrak{S}_m\right)}^{\mathfrak{S}_{nm}}\left(\tau^*\left(\left(\operatorname{ch}_{\mathbb{C}}\right)^{-1}\left(a\right) \otimes \left(\operatorname{ch}_{\mathbb{C}}\right)^{-1}\left(b\right)\right)\right)\right) \tag{13.122.5}$$

(where we identify $R_{\mathbb{C}}\left(\mathfrak{S}_n\right) \otimes R_{\mathbb{C}}\left(\mathfrak{S}_m\right)$ with $R_{\mathbb{C}}\left(\mathfrak{S}_n \times \mathfrak{S}_m\right)$ along the canonical isomorphism $R_{\mathbb{C}}\left(\mathfrak{S}_n\right) \otimes R_{\mathbb{C}}\left(\mathfrak{S}_m\right) \to R_{\mathbb{C}}\left(\mathfrak{S}_n \times \mathfrak{S}_m\right)$) [894].

Proof of (13.122.5): Let $a \in \left(\Lambda_{\mathbb{C}}\right)_n$ and $b \in \left(\Lambda_{\mathbb{C}}\right)_m$. We need to prove the equality (13.122.5). Since this equality is $\mathbb{C}$-linear in each of a and b (because the operations $\boxdot$ and $\otimes$ are $\mathbb{C}$-bilinear, and the maps $\operatorname{ch}_{\mathbb{C}}$ $\left(\operatorname{ch}_{\mathbb{C}}\right)^{-1}$, $\operatorname{Ind}_{\Omega\left(\mathfrak{S}_n\times\mathfrak{S}_m\right)}^{\mathfrak{S}_{nm}}$ and τ^* are $\mathbb{C}$-linear), we can WLOG assume that a is an element of the basis $\left(p_\lambda\right)_{\lambda \in \operatorname{Par}_n}$ of the $\mathbb{C}$-vector space $\left(\Lambda_{\mathbb{C}}\right)_n$, and that b is an element of the basis $\left(p_\lambda\right)_{\lambda \in \operatorname{Par}_m}$ of the $\mathbb{C}$-vector space $\left(\Lambda_{\mathbb{C}}\right)_m$. Assume this. Then, we can write a and b as $a = p_\gamma$ and $b = p_\eta$ for some $\gamma \in \operatorname{Par}_n$ and some $\eta \in \operatorname{Par}_m$. Consider these γ and η.

Choose some permutation $g \in \mathfrak{S}_n$ which has cycle type γ. (Such a g clearly exists.) Then, $\operatorname{ch}_{\mathbb{C}}\left(\alpha_{\mathfrak{S}_n,g}\right) = p_\gamma$ (according to (13.122.2)). Hence, $\left(\operatorname{ch}_{\mathbb{C}}\right)^{-1}\left(p_\gamma\right) = \alpha_{\mathfrak{S}_n,g}$, so that $\left(\operatorname{ch}_{\mathbb{C}}\right)^{-1}\left(\underbrace{a}_{=p_\gamma}\right) = \left(\operatorname{ch}_{\mathbb{C}}\right)^{-1}\left(p_\gamma\right) = \alpha_{\mathfrak{S}_n,g}$.

[894] Notice that the term $\tau^*\left(\left(\operatorname{ch}_{\mathbb{C}}\right)^{-1}\left(a\right) \otimes \left(\operatorname{ch}_{\mathbb{C}}\right)^{-1}\left(b\right)\right)$ on the right hand side of (13.122.5) is well-defined. (This is because

$$\begin{aligned}\left(\operatorname{ch}_{\mathbb{C}}\right)^{-1}\left(\underbrace{a}_{\in\left(\Lambda_{\mathbb{C}}\right)_n}\right) \otimes \left(\operatorname{ch}_{\mathbb{C}}\right)^{-1}\left(\underbrace{b}_{\in\left(\Lambda_{\mathbb{C}}\right)_m}\right) &\in \underbrace{\left(\operatorname{ch}_{\mathbb{C}}\right)^{-1}\left(\left(\Lambda_{\mathbb{C}}\right)_n\right)}_{=R_{\mathbb{C}}\left(\mathfrak{S}_n\right)} \otimes \underbrace{\left(\operatorname{ch}_{\mathbb{C}}\right)^{-1}\left(\left(\Lambda_{\mathbb{C}}\right)_m\right)}_{=R_{\mathbb{C}}\left(\mathfrak{S}_m\right)} \\ &= R_{\mathbb{C}}\left(\mathfrak{S}_n\right) \otimes R_{\mathbb{C}}\left(\mathfrak{S}_m\right) = R_{\mathbb{C}}\left(\mathfrak{S}_n \times \mathfrak{S}_m\right).\end{aligned}$$

)

Choose some permutation $h \in \mathfrak{S}_m$ which has cycle type η. (Such an h clearly exists.) Then, $\operatorname{ch}_{\mathbb{C}}\left(\alpha_{\mathfrak{S}_m,h}\right) = p_\eta$ (according to (13.122.2), applied to m, h and η instead of n, g and λ). Hence, $\left(\operatorname{ch}_{\mathbb{C}}\right)^{-1}(p_\eta) = \alpha_{\mathfrak{S}_m,h}$, so that $\left(\operatorname{ch}_{\mathbb{C}}\right)^{-1}\left(\underbrace{b}_{=p_\eta}\right) = \left(\operatorname{ch}_{\mathbb{C}}\right)^{-1}(p_\eta) = \alpha_{\mathfrak{S}_m,h}$. Thus,

$$\underbrace{\left(\operatorname{ch}_{\mathbb{C}}\right)^{-1}(a)}_{=\alpha_{\mathfrak{S}_n,g}} \otimes \underbrace{\left(\operatorname{ch}_{\mathbb{C}}\right)^{-1}(b)}_{=\alpha_{\mathfrak{S}_m,h}} = \alpha_{\mathfrak{S}_n,g} \otimes \alpha_{\mathfrak{S}_m,h}. \tag{13.122.6}$$

Exercise 4.4.3(c) (applied to $G_1 = \mathfrak{S}_n$, $G_2 = \mathfrak{S}_m$, $h_1 = g$ and $h_2 = h$) yields that the canonical isomorphism $R_{\mathbb{C}}(\mathfrak{S}_n) \otimes R_{\mathbb{C}}(\mathfrak{S}_m) \to R_{\mathbb{C}}(\mathfrak{S}_n \times \mathfrak{S}_m)$ sends $\alpha_{\mathfrak{S}_n,g} \otimes \alpha_{\mathfrak{S}_m,h}$ to $\alpha_{\mathfrak{S}_n \times \mathfrak{S}_m,(g,h)}$. Since we are regarding this isomorphism as an identity (because we have identified $R_{\mathbb{C}}(\mathfrak{S}_n) \otimes R_{\mathbb{C}}(\mathfrak{S}_m)$ with $R_{\mathbb{C}}(\mathfrak{S}_n \times \mathfrak{S}_m)$ along this isomorphism), this yields $\alpha_{\mathfrak{S}_n,g} \otimes \alpha_{\mathfrak{S}_m,h} = \alpha_{\mathfrak{S}_n \times \mathfrak{S}_m,(g,h)}$. Thus, (13.122.6) becomes

$$\left(\operatorname{ch}_{\mathbb{C}}\right)^{-1}(a) \otimes \left(\operatorname{ch}_{\mathbb{C}}\right)^{-1}(b) = \alpha_{\mathfrak{S}_n,g} \otimes \alpha_{\mathfrak{S}_m,h} = \alpha_{\mathfrak{S}_n \times \mathfrak{S}_m,(g,h)}.$$

Applying the map τ^* to both sides of this equality, we obtain

$$\tau^*\left(\left(\operatorname{ch}_{\mathbb{C}}\right)^{-1}(a) \otimes \left(\operatorname{ch}_{\mathbb{C}}\right)^{-1}(b)\right) = \tau^*\left(\alpha_{\mathfrak{S}_n \times \mathfrak{S}_m,(g,h)}\right). \tag{13.122.7}$$

But $\overline{\Omega}((g,h)) = \Omega((g,h))$ (by the definition of $\overline{\Omega}((g,h))$), and so $(g,h) = \underbrace{\overline{\Omega}^{-1}}_{=\tau}(\Omega((g,h))) = \tau(\Omega((g,h)))$. Thus,

$$\tau^*\left(\alpha_{\mathfrak{S}_n \times \mathfrak{S}_m,(g,h)}\right) = \tau^*\left(\alpha_{\mathfrak{S}_n \times \mathfrak{S}_m,\tau(\Omega((g,h)))}\right) = \alpha_{\Omega(\mathfrak{S}_n \times \mathfrak{S}_m),\Omega((g,h))}$$

(by Lemma 13.122.3(d), applied to $\Omega(\mathfrak{S}_n \times \mathfrak{S}_m)$, $\mathfrak{S}_n \times \mathfrak{S}_m$ and $\Omega((g,h))$ instead of K, H and g). Thus, (13.122.7) becomes

$$\tau^*\left(\left(\operatorname{ch}_{\mathbb{C}}\right)^{-1}(a) \otimes \left(\operatorname{ch}_{\mathbb{C}}\right)^{-1}(b)\right) = \tau^*\left(\alpha_{\mathfrak{S}_n \times \mathfrak{S}_m,(g,h)}\right) = \alpha_{\Omega(\mathfrak{S}_n \times \mathfrak{S}_m),\Omega((g,h))}.$$

Applying the map $\operatorname{Ind}_{\Omega(\mathfrak{S}_n \times \mathfrak{S}_m)}^{\mathfrak{S}_{nm}}$ to both sides of this equality, we obtain

$$\begin{aligned} \operatorname{Ind}_{\Omega(\mathfrak{S}_n \times \mathfrak{S}_m)}^{\mathfrak{S}_{nm}}\left(\tau^*\left(\left(\operatorname{ch}_{\mathbb{C}}\right)^{-1}(a) \otimes \left(\operatorname{ch}_{\mathbb{C}}\right)^{-1}(b)\right)\right) &= \operatorname{Ind}_{\Omega(\mathfrak{S}_n \times \mathfrak{S}_m)}^{\mathfrak{S}_{nm}}\left(\alpha_{\Omega(\mathfrak{S}_n \times \mathfrak{S}_m),\Omega((g,h))}\right) \\ &= \alpha_{\mathfrak{S}_{nm},\Omega((g,h))} \end{aligned} \tag{13.122.8}$$

(by Exercise 4.4.3(b), applied to $\mathfrak{S}_{nm}$, $\Omega(\mathfrak{S}_n \times \mathfrak{S}_m)$ and $\Omega((g,h))$ instead of G, H and h).

Now, $\Omega((g,h)) \in \mathfrak{S}_{nm}$, so that $\Omega((g,h))$ is a permutation of $\{1,2,\ldots,nm\}$. Let κ be the cycle type of this permutation $\Omega((g,h))$. Then, (13.122.2) (applied to nm, κ and $\Omega((g,h))$ instead of n, λ and g) yields $\operatorname{ch}_{\mathbb{C}}\left(\alpha_{\mathfrak{S}_{nm},\Omega((g,h))}\right) = p_\kappa$. But applying the map $\operatorname{ch}_{\mathbb{C}}$ to both sides of the identity (13.122.8), we obtain

$$\begin{aligned} &\operatorname{ch}_{\mathbb{C}}\left(\operatorname{Ind}_{\Omega(\mathfrak{S}_n \times \mathfrak{S}_m)}^{\mathfrak{S}_{nm}}\left(\tau^*\left(\left(\operatorname{ch}_{\mathbb{C}}\right)^{-1}(a) \otimes \left(\operatorname{ch}_{\mathbb{C}}\right)^{-1}(b)\right)\right)\right) \\ &= \operatorname{ch}_{\mathbb{C}}\left(\alpha_{\mathfrak{S}_{nm},\Omega((g,h))}\right) = p_\kappa. \end{aligned} \tag{13.122.9}$$

Let us now show that $a \boxdot b = p_\kappa$.

In fact, $\underbrace{\Omega}_{=\varphi^* \circ \operatorname{cross}}((g,h)) = (\varphi^* \circ \operatorname{cross})((g,h)) = \varphi^*\left(\underbrace{\operatorname{cross}((g,h))}_{\substack{=g \times h \\ \text{(by the definition} \\ \text{of cross)}}}\right) = \varphi^*(g \times h)$. But Lemma 13.122.5(b) (applied to $U = X \times Y$, $V = \{1,2,\ldots,nm\}$ and $\pi = g \times h$) yields that the cycle type of $\varphi^*(g \times h)$ equals the cycle type of $g \times h$. In other words, the cycle type of $\Omega((g,h))$ equals the cycle type of $g \times h$ (since $\Omega((g,h)) = \varphi^*(g \times h)$). In other words, κ equals the cycle type of $g \times h$ (since κ is the cycle type of $\Omega((g,h))$).

So we know that $g \in \mathfrak{S}_n = \mathfrak{S}_X$ and $h \in \mathfrak{S}_m = \mathfrak{S}_Y$ are permutations, and that γ, η and κ are the cycle types of the permutations g, h and $g \times h$. Thus, Lemma 13.122.2 (applied to γ and η instead of λ and μ)

yields $p_\gamma \boxdot p_\eta = p_\kappa$. Now,

$$\underbrace{a}_{=p_\gamma} \boxdot \underbrace{b}_{=p_\eta} = p_\gamma \boxdot p_\eta = p_\kappa = \operatorname{ch}_{\mathbb{C}}\left(\operatorname{Ind}_{\Omega(\mathfrak{S}_n\times\mathfrak{S}_m)}^{\mathfrak{S}_{nm}}\left(\tau^*\left(\left(\operatorname{ch}_{\mathbb{C}}\right)^{-1}(a)\otimes\left(\operatorname{ch}_{\mathbb{C}}\right)^{-1}(b)\right)\right)\right)$$

(by (13.122.9)). This proves (13.122.5).

Now that (13.122.5) is proven, it is easy to complete the solution of Exercise 4.4.9(c). Recall that μ and ν are two partitions such that $\mu \in \operatorname{Par}_m$, $\nu \in \operatorname{Par}_n$, $m \geq 1$ and $n \geq 1$. We need to show that $s_\mu \boxdot s_\nu \in \sum_{\lambda\in\operatorname{Par}} \mathbb{N} s_\lambda$.

Applying (13.122.3) to $\lambda = \mu$, we obtain $\operatorname{ch}_{\mathbb{C}}(\chi^\mu) = s_\mu$, so that $(\operatorname{ch}_{\mathbb{C}})^{-1}(s_\mu) = \chi^\mu$. The same argument (but with μ replaced by ν) shows that $(\operatorname{ch}_{\mathbb{C}})^{-1}(s_\nu) = \chi^\nu$.

We know that χ^μ is an irreducible complex character of $\mathbb{C}\mathfrak{S}_m$. In other words, there exists a simple $\mathbb{C}\mathfrak{S}_m$-module M such that $\chi^\mu = \chi_M$. Consider this M. Then, $(\operatorname{ch}_{\mathbb{C}})^{-1}(s_\mu) = \chi^\mu = \chi_M$.

We know that χ^ν is an irreducible complex character of $\mathbb{C}\mathfrak{S}_n$. In other words, there exists a simple $\mathbb{C}\mathfrak{S}_n$-module N such that $\chi^\nu = \chi_N$. Consider this N. Then, $(\operatorname{ch}_{\mathbb{C}})^{-1}(s_\nu) = \chi^\nu = \chi_N$.

Recall that we are identifying $R_{\mathbb{C}}(\mathfrak{S}_n) \otimes R_{\mathbb{C}}(\mathfrak{S}_m)$ with $R_{\mathbb{C}}(\mathfrak{S}_n \times \mathfrak{S}_m)$ along the canonical isomorphism $R_{\mathbb{C}}(\mathfrak{S}_n) \otimes R_{\mathbb{C}}(\mathfrak{S}_m) \to R_{\mathbb{C}}(\mathfrak{S}_n \times \mathfrak{S}_m)$. Thus, for any finite-dimensional $\mathbb{C}\mathfrak{S}_n$-module U and any finite-dimensional $\mathbb{C}\mathfrak{S}_m$-module V, we have $\chi_{U\otimes V} = \chi_U \otimes \chi_V$ (since this isomorphism sends $\chi_U \otimes \chi_V$ to $\chi_{U\otimes V}$). Applying this to $U = N$ and $V = M$, we obtain $\chi_{N\otimes M} = \chi_N \otimes \chi_M$. Thus, $\chi_N \otimes \chi_M = \chi_{N\otimes M}$.

Lemma 13.122.3(b) (applied to $\Omega(\mathfrak{S}_n \times \mathfrak{S}_m)$, $\mathfrak{S}_n \times \mathfrak{S}_m$ and $N \otimes M$ instead of K, H and U) yields $\chi_{(N\otimes M)^\tau} = \tau^*(\chi_{N\otimes M})$. Thus, $\tau^*(\chi_{N\otimes M}) = \chi_{(N\otimes M)^\tau}$. Hence,

$$\tau^*\left(\underbrace{\chi_N \otimes \chi_M}_{=\chi_{N\otimes M}}\right) = \tau^*(\chi_{N\otimes M}) = \chi_{(N\otimes M)^\tau}. \tag{13.122.10}$$

But (4.1.5) (applied to $\mathfrak{S}_{nm}$, $\Omega(\mathfrak{S}_n \times \mathfrak{S}_m)$ and $(N\otimes M)^\tau$ instead of G, H and U) yields $\chi_{\operatorname{Ind}_{\Omega(\mathfrak{S}_n\times\mathfrak{S}_m)}^{\mathfrak{S}_{nm}}((N\otimes M)^\tau)} = \operatorname{Ind}_{\Omega(\mathfrak{S}_n\times\mathfrak{S}_m)}^{\mathfrak{S}_{nm}}\left(\chi_{(N\otimes M)^\tau}\right)$. Let P denote the finite-dimensional $\mathbb{C}\mathfrak{S}_{nm}$-module $\operatorname{Ind}_{\Omega(\mathfrak{S}_n\times\mathfrak{S}_m)}^{\mathfrak{S}_{nm}}((N\otimes M)^\tau)$. Then, $P = \operatorname{Ind}_{\Omega(\mathfrak{S}_n\times\mathfrak{S}_m)}^{\mathfrak{S}_{nm}}((N\otimes M)^\tau)$, so that

$$\chi_P = \chi_{\operatorname{Ind}_{\Omega(\mathfrak{S}_n\times\mathfrak{S}_m)}^{\mathfrak{S}_{nm}}((N\otimes M)^\tau)} = \operatorname{Ind}_{\Omega(\mathfrak{S}_n\times\mathfrak{S}_m)}^{\mathfrak{S}_{nm}}\left(\underbrace{\chi_{(N\otimes M)^\tau}}_{\substack{=\tau^*(\chi_N\otimes\chi_M)\\ (\text{by } (13.122.10))}}\right) = \operatorname{Ind}_{\Omega(\mathfrak{S}_n\times\mathfrak{S}_m)}^{\mathfrak{S}_{nm}}\left(\tau^*(\chi_N\otimes\chi_M)\right).$$

Hence,

$$\operatorname{Ind}_{\Omega(\mathfrak{S}_n\times\mathfrak{S}_m)}^{\mathfrak{S}_{nm}}\left(\tau^*(\chi_N\otimes\chi_M)\right) = \chi_P. \tag{13.122.11}$$

We have $s_\mu \in (\Lambda_{\mathbb{C}})_m$ (since $\mu \in \operatorname{Par}_m$) and $s_\nu \in (\Lambda_{\mathbb{C}})_n$ (since $\nu \in \operatorname{Par}_n$). Thus, we can apply (13.122.5) to $a = s_\nu$ and $b = s_\mu$.

Exercise 4.4.9(a) yields that $\Lambda_{\mathbb{Q}}$, equipped with the binary operation $\boxdot$, becomes a commutative $\mathbb{Q}$-algebra with unity p_1. Thus, the operation $\boxdot$ is commutative. Hence,

$$\begin{aligned} s_\mu \boxdot s_\nu = s_\nu \boxdot s_\mu &= \operatorname{ch}_{\mathbb{C}}\left(\operatorname{Ind}_{\Omega(\mathfrak{S}_n\times\mathfrak{S}_m)}^{\mathfrak{S}_{nm}}\left(\tau^*\left(\underbrace{(\operatorname{ch}_{\mathbb{C}})^{-1}(s_\nu)}_{=\chi_N}\otimes\underbrace{(\operatorname{ch}_{\mathbb{C}})^{-1}(s_\mu)}_{=\chi_M}\right)\right)\right) \\ &\qquad (\text{by } (13.122.5), \text{ applied to } a = s_\nu \text{ and } b = s_\mu) \\ &= \operatorname{ch}_{\mathbb{C}}\left(\underbrace{\operatorname{Ind}_{\Omega(\mathfrak{S}_n\times\mathfrak{S}_m)}^{\mathfrak{S}_{nm}}\left(\tau^*(\chi_N\otimes\chi_M)\right)}_{\substack{=\chi_P\\ (\text{by } (13.122.11))}}\right) = \operatorname{ch}_{\mathbb{C}}(\chi_P) \in \sum_{\lambda\in\operatorname{Par}} \mathbb{N} s_\lambda \qquad (\text{by } (13.122.4)). \end{aligned}$$

This solves Exercise 4.4.9(c).

[*Remark:* The above solution can be simplified using the results of Exercise 4.1.14. We leave the details of this simplification to the reader, and only mention some highlights. Let us use the notations of Exercise 4.1.14. Then, the awkward equality (13.122.5) can be replaced by the simpler equality

$$a \boxdot b = \operatorname{ch}_{\mathbb{C}} \left(\operatorname{Ind}_{\Omega} \left(\left(\operatorname{ch}_{\mathbb{C}}\right)^{-1} (a) \otimes \left(\operatorname{ch}_{\mathbb{C}}\right)^{-1} (b) \right) \right) . \tag{13.122.12}$$

This makes the definition of the maps $\overline{\Omega}$ and τ unnecessary. The WLOG assumption that we don't have $\min\{m, n\} = 0$ becomes unnecessary as well; the injectivity of the map cross can no longer be guaranteed without this assumption, but we do not need this map to be injective anymore. We no longer need Lemma 13.122.1(b), Lemma 13.122.3 and Lemma 13.122.4. However, we need to use Exercise 4.1.14(b) instead of (4.1.5), and we have to use Exercise 4.4.3(f) instead of Exercise 4.4.3(b).]

(d) We need to prove that $f \boxdot g \in \Lambda$ for any $f \in \Lambda$ and $g \in \Lambda$. Since the binary operation $\boxdot$ is $\mathbb{Z}$-bilinear (because it is $\mathbb{Q}$-bilinear), it is clearly enough to prove that $s_\mu \boxdot s_\nu \in \Lambda$ for any partitions μ and ν (since $(s_\lambda)_{\lambda \in \operatorname{Par}}$ is a $\mathbb{Z}$-basis of Λ). But this follows from the fact that

$$\begin{aligned} s_\mu \boxdot s_\nu &\in \sum_{\lambda \in \operatorname{Par}} \mathbb{N} s_\lambda \qquad \text{(by Exercise 4.4.9(c))} \\ &\subset \Lambda \end{aligned}$$

for any partitions μ and ν. Thus, Exercise 4.4.9(d) is solved.

13.123. **Solution to Exercise 4.6.4.** *Solution to Exercise 4.6.4.* (a) We shall prove the following fact:

Lemma 13.123.1. *Let q be a prime power. For every positive integer n, we have*

$$(\text{the number of all irreducible monic degree-}n\text{ polynomials in } \mathbb{F}_q[x]) = \frac{1}{n} \sum_{d \mid n} \mu\left(\frac{n}{d}\right) q^d .$$

Proof of Lemma 13.123.1. We first recall that every positive integer n satisfies

$$\sum_{d \mid n} \mu(d) = \delta_{n,1} . \tag{13.123.1}$$

(This is precisely the equality (13.84.3), which was proven in the solution of Exercise 2.9.6.)

For every positive integer n, define a nonnegative integer $\operatorname{irr} n$ by

$$\operatorname{irr} n = (\text{the number of all irreducible monic degree-}n\text{ polynomials in } \mathbb{F}_q[x]) . \tag{13.123.2}$$

Let $\mathfrak{P}$ denote the set of all irreducible monic polynomials in $\mathbb{F}_q[x]$. Then, every $p \in \mathfrak{P}$ is an irreducible polynomial and thus satisfies $\deg p \geq 1$. Also, the irreducible monic polynomials in $\mathbb{F}_q[x]$ are exactly the elements of $\mathfrak{P}$ (since $\mathfrak{P}$ is the set of all irreducible monic polynomials in $\mathbb{F}_q[x]$). Hence, for every positive integer n, we have

$$\begin{aligned} \operatorname{irr} n &= (\text{the number of all irreducible monic degree-}n\text{ polynomials in } \mathbb{F}_q[x]) \\ &= (\text{the number of all irreducible monic polynomials } p \text{ in } \mathbb{F}_q[x] \text{ such that } \deg p = n) \\ &= (\text{the number of all } p \in \mathfrak{P} \text{ such that } \deg p = n) \\ &\qquad (\text{since the irreducible monic polynomials in } \mathbb{F}_q[x] \text{ are exactly the elements of } \mathfrak{P}) . \end{aligned}$$

In other words, for every positive integer n, we have

$$(\text{the number of all } p \in \mathfrak{P} \text{ such that } \deg p = n) = \operatorname{irr} n . \tag{13.123.3}$$

Let $\mathfrak{N}$ be the set of all families $(k_p)_{p \in \mathfrak{P}} \in \mathbb{N}^{\mathfrak{P}}$ of nonnegative integers (indexed by the polynomials belonging to $\mathfrak{P}$) such that all but finitely many $p \in \mathfrak{P}$ satisfy $k_p = 0$. Every monic polynomial $P \in \mathbb{F}_q[x]$ has a unique factorization into irreducible monic polynomials. In other words, every monic polynomial $P \in \mathbb{F}_q[x]$ can be written in the form $P = \prod_{p\in\mathfrak{P}} p^{k_p}$ for a unique family $(k_p)_{p\in\mathfrak{P}} \in \mathfrak{N}$. Thus, the map

$$\begin{aligned} \mathfrak{N} &\to \{P \in \mathbb{F}_q[x] \mid P \text{ is monic}\}, \\ (k_p)_{p\in\mathfrak{P}} &\mapsto \prod_{p\in\mathfrak{P}} p^{k_p} \end{aligned}$$

is bijective. Thus, in the ring $\mathbb{Q}[[t]]$ of formal power series, we have

$$\sum_{\substack{P\in\mathbb{F}_q[x];\\ P \text{ is monic}}} t^{\deg P} = \sum_{(k_p)_{p\in\mathfrak{P}}\in\mathfrak{N}} t^{\deg\left(\prod_{p\in\mathfrak{P}} p^{k_p}\right)}. \tag{13.123.4}$$

But every $(k_p)_{p\in\mathfrak{P}} \in \mathfrak{N}$ satisfies

$$\begin{aligned} t^{\deg\left(\prod_{p\in\mathfrak{P}} p^{k_p}\right)} &= t^{\sum_{p\in\mathfrak{P}} k_p \deg p} \qquad \left(\text{since } \deg\left(\prod_{p\in\mathfrak{P}} p^{k_p}\right) = \sum_{p\in\mathfrak{P}} k_p \deg p\right) \\ &= \prod_{p\in\mathfrak{P}} t^{k_p \deg p} = \prod_{p\in\mathfrak{P}} \left(t^{\deg p}\right)^{k_p}. \end{aligned}$$

Thus, (13.123.4) becomes

$$\begin{aligned} \sum_{\substack{P\in\mathbb{F}_q[x];\\ P \text{ is monic}}} t^{\deg P} &= \sum_{(k_p)_{p\in\mathfrak{P}}\in\mathfrak{N}} \underbrace{t^{\deg\left(\prod_{p\in\mathfrak{P}} p^{k_p}\right)}}_{=\prod_{p\in\mathfrak{P}} \left(t^{\deg p}\right)^{k_p}} = \sum_{(k_p)_{p\in\mathfrak{P}}\in\mathfrak{N}} \prod_{p\in\mathfrak{P}} \left(t^{\deg p}\right)^{k_p} \\ &= \prod_{p\in\mathfrak{P}} \sum_{k\in\mathbb{N}} \left(t^{\deg p}\right)^{k} \qquad \text{(by the product rule)} \\ &= \prod_{n\geq 1} \prod_{\substack{p\in\mathfrak{P};\\ \deg p = n}} \sum_{k\in\mathbb{N}} \left(\underbrace{t^{\deg p}}_{\substack{=t^n \\ (\text{since } \deg p = n)}} \right)^{k} \qquad (\text{since } \deg p \geq 1 \text{ for every } p \in \mathfrak{P}) \\ &= \prod_{n\geq 1} \prod_{\substack{p\in\mathfrak{P};\\ \deg p = n}} \underbrace{\sum_{k\in\mathbb{N}} \left(t^n\right)^k}_{=\dfrac{1}{1-t^n}} = \prod_{n\geq 1} \underbrace{\underbrace{\prod_{\substack{p\in\mathfrak{P};\\ \deg p = n}} \frac{1}{1-t^n}}_{=\left(\dfrac{1}{1-t^n}\right)^{(\text{the number of all } p\in\mathfrak{P} \text{ such that } \deg p = n)}}}_{\substack{=\left(\dfrac{1}{1-t^n}\right)^{\operatorname{irr} n} \\ (\text{by (13.123.3)})}} \\ &= \prod_{n\geq 1} \left(\frac{1}{1-t^n}\right)^{\operatorname{irr} n}. \end{aligned}$$

Compared with

$$\sum_{\substack{P\in\mathbb{F}_q[x];\\ P\text{ is monic}}} t^{\deg P} = \sum_{n\in\mathbb{N}} \sum_{\substack{P\in\mathbb{F}_q[x];\\ P\text{ is monic;}\\ \deg P=n}} \underbrace{t^{\deg P}}_{\substack{=t^n\\ (\text{since }\deg P=n)}} = \sum_{n\in\mathbb{N}} \underbrace{\sum_{\substack{P\in\mathbb{F}_q[x];\\ P\text{ is monic;}\\ \deg P=n}} t^n}_{=(\text{the number of all monic }P\in\mathbb{F}_q[x]\text{ such that }\deg P=n)\cdot t^n}$$

$$= \sum_{n\in\mathbb{N}} \underbrace{(\text{the number of all monic } P\in\mathbb{F}_q[x] \text{ such that } \deg P = n)}_{\substack{=q^n\\ (\text{because specifying a monic }P\in\mathbb{F}_q[x]\text{ such that }\deg P=n\\ \text{is equivalent to specifying its coefficients before }x^0,\ x^1,\ \ldots,\ x^{n-1},\\ \text{and each of these coefficients can be chosen freely from }q\\ \text{possible values})}} \cdot t^n$$

$$= \sum_{n\in\mathbb{N}} q^n t^n = \sum_{n\in\mathbb{N}} (qt)^n = \frac{1}{1-qt},$$

this yields

$$\frac{1}{1-qt} = \prod_{n\geq 1} \left(\frac{1}{1-t^n}\right)^{\operatorname{irr} n}.$$

Taking the logarithm of both sides of this identity, we obtain

$$\log\frac{1}{1-qt} = \log\left(\prod_{n\geq 1}\left(\frac{1}{1-t^n}\right)^{\operatorname{irr} n}\right) = \sum_{n\geq 1} (\operatorname{irr} n)\cdot \underbrace{\log\left(\frac{1}{1-t^n}\right)}_{\substack{=-\log(1-t^n)=\sum_{u\geq 1}\frac{1}{u}(t^n)^u\\ (\text{by the Mercator series for the logarithm})}}$$

$$= \sum_{n\geq 1} (\operatorname{irr} n)\cdot \sum_{u\geq 1} \frac{1}{u}(t^n)^u = \sum_{n\geq 1}\sum_{u\geq 1} (\operatorname{irr} n)\frac{1}{u}\underbrace{(t^n)^u}_{=t^{nu}} = \sum_{n\geq 1}\sum_{u\geq 1}(\operatorname{irr} n)\frac{1}{u}t^{nu}$$

$$= \underbrace{\sum_{n\geq 1}\sum_{\substack{v\geq 1;\\ n\mid v}}}_{=\sum_{v\geq 1}\sum_{n\mid v}} (\operatorname{irr} n)\underbrace{\frac{1}{v/n}}_{=\frac{n}{v}} t^v \qquad (\text{here, we substituted } v/n \text{ for } u \text{ in the second sum})$$

$$= \sum_{v\geq 1}\sum_{n\mid v}(\operatorname{irr} n)\frac{n}{v}t^v = \sum_{n\geq 1}\sum_{d\mid n}(\operatorname{irr} d)\frac{d}{n}t^n$$

(here, we renamed the summation indices v and n as n and d). Since

$$\log\frac{1}{1-qt} = -\log(1-qt) = \sum_{n\geq 1}\frac{1}{n}(qt)^n \qquad (\text{by the Mercator series for the logarithm})$$
$$= \sum_{n\geq 1}\frac{1}{n}q^n t^n,$$

this rewrites as

$$\sum_{n\geq 1}\frac{1}{n}q^n t^n = \sum_{n\geq 1}\sum_{d\mid n}(\operatorname{irr} d)\frac{d}{n}t^n.$$

Comparing coefficients, we conclude that every positive integer n satisfies

$$\frac{1}{n}q^n = \sum_{d\mid n}(\operatorname{irr} d)\frac{d}{n}.$$

Multiplying this with n, we obtain

$$q^n = \sum_{d\mid n}(\operatorname{irr} d)\, d. \tag{13.123.5}$$

Now, every positive integer n satisfies

$$\begin{aligned}\sum_{d|n}\mu(d)\,q^{n/d} &= \sum_{e|n}\mu(e)\underbrace{q^{n/e}}_{\substack{=\sum_{d|n/e}(\operatorname{irr} d)d \\ \text{(by (13.123.5), applied} \\ \text{to } n/e \text{ instead of } n)}} = \sum_{e|n}\mu(e)\sum_{d|n/e}(\operatorname{irr} d)\,d \\ &= \underbrace{\sum_{e|n}\sum_{d|n/e}}_{=\sum_{d|n}\sum_{e|n/d}}\mu(e)(\operatorname{irr} d)\,d = \sum_{d|n}\underbrace{\sum_{e|n/d}\mu(e)}_{\substack{=\delta_{n/d,1} \\ \text{(by (13.123.1), applied} \\ \text{to } n/d \text{ instead of } n)}}(\operatorname{irr} d)\,d \\ &= \sum_{d|n}\underbrace{\delta_{n/d,1}}_{=\delta_{n,d}}(\operatorname{irr} d)\,d = \sum_{d|n}\delta_{n,d}(\operatorname{irr} d)\,d = (\operatorname{irr} n)\,n.\end{aligned}$$

Dividing this by n, we obtain $\frac{1}{n}\sum_{d|n}\mu(d)\,q^{n/d} = \operatorname{irr} n$. Now, (13.123.2) yields

$$(\text{the number of all irreducible monic degree-}n\text{ polynomials in }\mathbb{F}_q[x]) = \operatorname{irr} n = \frac{1}{n}\sum_{d|n}\mu(d)\,q^{n/d}.$$

This proves Lemma 13.123.1. □

Now, let us solve Exercise 4.6.4(a). Let $n \geq 2$ be an integer. We know that $|\mathcal{F}_n|$ is the number of irreducible monic degree-n polynomials $f(x) \neq x$ in $\mathbb{F}_q[x]$ with nonzero constant term. Since the irreducible monic degree-n polynomials $f(x) \neq x$ in $\mathbb{F}_q[x]$ with nonzero constant term are precisely the irreducible monic degree-n polynomials in $\mathbb{F}_q[x]$ [895], this statement rewrites as follows: $|\mathcal{F}_n|$ is the number of irreducible monic degree-n polynomials in $\mathbb{F}_q[x]$. In other words,

$$\begin{aligned}|\mathcal{F}_n| &= (\text{the number of all irreducible monic degree-}n\text{ polynomials in }\mathbb{F}_q[x]) \\ &= \frac{1}{n}\sum_{d|n}\mu\left(\frac{n}{d}\right)q^d \qquad \text{(by Lemma 13.123.1)}.\end{aligned}$$

This solves Exercise 4.6.4(a).

(b) We shall use the results of Exercise 6.1.34 (which provides a more systematic introduction to necklaces).

Let $\mathfrak{A}$ be the set $\mathbb{F}_q$. Consider the notion of an n-necklace defined in Exercise 6.1.34. The "n-necklaces" (as defined in Exercise 6.1.34) are precisely the "necklaces with n beads of q colors" (as defined in Exercise 4.6.4(b))[896]. Moreover, it is easy to see that the "aperiodic n-necklaces" (as defined in Exercise 6.1.34) are

[895] *Proof.* Let f be an irreducible polynomial monic degree-n polynomial in $\mathbb{F}_q[x]$. We will now show that $f(x) \neq x$ and that f has nonzero constant term.

Since f is a degree-n polynomial, we have $\deg f = n \geq 2 > 1 = \deg x$, thus $\deg f \neq \deg x$ and therefore $f \neq x$. In other words, $f(x) \neq x$.

Now, let us assume (for the sake of contradiction) that the constant term of f is zero. Then, the polynomial x divides f in $\mathbb{F}_q[x]$. But since f is irreducible, the polynomial f is a scalar multiple of every non-constant polynomial which divides f. In particular, f is a scalar multiple of x (since x is a non-constant polynomial which divides f). Consequently, $\deg f = \deg x$, which contradicts $\deg f \neq \deg x$. This contradiction proves that our assumption (that the constant term of f is zero) was wrong. Hence, the constant term of f is nonzero. In other words, f has nonzero constant term.

Now, let us forget that we fixed f. We thus have proven that every irreducible monic degree-n polynomial f in $\mathbb{F}_q[x]$ satisfies $f(x) \neq x$ and has nonzero constant term. Thus, all irreducible monic degree-n polynomials in $\mathbb{F}_q[x]$ are irreducible monic degree-n polynomials $f(x) \neq x$ in $\mathbb{F}_q[x]$ with nonzero constant term. Combining this statement with the (obvious) converse statement (which states that all irreducible monic degree-n polynomials $f(x) \neq x$ in $\mathbb{F}_q[x]$ with nonzero constant term are irreducible monic degree-n polynomials in $\mathbb{F}_q[x]$), we conclude that the irreducible monic degree-n polynomials $f(x) \neq x$ in $\mathbb{F}_q[x]$ with nonzero constant term are precisely the irreducible monic degree-n polynomials in $\mathbb{F}_q[x]$. Qed.

[896] *Proof.* Consider the group C, its generator c and the action of C on $\mathfrak{A}^n$ which are defined in Exercise 6.1.34. Then, the "n-necklaces" (as defined in Exercise 6.1.34) are the orbits of the C-action on $\mathfrak{A}^n$. In other words, the "n-necklaces" (as defined in Exercise 6.1.34) are the equivalence classes of n-tuples $(a_1, a_2, ..., a_n) \in \mathfrak{A}^n$ with respect to cyclic rotation (because C acts on $\mathfrak{A}^n$ by cyclic rotation). But the same can be said about the "necklaces with n beads of q colors" (as defined in Exercise 4.6.4(b)). Thus, the "n-necklaces" (as defined in Exercise 6.1.34) are precisely the "necklaces with n beads of q colors" (as defined in Exercise 4.6.4(b)), qed.

precisely the "primitive necklaces with n beads of q colors" (as defined in Exercise 4.6.4(b))[897]. Hence,

$$\begin{aligned}&(\text{the number of all aperiodic } n\text{-necklaces})\\&= (\text{the number of all primitive necklaces with } n \text{ beads of } q \text{ colors}),\end{aligned}$$

so that

$$\begin{aligned}&(\text{the number of all primitive necklaces with } n \text{ beads of } q \text{ colors})\\&= (\text{the number of all aperiodic } n\text{-necklaces}) = \frac{1}{n}\sum_{d\mid n}\mu\left(d\right)\underbrace{\left|\mathfrak{A}\right|^{n/d}}_{\substack{=q^{n/d}\\ (\text{since } \left|\mathfrak{A}\right|=q)}} \qquad (\text{by Exercise 6.1.34(f)})\\&= \frac{1}{n}\sum_{d\mid n}\mu\left(d\right)q^{n/d}.\end{aligned}$$

This solves Exercise 4.6.4(b).

13.124. **Solution to Exercise 4.9.6.** *Solution to Exercise 4.9.6.* Let us first notice that any $k \in \mathbb{N}$ and any k partitions $\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(k)}$ satisfy

$$\underline{1}_{J_{\lambda^{(1)}}}\underline{1}_{J_{\lambda^{(2)}}}\cdots\underline{1}_{J_{\lambda^{(k)}}} = \sum_{\lambda\in\operatorname{Par}} g^{\lambda}_{\lambda^{(1)},\lambda^{(2)},\ldots,\lambda^{(k)}}\left(q\right)\underline{1}_{J_\lambda} = \sum_{\mu\in\operatorname{Par}} g^{\mu}_{\lambda^{(1)},\lambda^{(2)},\ldots,\lambda^{(k)}}\left(q\right)\underline{1}_{J_\mu} \tag{13.124.1}$$

(here, we renamed the summation index λ as μ). Any two partitions μ and ν satisfy

$$\underline{1}_{J_\mu}\underline{1}_{J_\nu} = \sum_{\lambda\in\operatorname{Par}} g^{\lambda}_{\mu,\nu}\left(q\right)\underline{1}_{J_\lambda} = \sum_{\tau\in\operatorname{Par}} g^{\tau}_{\mu,\nu}\left(q\right)\underline{1}_{J_\tau} \tag{13.124.2}$$

(here, we renamed the summation index λ as τ).

(a) We shall prove the statement of Exercise 4.9.6(a) by induction over k. The base cases ($k = 0$ and $k = 1$) are left to the reader. We will now handle the induction step. So let us solve Exercise 4.9.6(a) for some positive integer $k > 1$, assuming (as the induction hypothesis) that Exercise 4.9.6(a) is already solved for $k - 1$ instead of k.

From (13.124.1), we obtain

$$\begin{aligned}&\sum_{\mu\in\operatorname{Par}} g^{\mu}_{\lambda^{(1)},\lambda^{(2)},\ldots,\lambda^{(k)}}\left(q\right)\underline{1}_{J_\mu}\\&= \underline{1}_{J_{\lambda^{(1)}}}\underline{1}_{J_{\lambda^{(2)}}}\cdots\underline{1}_{J_{\lambda^{(k)}}} = \underbrace{\underline{1}_{J_{\lambda^{(1)}}}\underline{1}_{J_{\lambda^{(2)}}}\cdots\underline{1}_{J_{\lambda^{(k-1)}}}}_{\substack{=\sum_{\mu\in\operatorname{Par}} g^{\mu}_{\lambda^{(1)},\lambda^{(2)},\ldots,\lambda^{(k-1)}}(q)\underline{1}_{J_\mu}\\ (\text{by (13.124.1), applied to } k-1 \text{ instead of } k)}}\cdot\underline{1}_{J_{\lambda^{(k)}}}\\&= \left(\sum_{\mu\in\operatorname{Par}} g^{\mu}_{\lambda^{(1)},\lambda^{(2)},\ldots,\lambda^{(k-1)}}\left(q\right)\underline{1}_{J_\mu}\right)\cdot\underline{1}_{J_{\lambda^{(k)}}} = \sum_{\mu\in\operatorname{Par}} g^{\mu}_{\lambda^{(1)},\lambda^{(2)},\ldots,\lambda^{(k-1)}}\left(q\right)\underbrace{\underline{1}_{J_\mu}\underline{1}_{J_{\lambda^{(k)}}}}_{\substack{=\sum_{\tau\in\operatorname{Par}} g^{\tau}_{\mu,\lambda^{(k)}}(q)\underline{1}_{J_\tau}\\ (\text{by (13.124.2), applied to } \nu=\lambda^{(k)})}}\\&= \sum_{\mu\in\operatorname{Par}} g^{\mu}_{\lambda^{(1)},\lambda^{(2)},\ldots,\lambda^{(k-1)}}\left(q\right)\sum_{\tau\in\operatorname{Par}} g^{\tau}_{\mu,\lambda^{(k)}}\left(q\right)\underline{1}_{J_\tau} = \sum_{\tau\in\operatorname{Par}}\sum_{\mu\in\operatorname{Par}} g^{\mu}_{\lambda^{(1)},\lambda^{(2)},\ldots,\lambda^{(k-1)}}\left(q\right)g^{\tau}_{\mu,\lambda^{(k)}}\left(q\right)\underline{1}_{J_\tau}.\end{aligned}$$

Comparing coefficients in front of $\underline{1}_{J_\lambda}$ on both sides of this equality, we obtain

$$g^{\lambda}_{\lambda^{(1)},\lambda^{(2)},\ldots,\lambda^{(k)}}\left(q\right) = \sum_{\mu\in\operatorname{Par}} g^{\mu}_{\lambda^{(1)},\lambda^{(2)},\ldots,\lambda^{(k-1)}}\left(q\right)g^{\lambda}_{\mu,\lambda^{(k)}}\left(q\right). \tag{13.124.3}$$

Our $\mathbb{F}_q$-vector space V is n-dimensional. Hence, we can WLOG assume that $V = \mathbb{F}_q^n$ (because nothing changes if we map V isomorphically to $\mathbb{F}_q^n$ and change g accordingly). Assume this.

[897] This follows from Exercise 6.1.34(b) (because for an n-tuple $(w_1, w_2, ..., w_n) \in \mathfrak{A}^n$, the statement that no nontrivial rotation (in $\mathbb{Z}/n\mathbb{Z}$) fixes w is equivalent to the statement that every $k \in \{1, 2, \ldots, n-1\}$ satisfies $(w_{k+1}, w_{k+2}, \ldots, w_n, w_1, w_2, \ldots, w_k) \neq w$).

In the following, an *ender* will mean a g-stable $\mathbb{F}_q$-vector subspace $W \subset \mathbb{F}_q^n$ for which the induced map $\bar{g}$ on the quotient space $\mathbb{F}_q^n / W$ has Jordan type $\lambda^{(k)}$. Notice that if $0 = V_0 \subset V_1 \subset V_2 \subset \cdots \subset V_k = V$ is a $\left(\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(k)}\right)$-compatible g-flag, then V_{k-1} is an ender (because the definition of a $\left(\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(k)}\right)$-compatible g-flag shows that the endomorphism of $\underbrace{\mathbb{F}_q^n}_{=V=V_k} / V_{k-1} = V_k / V_{k-1}$ induced by g has Jordan type $\lambda^{(k)}$). (This is why we have chosen the name "ender" – an ender is the last proper subspace in a $\left(\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(k)}\right)$-compatible g-flag.)

Whenever h is a unipotent endomorphism of a finite-dimensional vector space, we let $\operatorname{type} h$ denote the Jordan type of h.

Proposition 4.9.4 (applied to $\nu = \lambda^{(k)}$) yields that for every $\mu \in \operatorname{Par}$, the number $g^{\lambda}_{\mu, \lambda^{(k)}}(q)$ counts the g-stable $\mathbb{F}_q$-subspaces $W \subset \mathbb{F}_q^n$ for which the restriction $g|W$ acts with Jordan type μ, and the induced map $\bar{g}$ on the quotient space $\mathbb{F}_q^n / W$ has Jordan type $\lambda^{(k)}$ [898]. In other words, for every $\mu \in \operatorname{Par}$, the number $g^{\lambda}_{\mu, \lambda^{(k)}}(q)$ counts the enders W for which the restriction $g|W$ acts with Jordan type μ. In other words, for every $\mu \in \operatorname{Par}$, the number $g^{\lambda}_{\mu, \lambda^{(k)}}(q)$ counts the enders W for which $\operatorname{type}(g|W) = \mu$. In other words, for every $\mu \in \operatorname{Par}$, we have

$$g^{\lambda}_{\mu, \lambda^{(k)}}(q) = \sum_{\substack{W \text{ is an ender;} \\ \operatorname{type}(g|W) = \mu}} 1.$$

Hence, (13.124.3) becomes

$$g^{\lambda}_{\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(k)}}(q) = \sum_{\mu \in \operatorname{Par}} g^{\mu}_{\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(k-1)}}(q) \underbrace{g^{\lambda}_{\mu, \lambda^{(k)}}(q)}_{= \sum_{\substack{W \text{ is an ender;} \\ \operatorname{type}(g|W) = \mu}} 1} = \sum_{\mu \in \operatorname{Par}} g^{\mu}_{\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(k-1)}}(q) \sum_{\substack{W \text{ is an ender;} \\ \operatorname{type}(g|W) = \mu}} 1$$

$$= \sum_{\mu \in \operatorname{Par}} \sum_{\substack{W \text{ is an ender;} \\ \operatorname{type}(g|W) = \mu}} g^{\mu}_{\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(k-1)}}(q) = \sum_{W \text{ is an ender}} g^{\operatorname{type}(g|W)}_{\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(k-1)}}(q). \tag{13.124.4}$$

Now, let us fix an ender W. By the definition of an ender, this W is a g-stable $\mathbb{F}_q$-subspace of $\mathbb{F}_q^n$ for which the induced map $\bar{g}$ on the quotient space $\mathbb{F}_q^n / W$ has Jordan type $\lambda^{(k)}$. Let $m = \dim W$.

By the induction hypothesis, we can apply Exercise 4.9.6(a) to $\operatorname{type}(g|W)$, m, W, $g|W$, $k-1$ and $\left(\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(k-1)}\right)$ instead of λ, n, V, g, k and $\left(\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(k)}\right)$. As a result, we conclude that

$$g^{\operatorname{type}(g|W)}_{\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(k-1)}}(q) \text{ is the number of } \left(\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(k-1)}\right)\text{-compatible } g|W\text{-flags.} \tag{13.124.5}$$

Now, forget that we fixed W. Combining (13.124.4) with (13.124.5), we see that $g^{\lambda}_{\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(k)}}(q)$ is the number of all pairs $(W, \mathbf{F})$, where W is an ender and $\mathbf{F}$ is a $\left(\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(k-1)}\right)$-compatible $g|W$-flag. But (in order to complete the induction step) we have to prove that $g^{\lambda}_{\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(k)}}(q)$ is the number of $\left(\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(k)}\right)$-compatible g-flags. So an obvious way to prove this would be to find a bijection between $\left(\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(k)}\right)$-compatible g-flags and pairs $(W, \mathbf{F})$, where W is an ender and $\mathbf{F}$ is a $\left(\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(k-1)}\right)$-compatible $g|W$-flag. This bijection is very easy to define: It sends a $\left(\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(k)}\right)$-compatible g-flag $0 = V_0 \subset V_1 \subset V_2 \subset \cdots \subset V_k = V$ to the pair $(V_{k-1}, \mathbf{F})$, where $\mathbf{F}$ is the $\left(\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(k-1)}\right)$-compatible $g|V_{k-1}$-flag $0 = V_0 \subset V_1 \subset V_2 \subset \cdots \subset V_{k-1} = V_{k-1}$. Showing that this bijection is well-defined and bijective is completely straightforward[899]; thus, the induction step is complete, and the solution of Exercise 4.9.6(a) is finished.

(b) We shall prove the statement of Exercise 4.9.6(b) by induction over k. The base cases ($k = 0$ and $k = 1$) are left to the reader. We will now handle the induction step. So let us solve Exercise 4.9.6(b) for some positive integer $k > 1$, assuming (as the induction hypothesis) that Exercise 4.9.6(b) is already solved for $k - 1$ instead of k.

[898] Note that the variable that I am calling W here has been denoted by V in Proposition 4.9.4.

[899] The inverse map takes a pair $(W, \mathbf{F})$, where W is an ender and $\mathbf{F}$ is a $\left(\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(k-1)}\right)$-compatible $g|W$-flag $0 = V_0 \subset V_1 \subset V_2 \subset \cdots \subset V_{k-1} = W$, and maps it to the $\left(\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(k)}\right)$-compatible g-flag $0 = V_0 \subset V_1 \subset V_2 \subset \cdots \subset V_{k-1} \subset V = V$.

We need to prove that $g^{\lambda}_{\lambda^{(1)},\lambda^{(2)},\ldots,\lambda^{(k)}}(q) = 0$ unless $\left|\lambda^{(1)}\right| + \left|\lambda^{(2)}\right| + \cdots + \left|\lambda^{(k)}\right| = |\lambda|$ and $\lambda^{(1)} + \lambda^{(2)} + \cdots + \lambda^{(k)} \triangleright \lambda$. In other words, we need to prove that if $g^{\lambda}_{\lambda^{(1)},\lambda^{(2)},\ldots,\lambda^{(k)}}(q) \neq 0$, then

$$\left|\lambda^{(1)}\right| + \left|\lambda^{(2)}\right| + \cdots + \left|\lambda^{(k)}\right| = |\lambda| \qquad \text{and} \qquad \lambda^{(1)} + \lambda^{(2)} + \cdots + \lambda^{(k)} \triangleright \lambda. \tag{13.124.6}$$

So let us assume that $g^{\lambda}_{\lambda^{(1)},\lambda^{(2)},\ldots,\lambda^{(k)}}(q) \neq 0$. Due to (13.124.3), this rewrites as

$$\sum_{\mu \in \operatorname{Par}} g^{\mu}_{\lambda^{(1)},\lambda^{(2)},\ldots,\lambda^{(k-1)}}(q)\, g^{\lambda}_{\mu,\lambda^{(k)}}(q) \neq 0.$$

Hence, there exists a $\mu \in \operatorname{Par}$ satisfying $g^{\mu}_{\lambda^{(1)},\lambda^{(2)},\ldots,\lambda^{(k-1)}}(q) \neq 0$ and $g^{\lambda}_{\mu,\lambda^{(k)}}(q) \neq 0$. Consider this μ.

By the induction hypothesis, we can apply Exercise 4.9.6(b) to μ, $k-1$ and $\left(\lambda^{(1)}, \lambda^{(2)}, ..., \lambda^{(k-1)}\right)$ instead of λ, k and $\left(\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(k)}\right)$. Thus, we conclude that $g^{\mu}_{\lambda^{(1)},\lambda^{(2)},\ldots,\lambda^{(k-1)}}(q) = 0$ unless $\left|\lambda^{(1)}\right| + \left|\lambda^{(2)}\right| + \cdots + \left|\lambda^{(k-1)}\right| = |\mu|$ and $\lambda^{(1)} + \lambda^{(2)} + \cdots + \lambda^{(k-1)} \triangleright \mu$. Since we have $g^{\mu}_{\lambda^{(1)},\lambda^{(2)},\ldots,\lambda^{(k-1)}}(q) \neq 0$, we therefore must have $\left|\lambda^{(1)}\right| + \left|\lambda^{(2)}\right| + \cdots + \left|\lambda^{(k-1)}\right| = |\mu|$ and $\lambda^{(1)} + \lambda^{(2)} + \cdots + \lambda^{(k-1)} \triangleright \mu$.

But let $n = |\lambda|$. Fix a unipotent endomorphism g of $\mathbb{F}_q^n$ having Jordan type λ (such a g clearly exists). The number $g^{\lambda}_{\mu,\lambda^{(k)}}(q)$ counts the g-stable $\mathbb{F}_q$-subspaces $W \subset \mathbb{F}_q^n$ for which the restriction $g|W$ acts with Jordan type μ, and the induced map $\bar{g}$ on the quotient space $\mathbb{F}_q^n/W$ has Jordan type $\lambda^{(k)}$ (by Proposition 4.9.4, applied to $\nu = \lambda^{(k)}$). Since $g^{\lambda}_{\mu,\lambda^{(k)}}(q) \neq 0$, this yields that there exists such a subspace W. If we define a nilpotent endomorphism f of $\mathbb{F}_q^n$ by $f = g - \operatorname{id}_{\mathbb{F}_q^n}$ (this is indeed nilpotent because g is unipotent), then this W is also an f-stable $\mathbb{F}_q$-subspace of $\mathbb{F}_q^n$ (because any g-stable subspace of $\mathbb{F}_q^n$ is f-stable) for which the restriction $f|W$ acts with Jordan type μ (since $f|W = g|W - \operatorname{id}_W$), and the induced map $\overline{f}$ on the quotient space $\mathbb{F}_q^n/W$ has Jordan type $\lambda^{(k)}$ (since $\overline{f} = \overline{g} - \operatorname{id}_{\mathbb{F}_q^n/W}$). Therefore, Exercise 2.9.22(b) (applied to $\mathbb{K} = \mathbb{F}_q$, $V = \mathbb{F}_q^n$, $U = W$ and $\nu = \lambda^{(k)}$) yields $c^{\lambda}_{\mu,\lambda^{(k)}} \neq 0$. Consequently, $|\lambda| = |\mu| + \left|\lambda^{(k)}\right|$. Exercise 2.9.17(c) (applied to $|\lambda|$, $|\mu|$ and $\lambda^{(k)}$ instead of n, k and ν) thus yields $\mu + \lambda^{(k)} \triangleright \lambda \triangleright \mu \sqcup \lambda^{(k)}$.

Exercise 2.9.17(d) (applied to $|\mu|$, $\left|\lambda^{(k)}\right|$, $\lambda^{(1)} + \lambda^{(2)} + \cdots + \lambda^{(k-1)}$, μ, $\lambda^{(k)}$ and $\lambda^{(k)}$ instead of n, m, α, β, γ and δ) yields

$$\left(\lambda^{(1)} + \lambda^{(2)} + \cdots + \lambda^{(k-1)}\right) + \lambda^{(k)} \triangleright \mu + \lambda^{(k)}$$

(since $\lambda^{(1)} + \lambda^{(2)} + \cdots + \lambda^{(k-1)} \triangleright \mu$ and $\lambda^{(k)} \triangleright \lambda^{(k)}$).

Now,

$$\left|\lambda^{(1)}\right| + \left|\lambda^{(2)}\right| + \cdots + \left|\lambda^{(k)}\right| = \underbrace{\left|\lambda^{(1)}\right| + \left|\lambda^{(2)}\right| + \cdots + \left|\lambda^{(k-1)}\right|}_{=|\mu|} + \left|\lambda^{(k)}\right| = |\mu| + \left|\lambda^{(k)}\right| = |\lambda|$$

and

$$\lambda^{(1)} + \lambda^{(2)} + \cdots + \lambda^{(k)} = \left(\lambda^{(1)} + \lambda^{(2)} + \cdots + \lambda^{(k-1)}\right) + \lambda^{(k)} \triangleright \mu + \lambda^{(k)} \triangleright \lambda.$$

Thus, we have proven (13.124.6), and so the induction step is complete. We thus have solved Exercise 4.9.6(b).

(c) Let $n = |\lambda|$. Fix a unipotent endomorphism g of $\mathbb{F}_q^n$ having Jordan type λ (such a g clearly exists).

Exercise 4.9.6(a) (applied to $V = \mathbb{F}_q^n$, $k = \ell$ and $\lambda^{(i)} = \left(1^{(\lambda^t)_i}\right)$) shows that $g^{\lambda}_{\left(1^{(\lambda^t)_1}\right),\left(1^{(\lambda^t)_2}\right),\ldots,\left(1^{(\lambda^t)_\ell}\right)}(q)$ is the number of $\left(\left(1^{(\lambda^t)_1}\right), \left(1^{(\lambda^t)_2}\right), \ldots, \left(1^{(\lambda^t)_\ell}\right)\right)$-compatible g-flags. Hence, in order to prove $g^{\lambda}_{\left(1^{(\lambda^t)_1}\right),\left(1^{(\lambda^t)_2}\right),\ldots,\left(1^{(\lambda^t)_\ell}\right)}(q) \neq 0$ (and thus, to solve Exercise 4.9.6(c)), it will be enough to prove that there exists at least one $\left(\left(1^{(\lambda^t)_1}\right), \left(1^{(\lambda^t)_2}\right), \ldots, \left(1^{(\lambda^t)_\ell}\right)\right)$-compatible g-flag.

Let $f = g - \operatorname{id}_{\mathbb{F}_q^n}$. Then, f is a nilpotent endomorphism of $\mathbb{F}_q^n$ having Jordan type λ (since g is a unipotent endomorphism of $\mathbb{F}_q^n$ having Jordan type λ). Also, $g = f + \operatorname{id}_{\mathbb{F}_q^n}$ (since $f = g - \operatorname{id}_{\mathbb{F}_q^n}$).

Let us define a sequence $V_0 \subset V_1 \subset V_2 \subset \cdots \subset V_\ell$ of $\mathbb{F}_q$-vector subspaces of $\mathbb{F}_q^n$ by setting

$$V_i = \ker\left(f^i\right) \qquad \text{for every } i \in \{0, 1, ..., \ell\}.$$

Then, $V_0 = \ker \left(\underbrace{f^0}_{=\mathrm{id}} \right) = \ker \mathrm{id} = 0$. Also, it follows readily from Exercise 2.9.22(a) that

$$\dim \left(\ker \left(f^k \right) \right) = \left(\lambda^t \right)_1 + \left(\lambda^t \right)_2 + \ldots + \left(\lambda^t \right)_k \qquad \text{for every } k \in \mathbb{N}. \tag{13.124.7}$$

[900] Applying this to $k = \ell$, we obtain

$$\begin{aligned} \dim \left(\ker \left(f^\ell \right) \right) &= \left(\lambda^t \right)_1 + \left(\lambda^t \right)_2 + \ldots + \left(\lambda^t \right)_\ell = \left| \lambda^t \right| \qquad \left(\text{since } \lambda^t = \left(\left(\lambda^t \right)_1, \left(\lambda^t \right)_2, \ldots, \left(\lambda^t \right)_\ell \right) \right) \\ &= |\lambda| = n = \dim \left(\mathbb{F}_q^n \right) \end{aligned}$$

which yields $\ker \left(f^\ell \right) = \mathbb{F}_q^n$ (since $\ker \left(f^\ell \right) \subset \mathbb{F}_q^n$). Thus, $V_\ell = \ker \left(f^\ell \right) = \mathbb{F}_q^n$. Since $V_0 = 0$ and $V_\ell = \mathbb{F}_q^n$, our sequence $V_0 \subset V_1 \subset V_2 \subset \cdots \subset V_\ell$ thus can be written as $0 = V_0 \subset V_1 \subset V_2 \subset \cdots \subset V_\ell = \mathbb{F}_q^n$.

For every $i \in \{0, 1, ..., \ell\}$, the $\mathbb{F}_q$-vector subspace V_i of $\mathbb{F}_q^n$ is g-invariant (since $V_i = \ker \left(f^i \right)$ is f-invariant and thus g-invariant). Also, for every $i \in \{1, 2, ..., \ell\}$, we have $\underbrace{V_i}_{=\ker(f^i)} / \underbrace{V_{i-1}}_{=\ker(f^{i-1})} = \ker \left(f^i \right) / \ker \left(f^{i-1} \right)$; thus, the endomorphism of V_i / V_{i-1} induced by f is the zero map, and therefore the endomorphism of V_i / V_{i-1} induced by g is the identity map (since $g = f + \mathrm{id}_{\mathbb{F}_q^n}$). Hence, for every $i \in \{1, 2, ..., \ell\}$, this latter endomorphism has Jordan type $\left(1^{\dim(V_i / V_{i-1})} \right)$. But since every $i \in \{1, 2, ..., \ell\}$ satisfies

$$\begin{aligned} \dim \left(\underbrace{V_i / V_{i-1}}_{=\ker(f^i) / \ker(f^{i-1})} \right) &= \dim \left(\ker \left(f^i \right) / \ker \left(f^{i-1} \right) \right) = \underbrace{\dim \left(\ker \left(f^i \right) \right)}_{\substack{=(\lambda^t)_1 + (\lambda^t)_2 + \ldots + (\lambda^t)_i \\ \text{(by (13.124.7),} \\ \text{applied to } k=i)}} - \underbrace{\dim \left(\ker \left(f^{i-1} \right) \right)}_{\substack{=(\lambda^t)_1 + (\lambda^t)_2 + \ldots + (\lambda^t)_{i-1} \\ \text{(by (13.124.7),} \\ \text{applied to } k=i-1)}} \\ &= \left(\left(\lambda^t \right)_1 + \left(\lambda^t \right)_2 + \ldots + \left(\lambda^t \right)_i \right) - \left(\left(\lambda^t \right)_1 + \left(\lambda^t \right)_2 + \ldots + \left(\lambda^t \right)_{i-1} \right) = \left(\lambda^t \right)_i, \end{aligned}$$

this rewrites as follows: For every $i \in \{1, 2, ..., \ell\}$, the endomorphism of V_i / V_{i-1} induced by g has Jordan type $\left(1^{(\lambda^t)_i} \right)$.

Altogether, we now know that $0 = V_0 \subset V_1 \subset V_2 \subset \cdots \subset V_\ell = \mathbb{F}_q^n$ is a sequence of g-invariant $\mathbb{F}_q$-vector subspaces V_i of $\mathbb{F}_q^n$ such that for every $i \in \{1, 2, \ldots, \ell\}$, the endomorphism of V_i / V_{i-1} induced by g has Jordan type $\left(1^{(\lambda^t)_i} \right)$. In other words, $0 = V_0 \subset V_1 \subset V_2 \subset \cdots \subset V_\ell = \mathbb{F}_q^n$ is a $\left(\left(1^{(\lambda^t)_1} \right), \left(1^{(\lambda^t)_2} \right), \ldots, \left(1^{(\lambda^t)_\ell} \right) \right)$-compatible g-flag (according the definition of the latter notion). Hence, there exists at least one $\left(\left(1^{(\lambda^t)_1} \right), \left(1^{(\lambda^t)_2} \right), \ldots, \left(1^{(\lambda^t)_\ell} \right) \right)$-compatible g-flag. Exercise 4.9.6(c) is solved.

(d) Write the partition λ as $\lambda = (\lambda_1, \lambda_2, ..., \lambda_\ell)$ with $\ell = \ell(\lambda)$. Then, $\left(\lambda^t \right)^t = \lambda = (\lambda_1, \lambda_2, ..., \lambda_\ell)$. Hence, Exercise 4.9.6(c) (applied to λ^t instead of λ) yields

$$g^{\lambda^t}_{(1^{\lambda_1}), (1^{\lambda_2}), \ldots, (1^{\lambda_\ell})}(q) \neq 0. \tag{13.124.8}$$

[900] *Proof of (13.124.7):* Let $k \in \mathbb{N}$. Let $N \in \mathbb{F}_q^{n \times n}$ be the matrix representing the endomorphism f of $\mathbb{F}_q^n$. Then, N is nilpotent (since f is nilpotent) and has Jordan type λ (since f has Jordan type λ), and thus satisfies $\dim \left(\ker \left(N^k \right) \right) = \left(\lambda^t \right)_1 + \left(\lambda^t \right)_2 + \ldots + \left(\lambda^t \right)_k$ (by Exercise 2.9.22(a)). But since N is a matrix representing the map f, we have $\dim \left(\ker \left(N^k \right) \right) = \dim \left(\ker \left(f^k \right) \right)$, so that $\dim \left(\ker \left(f^k \right) \right) = \dim \left(\ker \left(N^k \right) \right) = \left(\lambda^t \right)_1 + \left(\lambda^t \right)_2 + \ldots + \left(\lambda^t \right)_k$. This proves (13.124.7).

But since $\lambda = (\lambda_1, \lambda_2, ..., \lambda_\ell)$, we have $e_\lambda = e_{\lambda_1} e_{\lambda_2} \cdots e_{\lambda_\ell}$ and thus

$$\begin{aligned}
\varphi(e_\lambda) &= \varphi(e_{\lambda_1} e_{\lambda_2} \cdots e_{\lambda_\ell}) \\
&= \varphi(e_{\lambda_1}) \varphi(e_{\lambda_2}) \cdots \varphi(e_{\lambda_\ell}) \qquad (\text{since } \varphi \text{ is a } \mathbb{C}\text{-algebra homomorphism}) \\
&= \prod_{i=1}^{\ell} \underbrace{\varphi(e_{\lambda_i})}_{\substack{= q^{\binom{\lambda_i}{2}} \underline{1}_{J_{(1^{\lambda_i})}} \\ (\text{since Theorem 4.9.5 yields} \\ \varphi(e_p) = q^{\binom{p}{2}} \underline{1}_{J_{(1^p)}} \text{ for every } p \in \mathbb{N})}} = \prod_{i=1}^{\ell} \left(q^{\binom{\lambda_i}{2}} \underline{1}_{J_{(1^{\lambda_i})}} \right) \\
&= \left(\prod_{i=1}^{\ell} q^{\binom{\lambda_i}{2}} \right) \underbrace{\prod_{i=1}^{\ell} \underline{1}_{J_{(1^{\lambda_i})}}}_{\substack{= \underline{1}_{J_{(1^{\lambda_1})}} \underline{1}_{J_{(1^{\lambda_2})}} \cdots \underline{1}_{J_{(1^{\lambda_\ell})}} \\ = \sum_{\mu \in \operatorname{Par}} g^{\mu}_{(1^{\lambda_1}), (1^{\lambda_2}), ..., (1^{\lambda_\ell})}(q) \underline{1}_{J_\mu} \\ (\text{by (13.124.1), applied to } k = \ell \text{ and } \lambda^{(i)} = (1^{\lambda_i}))}} \\
&= \left(\prod_{i=1}^{\ell} q^{\binom{\lambda_i}{2}} \right) \sum_{\mu \in \operatorname{Par}} g^{\mu}_{(1^{\lambda_1}), (1^{\lambda_2}), ..., (1^{\lambda_\ell})}(q) \underline{1}_{J_\mu}. \qquad (13.124.9)
\end{aligned}$$

Now, we notice that

$$\lambda^t = (1^{\lambda_1}) + (1^{\lambda_2}) + \cdots + (1^{\lambda_\ell}) \qquad (13.124.10)$$

[901]. Now, for every partition μ, we have $g^{\mu}_{(1^{\lambda_1}), (1^{\lambda_2}), ..., (1^{\lambda_\ell})}(q) = 0$ unless $\mu \in \operatorname{Par}_n$ and $\lambda^t \triangleright \mu$ [902]. Hence, we can replace the summation sign "$\sum_{\mu \in \operatorname{Par}}$" on the right hand side of (13.124.9) by a more restricted

[901] *Proof of (13.124.10):* In the following, we use the so-called *Iverson bracket notation*: For every assertion $\mathcal{A}$, we let $[\mathcal{A}]$ denote the integer $\begin{cases} 1, & \text{if } \mathcal{A} \text{ is true;} \\ 0, & \text{if } \mathcal{A} \text{ is false} \end{cases}$. (This integer is called the *truth value* of $\mathcal{A}$.)

For every $p \in \mathbb{N}$ and $i \in \{1, 2, 3, ...\}$, we have

$$(1^p)_i = [p \geq i]. \qquad (13.124.11)$$

Now, every $i \in \{1, 2, 3, ...\}$ satisfies

$$\begin{aligned}
\left((1^{\lambda_1}) + (1^{\lambda_2}) + \cdots + (1^{\lambda_\ell}) \right)_i &= (1^{\lambda_1})_i + (1^{\lambda_2})_i + \cdots + (1^{\lambda_\ell})_i \qquad (\text{by the definition of } \mu + \nu \text{ for two partitions } \mu \text{ and } \nu) \\
&= \sum_{k=1}^{\ell} \underbrace{(1^{\lambda_k})_i}_{\substack{= [\lambda_k \geq i] \\ (\text{by (13.124.11)})}} = \sum_{k=1}^{\ell} [\lambda_k \geq i] \\
&= |\{j \in \{1, 2, ..., \ell\} \mid \lambda_j \geq i\}| = (\lambda^t)_i \qquad (\text{by (2.2.7)}).
\end{aligned}$$

Hence, $(1^{\lambda_1}) + (1^{\lambda_2}) + \cdots + (1^{\lambda_\ell}) = \lambda^t$, qed.

[902] *Proof.* Let μ be a partition. Exercise 4.9.6(b) (applied to ℓ, (1^{λ_i}) and μ instead of k, $\lambda^{(i)}$ and λ) shows that we have $g^{\mu}_{(1^{\lambda_1}), (1^{\lambda_2}), ..., (1^{\lambda_\ell})}(q) = 0$ unless $|(1^{\lambda_1})| + |(1^{\lambda_2})| + \cdots + |(1^{\lambda_\ell})| = |\mu|$ and $(1^{\lambda_1}) + (1^{\lambda_2}) + \cdots + (1^{\lambda_\ell}) \triangleright \mu$. Since $|(1^{\lambda_1})| + |(1^{\lambda_2})| + \cdots + |(1^{\lambda_\ell})| = \left| \underbrace{(1^{\lambda_1}) + (1^{\lambda_2}) + \cdots + (1^{\lambda_\ell})}_{\substack{= \lambda^t \\ (\text{by (13.124.10)})}} \right| = |\lambda^t| = |\lambda| = n$ and $(1^{\lambda_1}) + (1^{\lambda_2}) + \cdots + (1^{\lambda_\ell}) = \lambda^t$ (by (13.124.10)), this rewrites as follows: We have $g^{\mu}_{(1^{\lambda_1}), (1^{\lambda_2}), ..., (1^{\lambda_\ell})}(q) = 0$ unless $n = |\mu|$ and $\lambda^t \triangleright \mu$. In other words, we have $g^{\mu}_{(1^{\lambda_1}), (1^{\lambda_2}), ..., (1^{\lambda_\ell})}(q) = 0$ unless $\mu \in \operatorname{Par}_n$ and $\lambda^t \triangleright \mu$, qed.

summation “$\sum_{\mu\in\operatorname{Par}_n;\ \lambda^t\triangleright\mu}$” without changing the value of the sum (since all addends that we lose are 0). Thus, (13.124.9) rewrites as

$$\varphi\left(e_\lambda\right)=\left(\prod_{i=1}^{\ell} q^{\binom{\lambda_i}{2}}\right)\sum_{\mu\in\operatorname{Par}_n;\ \lambda^t\triangleright\mu} g^{\mu}_{\left(1^{\lambda_1}\right),\left(1^{\lambda_2}\right),\ldots,\left(1^{\lambda_\ell}\right)}\left(q\right)\underline{1}_{J_\mu}$$
$$=\sum_{\mu\in\operatorname{Par}_n;\ \lambda^t\triangleright\mu}\left(\prod_{i=1}^{\ell} q^{\binom{\lambda_i}{2}}\right) g^{\mu}_{\left(1^{\lambda_1}\right),\left(1^{\lambda_2}\right),\ldots,\left(1^{\lambda_\ell}\right)}\left(q\right)\underline{1}_{J_\mu}.$$

Setting $\alpha_{\lambda,\mu}=\left(\prod_{i=1}^{\ell} q^{\binom{\lambda_i}{2}}\right) g^{\mu}_{\left(1^{\lambda_1}\right),\left(1^{\lambda_2}\right),\ldots,\left(1^{\lambda_\ell}\right)}\left(q\right)$, we can rewrite this as $\varphi\left(e_\lambda\right)=\sum_{\mu\in\operatorname{Par}_n;\ \lambda^t\triangleright\mu}\alpha_{\lambda,\mu}\underline{1}_{J_\mu}$. Thus, we will be done solving Exercise 4.9.6(d) as soon as we can prove the inequality

$$\left(\prod_{i=1}^{\ell} q^{\binom{\lambda_i}{2}}\right) g^{\lambda^t}_{\left(1^{\lambda_1}\right),\left(1^{\lambda_2}\right),\ldots,\left(1^{\lambda_\ell}\right)}\left(q\right)\neq 0.$$

But the latter inequality follows from $q\neq 0$ and (13.124.8). Thus, Exercise 4.9.6(d) is solved.

(e) The map $\varphi:\Lambda_{\mathbb{C}}\to\mathcal{H}$ is graded, and thus, in order to prove that φ is injective, it is enough to show that the restriction $\varphi\mid_{\left(\Lambda_{\mathbb{C}}\right)_n}$ of φ to $\left(\Lambda_{\mathbb{C}}\right)_n$ is injective for every $n\in\mathbb{N}$. So let us fix $n\in\mathbb{N}$.

We know that $\left(e_\lambda\right)_{\lambda\in\operatorname{Par}_n}$ is a basis of the $\mathbb{C}$-vector space $\left(\Lambda_{\mathbb{C}}\right)_n$. [903] Hence, $\left(e_{\lambda^t}\right)_{\lambda\in\operatorname{Par}_n}$ also is a basis of the $\mathbb{C}$-vector space $\left(\Lambda_{\mathbb{C}}\right)_n$. Every $\lambda\in\operatorname{Par}_n$ satisfies

$$\varphi\left(e_{\lambda^t}\right)=\sum_{\mu\in\operatorname{Par}_n;\ \left(\lambda^t\right)^t\triangleright\mu}\alpha_{\lambda^t,\mu}\underline{1}_{J_\mu}$$

for some coefficients $\alpha_{\lambda^t,\mu}\in\mathbb{C}$ satisfying $\alpha_{\lambda^t,\left(\lambda^t\right)^t}\neq 0$ (according to Exercise 4.9.6(d)). In other words, every $\lambda\in\operatorname{Par}_n$ satisfies

$$\varphi\left(e_{\lambda^t}\right)=\sum_{\mu\in\operatorname{Par}_n;\ \lambda\triangleright\mu}\alpha_{\lambda^t,\mu}\underline{1}_{J_\mu}\qquad(13.124.12)$$

for some coefficients $\alpha_{\lambda^t,\mu}\in\mathbb{C}$ satisfying $\alpha_{\lambda^t,\lambda}\neq 0$ (since $\left(\lambda^t\right)^t=\lambda$).

Now, regard the set Par_n as a poset with the smaller-or-equal relation $\triangleright$.

The $\mathbb{C}$-vector space basis $\left(e_{\lambda^t}\right)_{\lambda\in\operatorname{Par}_n}$ of $\left(\Lambda_{\mathbb{C}}\right)_n$ and the $\mathbb{C}$-vector space basis $\left(\underline{1}_{J_\lambda}\right)_{\lambda\in\operatorname{Par}_n}$ of $\mathcal{H}_n$ are both indexed by the poset Par_n, and the $\operatorname{Par}_n\times\operatorname{Par}_n$-matrix that represents the map $\varphi\mid_{\left(\Lambda_{\mathbb{C}}\right)_n}$ with respect to these bases[904] is triangular[905] (by (13.124.12)). Furthermore, the diagonal entries of this triangular matrix are nonzero (due to $\alpha_{\lambda^t,\lambda}\neq 0$), and therefore invertible (in $\mathbb{C}$). Hence, this matrix is invertibly triangular, and thus invertible[906]. Therefore, the map $\varphi\mid_{\left(\Lambda_{\mathbb{C}}\right)_n}$ (which is represented by this matrix) is invertible (as a linear map $\left(\Lambda_{\mathbb{C}}\right)_n\to\mathcal{H}_n$) and thus injective. This completes the solution to Exercise 4.9.6(e).

13.125. **Solution to Exercise 5.2.13.** *Solution to Exercise 5.2.13.*

Alternative proof of Theorem 5.2.11. Let P be a labelled poset.

First of all, let f be any map $P\to\{1,2,3,\ldots\}$. We define a binary relation $\prec_f$ on the set P by letting $i\prec_f j$ hold if and only if

$$\left(f\left(i\right)<f\left(j\right)\ \text{ or }\ \left(f\left(i\right)=f\left(j\right)\ \text{ and }\ i<_{\mathbb{Z}} j\right)\right).$$

[903] This has been proven in the proof of Proposition 2.2.10. (Alternatively, this can be easily concluded from Proposition 2.2.10.)

[904] i.e., the $\operatorname{Par}_n\times\operatorname{Par}_n$-matrix whose $\left(\mu,\lambda\right)$-th entry is the $\underline{1}_{J_\mu}$-coordinate of $\varphi\left(e_{\lambda^t}\right)$

[905] See Definition 11.1.7 for the notation we are using here.

[906] by Proposition 11.1.10(d)

It is straightforward to see that this binary relation $\prec_f$ is the smaller relation of a total order on P. Let us define $\mathbf{w}(f)$ to be the set P endowed with this total order. Thus, $\mathbf{w}(f) = P$ as sets, but the smaller relation $<_{\mathbf{w}(f)}$ of the totally ordered set $\mathbf{w}(f)$ is the relation $\prec_f$.

Let us now forget that we fixed f. Thus, for every map $f : P \to \{1, 2, 3, \ldots\}$, we have constructed a binary relation $\prec_f$ on the set P and a totally ordered poset $\mathbf{w}(f)$. It is easy to see that, for every $f \in \mathcal{A}(P)$, we have

$$\mathbf{w}(f) \in \mathcal{L}(P) \tag{13.125.1}$$

[907].

Now, fix $w \in \mathcal{L}(P)$. Thus, w is a linear extension of P. That is, w is a totally ordered set with ground set P, and extends the poset P.

Let us first show that

$$\{g \in \mathcal{A}(P) \mid \mathbf{w}(g) = w\} \subset \mathcal{A}(w). \tag{13.125.2}$$

[*Proof of (13.125.2):* Let $f \in \{g \in \mathcal{A}(P) \mid \mathbf{w}(g) = w\}$. Hence, f is an element of $\mathcal{A}(P)$ and satisfies $\mathbf{w}(f) = w$. Therefore, $w = \mathbf{w}(f) = P$ as sets.

Our next goal is to prove $f \in \mathcal{A}(w)$. In other words, we want to prove that f is a w-partition.

Indeed, we claim that

$$(\text{if } i \in w \text{ and } j \in w \text{ satisfy } i <_w j \text{ and } i <_{\mathbb{Z}} j, \text{ then } f(i) \leq f(j)) \tag{13.125.3}$$

and

$$(\text{if } i \in w \text{ and } j \in w \text{ satisfy } i <_w j \text{ and } i >_{\mathbb{Z}} j, \text{ then } f(i) < f(j)). \tag{13.125.4}$$

Let us prove (13.125.3) first. So let i and j be two elements of w satisfying $i <_w j$ and $i <_{\mathbb{Z}} j$. We have $i <_w j$, thus $i <_{\mathbf{w}(f)} j$ (since $w = \mathbf{w}(f)$), and thus $i \prec_f j$ (since the relation $<_{\mathbf{w}(f)}$ is the relation $\prec_f$). By the definition of $\prec_f$, this means that we have $(f(i) < f(j)$ or $(f(i) = f(j)$ and $i <_{\mathbb{Z}} j))$. From this, we immediately obtain $f(i) \leq f(j)$. Thus, (13.125.3) is proven.

The proof of (13.125.4) is similar to the proof that we just gave for (13.125.3), but with a minor twist: In order to derive $f(i) < f(j)$ from $(f(i) < f(j)$ or $(f(i) = f(j)$ and $i <_{\mathbb{Z}} j))$, we need to recall the assumption $i >_{\mathbb{Z}} j$ (which rules out the possibility $(f(i) = f(j)$ and $i <_{\mathbb{Z}} j)$).

Thus, both (13.125.3) and (13.125.4) are proven. In other words, f is a w-partition. In yet other words, $f \in \mathcal{A}(w)$.

Let us now forget that we fixed f. We thus have proven that every $f \in \{g \in \mathcal{A}(P) \mid \mathbf{w}(g) = w\}$ satisfies $f \in \mathcal{A}(w)$. In other words, (13.125.2) is proven.]

Let us next show that

$$\mathcal{A}(w) \subset \{g \in \mathcal{A}(P) \mid \mathbf{w}(g) = w\}. \tag{13.125.5}$$

[*Proof of (13.125.5):* Let $f \in \mathcal{A}(w)$. Thus, f is a w-partition. We shall next prove that $f \in \{g \in \mathcal{A}(P) \mid \mathbf{w}(g) = w\}$.

Indeed, let us first make a general and trivial observation: If Q and R are two labelled posets such that $Q = R$ as sets, and if every two elements i and j of Q satisfying $i <_Q j$ satisfy $i <_R j$ (that is, the poset R is

[907] *Proof of (13.125.1):* Let $f \in \mathcal{A}(P)$. We need to show that $\mathbf{w}(f) \in \mathcal{L}(P)$. In other words, we need to show that $\mathbf{w}(f)$ is a linear extension of P. In order to show this, it clearly suffices to prove that every two elements i and j of P satisfying $i <_P j$ satisfy $i <_{\mathbf{w}(f)} j$ (because we already know that $\mathbf{w}(f)$ is a totally ordered set). So, let us fix two elements i and j of P satisfying $i <_P j$. We need to prove that $i <_{\mathbf{w}(f)} j$.

We have $i <_P j$ and thus $i \neq j$. Hence, either $i <_{\mathbb{Z}} j$ or $i >_{\mathbb{Z}} j$. In other words, we are in one of the following two Cases:

Case 1: We have $i <_{\mathbb{Z}} j$.

Case 2: We have $i >_{\mathbb{Z}} j$.

Let us first consider Case 1. In this case, $i <_{\mathbb{Z}} j$. But f is a P-partition (since $f \in \mathcal{A}(P)$), and thus, by the definition of a P-partition, we conclude that $f(i) \leq f(j)$ (since $i <_P j$ and $i <_{\mathbb{Z}} j$). In other words, either $f(i) < f(j)$ or $f(i) = f(j)$. Therefore, either $f(i) < f(j)$ or $(f(i) = f(j)$ and $i <_{\mathbb{Z}} j)$ (because we have $i <_{\mathbb{Z}} j$ by assumption). In other words, $i \prec_f j$. This rewrites as $i <_{\mathbf{w}(f)} j$ (since the relation $<_{\mathbf{w}(f)}$ is the relation $\prec_f$). Thus, $i <_{\mathbf{w}(f)} j$ is proven in Case 1.

Let us now consider Case 2. In this case, $i >_{\mathbb{Z}} j$. But f is a P-partition (since $f \in \mathcal{A}(P)$), and thus, by the definition of a P-partition, we conclude that $f(i) < f(j)$ (since $i <_P j$ and $i >_{\mathbb{Z}} j$). Therefore, either $f(i) < f(j)$ or $(f(i) = f(j)$ and $i <_{\mathbb{Z}} j)$. In other words, $i \prec_f j$. This rewrites as $i <_{\mathbf{w}(f)} j$ (since the relation $<_{\mathbf{w}(f)}$ is the relation $\prec_f$). Thus, $i <_{\mathbf{w}(f)} j$ is proven in Case 2.

We have now proven $i <_{\mathbf{w}(f)} j$ in both Cases 1 and 2; thus, $i <_{\mathbf{w}(f)} j$ always holds. This completes the proof of (13.125.1).

an extension of the poset Q), then every R-partition is a Q-partition.[908] Applying this to $Q = P$ and $R = w$, we conclude that every w-partition is a P-partition. Thus, f is a P-partition (since f is a w-partition). In other words, $f \in \mathcal{A}(P)$.

Next, we want to check that $\mathbf{w}(f) = w$.

It is easy to check that every two elements i and j of w satisfying $i <_w j$ satisfy

$$i <_{\mathbf{w}(f)} j \tag{13.125.6}$$

[909]. Now, let us again state a triviality: If Q and R are two totally ordered sets such that $Q = R$ as sets, and if every two elements i and j of Q satisfying $i <_Q j$ satisfy $i <_R j$, then $Q = R$ as totally ordered sets.[910] Applying this to $Q = w$ and $R = \mathbf{w}(f)$, we conclude that $w = \mathbf{w}(f)$ as totally ordered sets (since every two elements i and j of w satisfying $i <_w j$ satisfy $i <_{\mathbf{w}(f)} j$). In other words, $\mathbf{w}(f) = w$.

Now, we know that $f \in \mathcal{A}(P)$ and $\mathbf{w}(f) = w$. In other words, $f \in \{g \in \mathcal{A}(P) \mid \mathbf{w}(g) = w\}$.

Let us now forget that we fixed f. We thus have shown that $f \in \{g \in \mathcal{A}(P) \mid \mathbf{w}(g) = w\}$ for every $f \in \mathcal{A}(w)$. This proves (13.125.5).]

Combining (13.125.2) with (13.125.5), we obtain

$$\{g \in \mathcal{A}(P) \mid \mathbf{w}(g) = w\} = \mathcal{A}(w). \tag{13.125.7}$$

Let us now forget that we fixed w. We thus have proven (13.125.7) for every $w \in \mathcal{L}(P)$.

Now, the definition of $F_P(\mathbf{x})$ yields

$$\begin{aligned} F_P(\mathbf{x}) &= \sum_{f \in \mathcal{A}(P)} \mathbf{x}_f = \underbrace{\sum_{w \in \mathcal{L}(P)} \underbrace{\sum_{\substack{f \in \mathcal{A}(P); \\ \mathbf{w}(f) = w}}}_{\substack{= \sum_{f \in \{g \in \mathcal{A}(P) \mid \mathbf{w}(g) = w\}} \\ = \sum_{f \in \mathcal{A}(w)} \\ \text{(by (13.125.7))}}} \mathbf{x}_f}_{} \\ &\qquad \left(\text{since } \mathbf{w}(f) \in \mathcal{L}(P) \text{ for every } f \in \mathcal{A}(P) \text{ (by (13.125.1))}\right) \\ &= \sum_{w \in \mathcal{L}(P)} \underbrace{\sum_{f \in \mathcal{A}(w)} \mathbf{x}_f}_{\substack{= F_w(\mathbf{x}) \\ \text{(since } F_w(\mathbf{x}) = \sum_{f \in \mathcal{A}(w)} \mathbf{x}_f \\ \text{(by the definition of } F_w(\mathbf{x})\text{))}}} = \sum_{w \in \mathcal{L}(P)} F_w(\mathbf{x}). \end{aligned}$$

This completes our proof of Theorem 5.2.11. □

13.126. **Solution to Exercise 5.3.7.** *Solution to Exercise 5.3.7.* Let us first state a basic fact about totally ordered sets:

Lemma 13.126.1. *Let T be a totally ordered set, and let $<_T$ be the smaller relation of T. Let $(a_1, a_2, \ldots, a_n)$ be a finite list of* ***distinct*** *elements of T. Then, there is a* ***unique*** *permutation $\sigma \in \mathfrak{S}_n$ such that $a_{\sigma(1)} <_T a_{\sigma(2)} <_T \cdots <_T a_{\sigma(n)}$.*

[908] This is clear, because the requirements for an R-partition are at least as strong as the requirements for a Q-partition.

[909] *Proof of (13.125.6):* Let i and j be two elements of w satisfying $i <_w j$. We must prove that $i <_{\mathbf{w}(f)} j$.

We have $i <_w j$ and thus $i \neq j$. Hence, either $i <_{\mathbb{Z}} j$ or $i >_{\mathbb{Z}} j$. In other words, we are in one of the following two Cases:

Case 1: We have $i <_{\mathbb{Z}} j$.

Case 2: We have $i >_{\mathbb{Z}} j$.

Let us consider Case 1 first. In this case, we have $i <_{\mathbb{Z}} j$. Since f is a w-partition, we have $f(i) \leq f(j)$ (because $i <_w j$ and $i <_{\mathbb{Z}} j$). In other words, $(f(i) < f(j)$ or $f(i) = f(j))$. Hence, $(f(i) < f(j)$ or $(f(i) = f(j)$ and $i <_{\mathbb{Z}} j))$ (because we have assumed that $i <_{\mathbb{Z}} j$). Therefore, $i \prec_f j$ (because of the definition of "$i \prec_f j$"). In other words, $i <_{\mathbf{w}(f)} j$ (since the relation $<_{\mathbf{w}(f)}$ is the relation $\prec_f$). Thus, $i <_{\mathbf{w}(f)} j$ is proven in Case 1.

We can similarly prove $i <_{\mathbf{w}(f)} j$ in Case 2 (but now we obtain $f(i) < f(j)$ instead of $f(i) \leq f(j)$).

We have now shown that $i <_{\mathbf{w}(f)} j$ in both Cases 1 and 2. Thus, (13.125.6) is proven.

[910] In other words: If a totally ordered set R is an extension of a totally ordered set Q, then $Q = R$.

Lemma 13.126.1 is well-known (it essentially says that a finite list of distinct elements of a totally ordered set can be sorted into increasing order by a unique permutation). We shall use it to prove Proposition 5.3.2 later.

Next, we state an elementary property of permutations:

Proposition 13.126.2. *Let $n \in \mathbb{N}$. Let φ and ψ be two elements of $\mathfrak{S}_n$. Assume that for every two elements $a \in \{1, 2, \ldots, n\}$ and $b \in \{1, 2, \ldots, n\}$ satisfying $a < b$, we have*

$$(\varphi(a) < \varphi(b) \text{ if and only if } \psi(a) < \psi(b)). \tag{13.126.1}$$

Then, $\varphi = \psi$.

Proof of Proposition 13.126.2. If a and b are two elements of $\{1, 2, \ldots, n\}$ satisfying $a < b$, then we have the logical equivalence

$$(\varphi(a) < \varphi(b)) \iff (\psi(a) < \psi(b)) \tag{13.126.2}$$

(by (13.126.1)). We next will show that this equivalence holds even if we don't require $a < b$:

> *Observation 1:* Let $p \in \{1, 2, \ldots, n\}$ and $q \in \{1, 2, \ldots, n\}$. Then, we have the logical equivalence
> $$(\varphi(p) < \varphi(q)) \iff (\psi(p) < \psi(q)).$$

[*Proof of Observation 1:* Let us first prove the logical implication

$$(\varphi(p) < \varphi(q)) \implies (\psi(p) < \psi(q)). \tag{13.126.3}$$

[*Proof of (13.126.3):* Assume that $\varphi(p) < \varphi(q)$. We want to show that $\psi(p) < \psi(q)$.

If $p < q$, then (13.126.2) (applied to $a = p$ and $b = q$) yields the equivalence $(\varphi(p) < \varphi(q)) \iff (\psi(p) < \psi(q))$, and thus $\psi(p) < \psi(q)$ follows (since $\varphi(p) < \varphi(q)$). Hence, for the rest of the proof of $\psi(p) < \psi(q)$, we WLOG assume that we don't have $p < q$. Hence, we have $p \geq q$.

From $\varphi(p) < \varphi(q)$, we also obtain $\varphi(p) \neq \varphi(q)$, so that $p \neq q$. Combining this with $p \geq q$, we obtain $p > q$. Hence, $q < p$. Thus, (13.126.2) (applied to $a = q$ and $b = p$) yields the logical equivalence $(\varphi(q) < \varphi(p)) \iff (\psi(q) < \psi(p))$. Since we don't have $\varphi(q) < \varphi(p)$ (because $\varphi(p) < \varphi(q)$), we thus conclude that we don't have $\psi(q) < \psi(p)$. Therefore, we have $\psi(p) \leq \psi(q)$. But the map ψ is injective (since $\psi \in \mathfrak{S}_n$); therefore, from $p \neq q$, we obtain $\psi(p) \neq \psi(q)$. Combining this with $\psi(p) \leq \psi(q)$, we obtain $\psi(p) < \psi(q)$. This completes the proof of (13.126.3).]

So we have proven (13.126.3). The same argument (but with the roles of φ and ψ interchanged) yields the logical implication

$$(\psi(p) < \psi(q)) \implies (\varphi(p) < \varphi(q)).$$

Combining this with (13.126.3), we obtain the equivalence $(\varphi(p) < \varphi(q)) \iff (\psi(p) < \psi(q))$. This proves Observation 1.]

Now, let $q \in \{1, 2, \ldots, n\}$. Then, the map φ is a bijection $\{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$ (since $\varphi \in \mathfrak{S}_n$). Thus, we can substitute i for $\varphi(p)$ in the sum $\sum_{\substack{p \in \{1,2,\ldots,n\}; \\ \varphi(p) < \varphi(q)}} 1$. We thus obtain

$$\sum_{\substack{p \in \{1,2,\ldots,n\}; \\ \varphi(p) < \varphi(q)}} 1 = \sum_{\substack{i \in \{1,2,\ldots,n\}; \\ i < \varphi(q)}} 1 = \sum_{i \in \{1,2,\ldots,\varphi(q)-1\}} 1 = (\varphi(q) - 1) \cdot 1 = \varphi(q) - 1. \tag{13.126.4}$$

The same argument (applied to ψ instead of φ) yields

$$\sum_{\substack{p \in \{1,2,\ldots,n\}; \\ \psi(p) < \psi(q)}} 1 = \psi(q) - 1.$$

However, Observation 1 shows that the sums $\sum_{\substack{p \in \{1,2,\ldots,n\}; \\ \varphi(p) < \varphi(q)}} 1$ and $\sum_{\substack{p \in \{1,2,\ldots,n\}; \\ \psi(p) < \psi(q)}} 1$ range over the same values of p. Thus,

$$\sum_{\substack{p \in \{1,2,\ldots,n\}; \\ \varphi(p) < \varphi(q)}} 1 = \sum_{\substack{p \in \{1,2,\ldots,n\}; \\ \psi(p) < \psi(q)}} 1 = \psi(q) - 1.$$

Comparing this with (13.126.4), we obtain $\varphi(q) - 1 = \psi(q) - 1$. Hence, $\varphi(q) = \psi(q)$.

Now, forget that we fixed q. We thus have shown that $\varphi(q) = \psi(q)$ for each $q \in \{1, 2, \ldots, n\}$. In other words, $\varphi = \psi$. This proves Proposition 13.126.2. □

For later use (in a different solution further below), let us derive another proposition from Proposition 13.126.2. It relies on the following notation:

Definition 13.126.3. Let $n \in \mathbb{N}$. Let $\varphi \in \mathfrak{S}_n$. Define the *inversion set* $\operatorname{Inv}\varphi$ of φ to be the set $\left\{(i,j) \in \{1,2,\ldots,n\}^2 \mid i < j;\ \varphi(i) > \varphi(j)\right\}$.

(Note that this notation is not completely standard. Some authors, instead, define $\operatorname{Inv}\varphi$ to be $\left\{(\varphi(i), \varphi(j)) \mid (i,j) \in \{1,2,\ldots,n\}^2;\ i < j;\ \varphi(i) > \varphi(j)\right\}$. This is a different set, although of the same size.)

Proposition 13.126.4. *Let $n \in \mathbb{N}$. Let φ and ψ be two elements of $\mathfrak{S}_n$ satisfying $\operatorname{Inv}\varphi = \operatorname{Inv}\psi$. Then, $\varphi = \psi$.*

(Proposition 13.126.4 can be restated as follows: If $n \in \mathbb{N}$, then a permutation in $\mathfrak{S}_n$ is uniquely determined by its inversion set.)

Proposition 13.126.4 is a known fact in elementary combinatorics; let us quickly derive it from Proposition 13.126.2:

Proof of Proposition 13.126.4. Let $a \in \{1,2,\ldots,n\}$ and $b \in \{1,2,\ldots,n\}$ be such that $a < b$. We shall prove that $\varphi(a) < \varphi(b)$ if and only if $\psi(a) < \psi(b)$.

The pair (a,b) is an element of $\{1,2,\ldots,n\}^2$ satisfying $a < b$. Thus, (a,b) belongs to $\operatorname{Inv}\varphi$ if and only if it satisfies $\varphi(a) > \varphi(b)$ (by the definition of $\operatorname{Inv}\varphi$). In other words, we have the following logical equivalence:

$$((a,b) \in \operatorname{Inv}\varphi) \iff (\varphi(a) > \varphi(b)). \tag{13.126.5}$$

But the map φ is injective (since $\varphi \in \mathfrak{S}_n$). Thus, from $a \neq b$ (which follows from $a < b$), we obtain $\varphi(a) \neq \varphi(b)$. Hence, $\varphi(a) < \varphi(b)$ holds if and only if $\varphi(a) \leq \varphi(b)$. Hence, we have the following chain of equivalences:

$$(\varphi(a) < \varphi(b)) \iff (\varphi(a) \leq \varphi(b)) \iff \left(\text{not } \underbrace{(\varphi(a) > \varphi(b))}_{\substack{\iff (a,b)\in\operatorname{Inv}\varphi \\ \text{(by (13.126.5))}}}\right)$$
$$\iff (\text{not } ((a,b) \in \operatorname{Inv}\varphi)). \tag{13.126.6}$$

The same argument (applied to ψ instead of φ) yields the equivalence

$$(\psi(a) < \psi(b)) \iff (\text{not } ((a,b) \in \operatorname{Inv}\psi)). \tag{13.126.7}$$

Now, due to (13.126.6), we have the following chain of equivalences:

$$(\varphi(a) < \varphi(b)) \iff (\text{not } ((a,b) \in \operatorname{Inv}\varphi)) \iff (\text{not } ((a,b) \in \operatorname{Inv}\psi)) \qquad (\text{since } \operatorname{Inv}\varphi = \operatorname{Inv}\psi)$$
$$\iff (\psi(a) < \psi(b))$$

(by (13.126.7)). In other words, we have $\varphi(a) < \varphi(b)$ if and only if $\psi(a) < \psi(b)$.

Now, forget that we fixed a and b. We thus have shown that for every two elements $a \in \{1,2,\ldots,n\}$ and $b \in \{1,2,\ldots,n\}$ satisfying $a < b$, we have

$$(\varphi(a) < \varphi(b) \text{ if and only if } \psi(a) < \psi(b)).$$

Thus, Proposition 13.126.2 yields $\varphi = \psi$. This proves Proposition 13.126.4. □

We are now ready to prove Proposition 5.3.2:

Proof of Proposition 5.3.2. Define a binary relation $\prec$ on the set $\{1, 2, \ldots, n\}$ by letting $i \prec j$ hold if and only if

$$\left(w_i < w_j \text{ or } \left(w_i = w_j \text{ and } i <_{\mathbb{Z}} j\right)\right).$$

It is easy to see that this relation $\prec$ is the smaller relation of a total order. Let T denote the set $\{1, 2, \ldots, n\}$ endowed with this total order. Thus, $T = \{1, 2, \ldots, n\}$ as sets, but the smaller relation $<_T$ of the poset T is the relation $\prec$.

Clearly, $(1, 2, \ldots, n)$ is a finite list of distinct elements of this totally ordered set T. Hence, Lemma 13.126.1 (applied to $a_i = i$) yields that there is a **unique** permutation $\sigma \in \mathfrak{S}_n$ such that $\sigma(1) <_T \sigma(2) <_T \cdots <_T \sigma(n)$. Consider this σ, and denote it by γ. Thus, γ is a permutation in $\mathfrak{S}_n$ and satisfies $\gamma(1) <_T \gamma(2) <_T \cdots <_T \gamma(n)$. Hence, $\gamma^{-1} \in \mathfrak{S}_n$ as well.

We have $\gamma(1) <_T \gamma(2) <_T \cdots <_T \gamma(n)$. In other words, if i and j are two elements of $\{1, 2, \ldots, n\}$ satisfying $i < j$, then

$$\gamma(i) <_T \gamma(j). \tag{13.126.8}$$

Thus, for every two elements a and b of $\{1, 2, \ldots, n\}$ satisfying $a < b$, we have

$$\left(\gamma^{-1}(a) < \gamma^{-1}(b) \text{ if and only if } w_a \leq w_b\right). \tag{13.126.9}$$

[*Proof of (13.126.9):* Let a and b be two elements of $\{1, 2, \ldots, n\}$ satisfying $a < b$. Set $i = \gamma^{-1}(a)$ and $j = \gamma^{-1}(b)$. Thus, i and j are two elements of $\{1, 2, \ldots, n\}$ (since $\gamma^{-1} \in \mathfrak{S}_n$). Also, from $i = \gamma^{-1}(a)$, we obtain $\gamma(i) = a$. From $j = \gamma^{-1}(b)$, we obtain $\gamma(j) = b$.

We are in one of the following two cases:

Case 1: We have $i < j$.

Case 2: We have $i \geq j$.

Let us first consider Case 1. In this case, we have $i < j$. This rewrites as $\gamma^{-1}(a) < \gamma^{-1}(b)$ (since $i = \gamma^{-1}(a)$ and $j = \gamma^{-1}(b)$). On the other hand, from $i < j$, we obtain $\gamma(i) <_T \gamma(j)$ (by (13.126.8)). This rewrites as $a <_T b$ (since $\gamma(i) = a$ and $\gamma(j) = b$). This rewrites as $a \prec b$ (since the relation $<_T$ is the relation $\prec$). In other words, $\left(f(a) < f(b) \text{ or } \left(f(a) = f(b) \text{ and } a <_{\mathbb{Z}} b\right)\right)$ (by the definition of the relation $\prec$). Thus, $f(a) \leq f(b)$. But the definition of f yields $f(a) = w_a$ and $f(b) = w_b$. Thus, $w_a = f(a) \leq f(b) = w_b$. Now, we have shown that both statements $\left(\gamma^{-1}(a) < \gamma^{-1}(b)\right)$ and $(w_a \leq w_b)$ are true. Hence, we have $\left(\gamma^{-1}(a) < \gamma^{-1}(b) \text{ if and only if } w_a \leq w_b\right)$. This proves (13.126.9) in Case 1.

Let us now consider Case 2. In this case, we have $i \geq j$. This rewrites as $\gamma^{-1}(a) \geq \gamma^{-1}(b)$ (since $i = \gamma^{-1}(a)$ and $j = \gamma^{-1}(b)$). Hence, the statement $\left(\gamma^{-1}(a) < \gamma^{-1}(b)\right)$ is false.

We have $a < b$. Thus, we cannot have $b < a$. In other words, we cannot have $b <_{\mathbb{Z}} a$. Thus, we cannot have $\left(f(b) = f(a) \text{ and } b <_{\mathbb{Z}} a\right)$.

If we had $i = j$, then we would have $a = \gamma\left(\underbrace{i}_{=j}\right) = \gamma(j) = b$, which would contradict $a < b$. Hence, we cannot have $i = j$. Thus, we have $i \neq j$. Combining this with $i \geq j$, we obtain $i > j$. Therefore, $j < i$. Therefore, (13.126.8) (applied to j and i instead of i and j) yields $\gamma(j) <_T \gamma(i)$. This rewrites as $b <_T a$ (since $\gamma(i) = a$ and $\gamma(j) = b$). This rewrites as $b \prec a$ (since the relation $<_T$ is the relation $\prec$). In other words, $\left(f(b) < f(a) \text{ or } \left(f(b) = f(a) \text{ and } b <_{\mathbb{Z}} a\right)\right)$ (by the definition of the relation $\prec$). Hence, we must have $f(b) < f(a)$ (since we cannot have $\left(f(b) = f(a) \text{ and } b <_{\mathbb{Z}} a\right)$). But the definition of f yields $f(a) = w_a$ and $f(b) = w_b$. Thus, $w_b = f(b) < f(a) = w_a$. Hence, the statement $(w_a \leq w_b)$ is false. Now, we have shown that both statements $\left(\gamma^{-1}(a) < \gamma^{-1}(b)\right)$ and $(w_a \leq w_b)$ are false. Hence, we have $\left(\gamma^{-1}(a) < \gamma^{-1}(b) \text{ if and only if } w_a \leq w_b\right)$. This proves (13.126.9) in Case 2.

We have now proven (13.126.9) in each of the two Cases 1 and 2. Hence, (13.126.9) is always proven.]

So we know that γ^{-1} is a permutation in $\mathfrak{S}_n$, and that for every two elements a and b of $\{1, 2, \ldots, n\}$ satisfying $a < b$, we have $\left(\gamma^{-1}(a) < \gamma^{-1}(b) \text{ if and only if } w_a \leq w_b\right)$ (by (13.126.9)). Thus, there exists **at least one** permutation $\sigma \in \mathfrak{S}_n$ such that for every two elements a and b of $\{1, 2, \ldots, n\}$ satisfying $a < b$, we have $\left(\sigma(a) < \sigma(b) \text{ if and only if } w_a \leq w_b\right)$ (namely, $\sigma = \gamma^{-1}$).

It remains to prove that there exists **at most one** such permutation σ. In other words, it remains to prove that any two such permutations σ are equal. In other words, it remains to prove the following claim:

> *Claim 1:* Let φ and ψ be two permutations $\sigma \in \mathfrak{S}_n$ such that for every two elements a and b of $\{1, 2, \ldots, n\}$ satisfying $a < b$, we have $(\sigma(a) < \sigma(b)$ if and only if $w_a \leq w_b)$. Then, $\varphi = \psi$.

[*Proof of Claim 1:* We know that φ is a permutation $\sigma \in \mathfrak{S}_n$ such that for every two elements a and b of $\{1, 2, \ldots, n\}$ satisfying $a < b$, we have $(\sigma(a) < \sigma(b)$ if and only if $w_a \leq w_b)$. In other words, φ is a permutation in $\mathfrak{S}_n$, and has the property that for every two elements a and b of $\{1, 2, \ldots, n\}$ satisfying $a < b$, we have

$$(\varphi(a) < \varphi(b) \text{ if and only if } w_a \leq w_b). \tag{13.126.10}$$

Now, let a and b be two elements of $\{1, 2, \ldots, n\}$ satisfying $a < b$. Then, from (13.126.10), we obtain the logical equivalence $(\varphi(a) < \varphi(b)) \iff (w_a \leq w_b)$. The same argument (applied to ψ instead of φ) yields the logical equivalence $(\psi(a) < \psi(b)) \iff (w_a \leq w_b)$. Hence, we have the following chain of logical equivalences:

$$(\varphi(a) < \varphi(b)) \iff (w_a \leq w_b) \iff (\psi(a) < \psi(b))$$

(because of the equivalence $(\psi(a) < \psi(b)) \iff (w_a \leq w_b)$). In other words, we have

$$(\varphi(a) < \varphi(b) \text{ if and only if } \psi(a) < \psi(b)).$$

Now, forget that we fixed a and b. We thus have proven that for every two elements $a \in \{1, 2, \ldots, n\}$ and $b \in \{1, 2, \ldots, n\}$ satisfying $a < b$, we have

$$(\varphi(a) < \varphi(b) \text{ if and only if } \psi(a) < \psi(b)).$$

Thus, Proposition 13.126.2 yields that $\varphi = \psi$. This proves Claim 1.]

Hence, Proposition 5.3.2 is proven. □

Next, we shall show a lemma that will be crucial in our proof of Lemma 5.3.6:

Lemma 13.126.5. *Let $n \in \mathbb{N}$. Let $\tau \in \mathfrak{S}_n$. Let P be the labelled poset whose underlying set is $\{1, 2, \ldots, n\}$ and which (as a poset) is the total order $(\tau(1) < \tau(2) < \cdots < \tau(n))$ (that is, the order $<_P$ is given by $\tau(1) <_P \tau(2) <_P \cdots <_P \tau(n)$).*

Let $\mathfrak{A}$ denote the totally ordered set $\{1 < 2 < 3 < \cdots\}$ of positive integers. Let $f : P \to \mathfrak{A}$ be any map. Then, we have the following logical equivalence:

$$(f \in \mathcal{A}(P)) \iff \left(\operatorname{std}(f(1), f(2), \ldots, f(n)) = \tau^{-1}\right)$$

(where we treat $(f(1), f(2), \ldots, f(n))$ as a word in $\mathfrak{A}^n$).

Proof of Lemma 13.126.5. We have $P = \{1, 2, \ldots, n\}$ as sets. Also, the definition of the order $<_P$ yields

$$\tau(1) <_P \tau(2) <_P \cdots <_P \tau(n).$$

Hence, for any two elements i and j of $\{1, 2, \ldots, n\}$, we have

$$(\tau(i) <_P \tau(j) \text{ if and only if } i < j) \tag{13.126.11}$$

(where the "$<$" sign in "$i < j$" refers to the usual smaller relation $<_{\mathbb{Z}}$ of the totally ordered set $\mathbb{Z}$).

The map f is a map from P to $\mathfrak{A}$. In other words, the map f is a map from P to $\{1, 2, 3, \ldots\}$ (since $\mathfrak{A} = \{1, 2, 3, \ldots\}$).

Also, $(f(1), f(2), \ldots, f(n))$ is a word in $\mathfrak{A}^n$. Denote this word by w. Thus,

$$w = (f(1), f(2), \ldots, f(n)) \in \mathfrak{A}^n.$$

For each $i \in \{1, 2, \ldots, n\}$, the i-th letter of the word w has been denoted by w_i. Thus, $w = (w_1, w_2, \ldots, w_n)$, so that $(w_1, w_2, \ldots, w_n) = w = (f(1), f(2), \ldots, f(n))$. In other words,

$$w_i = f(i) \qquad \text{for each } i \in \{1, 2, \ldots, n\}. \tag{13.126.12}$$

The definition of the standardization $\operatorname{std} w$ shows that $\operatorname{std} w$ is the unique permutation $\sigma \in \mathfrak{S}_n$ such that for every two elements a and b of $\{1, 2, \ldots, n\}$ satisfying $a < b$, we have $(\sigma(a) < \sigma(b)$ if and only if $w_a \leq w_b)$. In particular, $\operatorname{std} w$ is such a permutation. In other words, $\operatorname{std} w$ is a permutation in $\mathfrak{S}_n$ and has the property that for every two elements a and b of $\{1, 2, \ldots, n\}$ satisfying $a < b$, we have

$$((\operatorname{std} w)(a) < (\operatorname{std} w)(b) \text{ if and only if } w_a \leq w_b). \tag{13.126.13}$$

Hence, for every two elements a and b of $\{1, 2, \ldots, n\}$ satisfying $a < b$, we have

$$\left((\operatorname{std} w)(a) < (\operatorname{std} w)(b) \text{ if and only if } f(a) \leq f(b)\right). \tag{13.126.14}$$

(Indeed, this just a restatement of (13.126.13), since (13.126.12) yields $w_a = f(a)$ and $w_b = f(b)$.)

We shall next prove the following two claims:

Claim 1: If $f \in \mathcal{A}(P)$, then $\operatorname{std} w = \tau^{-1}$.

Claim 2: If $\operatorname{std} w = \tau^{-1}$, then $f \in \mathcal{A}(P)$.

[*Proof of Claim 1:* Assume that $f \in \mathcal{A}(P)$. Thus, f is a P-partition (since $\mathcal{A}(P)$ is the set of all P-partitions). In other words, f is a map $P \to \{1, 2, 3, \ldots\}$ with the properties

$$(\text{if } i \in P \text{ and } j \in P \text{ satisfy } i <_P j \text{ and } i <_{\mathbb{Z}} j, \text{ then } f(i) \leq f(j)) \tag{13.126.15}$$

and

$$(\text{if } i \in P \text{ and } j \in P \text{ satisfy } i <_P j \text{ and } i >_{\mathbb{Z}} j, \text{ then } f(i) < f(j)) \tag{13.126.16}$$

(because this is how a P-partition is defined).

For every two elements $a \in \{1, 2, \ldots, n\}$ and $b \in \{1, 2, \ldots, n\}$ satisfying $a < b$, we have

$$\left((\operatorname{std} w)(a) < (\operatorname{std} w)(b) \text{ if and only if } \tau^{-1}(a) < \tau^{-1}(b)\right) \tag{13.126.17}$$

[911]. Hence, Proposition 13.126.2 (applied to $\varphi = \operatorname{std} w$ and $\psi = \tau^{-1}$) yields $\operatorname{std} w = \tau^{-1}$. Thus, Claim 1 is proven.]

[*Proof of Claim 2:* Assume that $\operatorname{std} w = \tau^{-1}$. The map $f : P \to \{1, 2, 3, \ldots\}$ has the following properties:

$$(\text{if } i \in P \text{ and } j \in P \text{ satisfy } i <_P j \text{ and } i <_{\mathbb{Z}} j, \text{ then } f(i) \leq f(j)) \tag{13.126.18}$$

[912] and

$$(\text{if } i \in P \text{ and } j \in P \text{ satisfy } i <_P j \text{ and } i >_{\mathbb{Z}} j, \text{ then } f(i) < f(j)) \tag{13.126.19}$$

[911] *Proof of (13.126.17):* Let $a \in \{1, 2, \ldots, n\}$ and $b \in \{1, 2, \ldots, n\}$ be such that $a < b$. We must prove (13.126.17).

We have $a \in \{1, 2, \ldots, n\} = P$ and $b \in \{1, 2, \ldots, n\} = P$.

Note that $a < b$. In other words, $a <_{\mathbb{Z}} b$. In other words, $b >_{\mathbb{Z}} a$.

Let $i = \tau^{-1}(a)$ and $j = \tau^{-1}(b)$. Thus, i and j belong to $\{1, 2, \ldots, n\}$. From $i = \tau^{-1}(a)$, we obtain $a = \tau(i)$. From $j = \tau^{-1}(b)$, we obtain $b = \tau(j)$. If we had $i = j$, then we would have $a = \tau\left(\underbrace{i}_{=j}\right) = \tau(j) = b$, which would contradict $a < b$. Thus, $i \neq j$. Hence, we are in one of the following two cases:

Case 1: We have $i < j$.

Case 2: We have $i > j$.

Let us first consider Case 1. In this case, we have $i < j$. Hence, (13.126.11) shows that $\tau(i) <_P \tau(j)$. In view of $\tau(i) = a$ and $\tau(j) = b$, this rewrites as $a <_P b$. Also, $a <_{\mathbb{Z}} b$. Hence, (13.126.15) (applied to a and b instead of i and j) yields $f(a) \leq f(b)$. Because of (13.126.14), this shows that $(\operatorname{std} w)(a) < (\operatorname{std} w)(b)$. Also, $\tau^{-1}(a) = i < j = \tau^{-1}(b)$. Thus, both statements $((\operatorname{std} w)(a) < (\operatorname{std} w)(b))$ and $\left(\tau^{-1}(a) < \tau^{-1}(b)\right)$ are true. Hence, we have $\left((\operatorname{std} w)(a) < (\operatorname{std} w)(b) \text{ if and only if } \tau^{-1}(a) < \tau^{-1}(b)\right)$. Thus, (13.126.17) is proven in Case 1.

Let us now consider Case 2. In this case, we have $i > j$. Hence, $j < i$. But (13.126.11) (applied to j and i instead of i and j) shows that $\tau(j) <_P \tau(i)$ if and only if $j < i$. Hence, we have $\tau(j) <_P \tau(i)$ (since we have $j < i$). In view of $\tau(i) = a$ and $\tau(j) = b$, this rewrites as $b <_P a$. Also, $b >_{\mathbb{Z}} a$. Hence, (13.126.16) (applied to b and a instead of i and j) yields $f(b) < f(a)$. In other words, we don't have $f(a) \leq f(b)$. Because of (13.126.14), this shows that we don't have $(\operatorname{std} w)(a) < (\operatorname{std} w)(b)$. Also, $\tau^{-1}(a) = i > j = \tau^{-1}(b)$. Hence, we don't have $\tau^{-1}(a) < \tau^{-1}(b)$. Thus, both statements $((\operatorname{std} w)(a) < (\operatorname{std} w)(b))$ and $\left(\tau^{-1}(a) < \tau^{-1}(b)\right)$ are false. Hence, we have $\left((\operatorname{std} w)(a) < (\operatorname{std} w)(b) \text{ if and only if } \tau^{-1}(a) < \tau^{-1}(b)\right)$. Thus, (13.126.17) is proven in Case 2.

We have now proven (13.126.17) in both Cases 1 and 2. Thus, (13.126.17) always holds.

[912] *Proof of (13.126.18):* Let $i \in P$ and $j \in P$ be such that $i <_P j$ and $i <_{\mathbb{Z}} j$. We must prove that $f(i) \leq f(j)$.

Let $a = \tau^{-1}(i)$ and $b = \tau^{-1}(j)$. Thus, $a = \tau^{-1}(i) \in \{1, 2, \ldots, n\} = P$ and $b = \tau^{-1}(j) \in \{1, 2, \ldots, n\} = P$. From $a = \tau^{-1}(i)$, we obtain $\tau(a) = i$. From $b = \tau^{-1}(j)$, we obtain $\tau(b) = j$.

We have $i <_P j$. In view of $i = \tau(a)$ and $j = \tau(b)$, this rewrites as $\tau(a) <_P \tau(b)$.

But (13.126.11) (applied to a and b instead of i and j) shows that $(\tau(a) <_P \tau(b) \text{ if and only if } a < b)$. Hence, we have $a < b$ (since we have $\tau(a) <_P \tau(b)$). Because of $\operatorname{std} w = \tau^{-1}$, we now have $\underbrace{(\operatorname{std} w)}_{=\tau^{-1}}(i) = \tau^{-1}(i) = a < b = \underbrace{\tau^{-1}}_{=\operatorname{std} w}(j) = (\operatorname{std} w)(j)$.

But $i <_{\mathbb{Z}} j$. In other words, $i < j$. Thus, (13.126.14) (applied to i and j instead of a and b) shows that $((\operatorname{std} w)(i) < (\operatorname{std} w)(j) \text{ if and only if } f(i) \leq f(j))$. Thus, $f(i) \leq f(j)$ (since $(\operatorname{std} w)(i) < (\operatorname{std} w)(j)$). This proves (13.126.18).

[913]. Thus, f is a P-partition (because this is how a P-partition is defined). In other words, $f \in \mathcal{A}(P)$ (since $\mathcal{A}(P)$ is the set of all P-partitions). This proves Claim 2.]

Combining Claim 1 with Claim 2, we obtain the logical equivalence $(f \in \mathcal{A}(P)) \iff \left(\operatorname{std} w = \tau^{-1}\right)$. In view of $w = (f(1), f(2), \ldots, f(n))$, this rewrites as

$$(f \in \mathcal{A}(P)) \iff \left(\operatorname{std}(f(1), f(2), \ldots, f(n)) = \tau^{-1}\right).$$

This proves Lemma 13.126.5. □

Next, we will use a trivial consequence of Proposition 5.2.10:

Lemma 13.126.6. *Let $n \in \mathbb{N}$. Let $\sigma \in \mathfrak{S}_n$. Let P be the labelled poset whose underlying set is $\{1, 2, \ldots, n\}$ and which (as a poset) is the total order $(\sigma(1) < \sigma(2) < \cdots < \sigma(n))$ (that is, the order $<_P$ is given by $\sigma(1) <_P \sigma(2) <_P \cdots <_P \sigma(n)$). Then, $F_P(\mathbf{x}) = L_{\gamma(\sigma)}$.*

Proof of Lemma 13.126.6. In Proposition 5.2.10, we have defined $\operatorname{Des} w$ for any labelled poset w that is a total order. Applying this definition to $w = P$, we obtain

$$\operatorname{Des} P = \{i \in \{1, 2, \ldots, n-1\} \mid \sigma(i) >_{\mathbb{Z}} \sigma(i+1)\}$$

(since P is the total order $(\sigma(1) < \sigma(2) < \cdots < \sigma(n))$). Comparing this with

$$\begin{aligned} \operatorname{Des} \sigma &= \{i \in \{1, 2, \ldots, n-1\} \mid \sigma(i) > \sigma(i+1)\} \qquad \text{(by the definition of } \operatorname{Des} \sigma) \\ &= \{i \in \{1, 2, \ldots, n-1\} \mid \sigma(i) >_{\mathbb{Z}} \sigma(i+1)\} \\ &\qquad \text{(since the greater relation } > \text{ of } \mathbb{Z} \text{ is the relation } >_{\mathbb{Z}}), \end{aligned}$$

we obtain $\operatorname{Des} \sigma = \operatorname{Des} P$.

Recall that $\gamma(\sigma)$ is the unique composition α of n satisfying $D(\alpha) = \operatorname{Des} \sigma$ (by the definition of $\gamma(\sigma)$). In other words, $\gamma(\sigma)$ is the unique composition $\alpha \in \operatorname{Comp}_n$ having partial sums $D(\alpha) = \operatorname{Des} \sigma$. In other words, $\gamma(\sigma)$ is the unique composition $\alpha \in \operatorname{Comp}_n$ having partial sums $D(\alpha) = \operatorname{Des} P$ (since $\operatorname{Des} \sigma = \operatorname{Des} P$).

Hence, Proposition 5.2.10 (applied to P, $\sigma(i)$ and $\gamma(\sigma)$ instead of w, w_i and α) yields that the generating function $F_P(\mathbf{x})$ equals the fundamental quasisymmetric function $L_{\gamma(\sigma)}$. Thus, $F_P(\mathbf{x}) = L_{\gamma(\sigma)}$. This proves Lemma 13.126.6. □

Proof of Lemma 5.3.6. We have $\mathfrak{A} = \{1 < 2 < 3 < \cdots\}$; thus, $\mathfrak{A} = \{1, 2, 3, \ldots\}$ as sets.

Let P be the labelled poset whose underlying set is $\{1, 2, \ldots, n\}$ and which (as a poset) is the total order $(\sigma(1) < \sigma(2) < \cdots < \sigma(n))$ (that is, the order $<_P$ is given by $\sigma(1) <_P \sigma(2) <_P \cdots <_P \sigma(n)$). Lemma 13.126.5 (applied to $\tau = \sigma$) yields that if $f : P \to \mathfrak{A}$ is any map, then we have the following logical equivalence:

$$(f \in \mathcal{A}(P)) \iff \left(\operatorname{std}(f(1), f(2), \ldots, f(n)) = \sigma^{-1}\right)$$

(where we treat $(f(1), f(2), \ldots, f(n))$ as a word in $\mathfrak{A}^n$). Hence, we have the following equality of summation signs:

$$\sum_{\substack{f : P \to \mathfrak{A}; \\ f \in \mathcal{A}(P)}} = \sum_{\substack{f : P \to \mathfrak{A}; \\ \operatorname{std}(f(1), f(2), \ldots, f(n)) = \sigma^{-1}}} . \tag{13.126.20}$$

[913] *Proof of (13.126.19):* Let $i \in P$ and $j \in P$ be such that $i <_P j$ and $i >_{\mathbb{Z}} j$. We must prove that $f(i) < f(j)$.

Assume the contrary. Thus, $f(i) \geq f(j)$. In other words, $f(j) \leq f(i)$.

Let $a = \tau^{-1}(i)$ and $b = \tau^{-1}(j)$. Thus, $a = \tau^{-1}(i) \in \{1, 2, \ldots, n\} = P$ and $b = \tau^{-1}(j) \in \{1, 2, \ldots, n\} = P$. From $a = \tau^{-1}(i)$, we obtain $\tau(a) = i$. From $b = \tau^{-1}(j)$, we obtain $\tau(b) = j$.

We have $i <_P j$. In view of $i = \tau(a)$ and $j = \tau(b)$, this rewrites as $\tau(a) <_P \tau(b)$.

But (13.126.11) (applied to a and b instead of i and j) shows that ($\tau(a) <_P \tau(b)$ if and only if $a < b$). Hence, we have $a < b$ (since we have $\tau(a) <_P \tau(b)$).

We have $i >_{\mathbb{Z}} j$. In other words, $i > j$. In other words, $j < i$. Hence, (13.126.14) (applied to j and i instead of a and b) shows that $((\operatorname{std} w)(j) < (\operatorname{std} w)(i)$ if and only if $f(j) \leq f(i))$. Thus, $(\operatorname{std} w)(j) < (\operatorname{std} w)(i)$ (since $f(j) \leq f(i)$). In view of $\operatorname{std} w = \tau^{-1}$, this rewrites as $\tau^{-1}(j) < \tau^{-1}(i)$. Thus, $b = \tau^{-1}(j) < \tau^{-1}(i) = a$. This contradicts $a < b$. This contradiction shows that our assumption was wrong. Hence, $f(i) < f(j)$. This proves (13.126.19).

But every P-partition is a function $P \to \{1,2,3,\ldots\}$. In other words, every P-partition is a function $P \to \mathfrak{A}$ (since $\{1,2,3,\ldots\} = \mathfrak{A}$). In other words, every $f \in \mathcal{A}(P)$ is a function $P \to \mathfrak{A}$ (since $\mathcal{A}(P)$ is the set of all P-partitions). Hence, we have the following equality of summation sums:

$$\sum_{f\in\mathcal{A}(P)} = \sum_{\substack{f:P\to\mathfrak{A};\\ f\in\mathcal{A}(P)}} = \sum_{\substack{f:P\to\mathfrak{A};\\ \operatorname{std}(f(1),f(2),\ldots,f(n))=\sigma^{-1}}}$$

(by (13.126.20)). But Lemma 13.126.6 yields $F_P(\mathbf{x}) = L_{\gamma(\sigma)}$. Hence,

$$\begin{aligned}
L_{\gamma(\sigma)} = F_P(\mathbf{x}) &= \underbrace{\sum_{f\in\mathcal{A}(P)}}_{=\sum_{\substack{f:P\to\mathfrak{A};\\ \operatorname{std}(f(1),f(2),\ldots,f(n))=\sigma^{-1}}}} \underbrace{\mathbf{x}_f}_{\substack{=\prod_{i\in P} x_{f(i)}\\ \text{(by the definition of } \mathbf{x}_f)}} \\
&\qquad \text{(by the definition of } F_P(\mathbf{x})) \\
&= \underbrace{\sum_{\substack{f:P\to\mathfrak{A};\\ \operatorname{std}(f(1),f(2),\ldots,f(n))=\sigma^{-1}}}}_{\substack{=\sum_{\substack{f:\{1,2,\ldots,n\}\to\mathfrak{A};\\ \operatorname{std}(f(1),f(2),\ldots,f(n))=\sigma^{-1}}}\\ \text{(since } P=\{1,2,\ldots,n\})}} \underbrace{\prod_{i\in P} x_{f(i)}}_{\substack{=x_{f(1)}x_{f(2)}\cdots x_{f(n)}\\ \text{(since } P=\{1,2,\ldots,n\})}}
\end{aligned}$$

$$= \sum_{\substack{f:\{1,2,\ldots,n\}\to\mathfrak{A};\\ \operatorname{std}(f(1),f(2),\ldots,f(n))=\sigma^{-1}}} x_{f(1)}x_{f(2)}\cdots x_{f(n)}. \tag{13.126.21}$$

But the map

$$\begin{aligned}
\{\text{functions } \{1,2,\ldots,n\} \to \mathfrak{A}\} &\to \mathfrak{A}^n, \\
f &\mapsto (f(1), f(2), \ldots, f(n))
\end{aligned}$$

is a bijection (indeed, this is just the standard bijection between the functions $\{1,2,\ldots,n\} \to \mathfrak{A}$ and the n-tuples of elements of $\mathfrak{A}$). Hence, we can substitute $(w_1, w_2, \ldots, w_n)$ for $(f(1), f(2), \ldots, f(n))$ in the sum on the right hand side of (13.126.21). We thus obtain

$$\sum_{\substack{f:\{1,2,\ldots,n\}\to\mathfrak{A};\\ \operatorname{std}(f(1),f(2),\ldots,f(n))=\sigma^{-1}}} x_{f(1)}x_{f(2)}\cdots x_{f(n)} = \sum_{\substack{(w_1,w_2,\ldots,w_n)\in\mathfrak{A}^n;\\ \operatorname{std}(w_1,w_2,\ldots,w_n)=\sigma^{-1}}} x_{w_1}x_{w_2}\cdots x_{w_n}.$$

Hence, (13.126.21) becomes

$$L_{\gamma(\sigma)} = \sum_{\substack{f:\{1,2,\ldots,n\}\to\mathfrak{A};\\ \operatorname{std}(f(1),f(2),\ldots,f(n))=\sigma^{-1}}} x_{f(1)}x_{f(2)}\cdots x_{f(n)} = \sum_{\substack{(w_1,w_2,\ldots,w_n)\in\mathfrak{A}^n;\\ \operatorname{std}(w_1,w_2,\ldots,w_n)=\sigma^{-1}}} x_{w_1}x_{w_2}\cdots x_{w_n}.$$

Comparing this with

$$\begin{aligned}
\sum_{\substack{w\in\mathfrak{A}^n;\\ \operatorname{std} w=\sigma^{-1}}} \mathbf{x}_w &= \sum_{\substack{(w_1,w_2,\ldots,w_n)\in\mathfrak{A}^n;\\ \operatorname{std}(w_1,w_2,\ldots,w_n)=\sigma^{-1}}} \underbrace{\mathbf{x}_{(w_1,w_2,\ldots,w_n)}}_{\substack{=x_{w_1}x_{w_2}\cdots x_{w_n}\\ \text{(by the definition of } \mathbf{x}_{(w_1,w_2,\ldots,w_n)})}} \\
&\qquad \text{(here, we have renamed the summation index } w \text{ as } (w_1,w_2,\ldots,w_n)) \\
&= \sum_{\substack{(w_1,w_2,\ldots,w_n)\in\mathfrak{A}^n;\\ \operatorname{std}(w_1,w_2,\ldots,w_n)=\sigma^{-1}}} x_{w_1}x_{w_2}\cdots x_{w_n},
\end{aligned}$$

we obtain $L_{\gamma(\sigma)} = \sum_{\substack{w\in\mathfrak{A}^n;\\ \operatorname{std} w=\sigma^{-1}}} \mathbf{x}_w$. This proves Lemma 5.3.6. □

Proposition 5.3.2 and Lemma 5.3.6 have now been proven. Thus, Exercise 5.3.7 is solved.

13.127. **Solution to Exercise 5.4.5.** *Solution to Exercise 5.4.5.* Consider the power series $\widetilde{H}(t)$ and $\xi(t)$ defined in (5.4.6). From (5.4.6), we know that

$$\sum_{n\geq 1} \xi_n t^n = \log \widetilde{H}(t) = \log\left(\sum_{n\geq 0} H_n t^n\right). \tag{13.127.1}$$

(a) The fact that ξ_n is primitive was proven for $\mathbf{k} = \mathbb{Q}$ in Remark 5.4.4, and can be proven in the same way for general $\mathbf{k}$. It remains to show that ξ_n is homogeneous of degree n for each $n \geq 1$. This can be done as follows:

Let us say that a power series $f \in A[[t]]$ over a graded $\mathbf{k}$-algebra A is *equigraded* if, for every $n \in \mathbb{N}$, the coefficient of f before t^n is homogeneous of degree n. Then, the set of all equigraded power series in $A[[t]]$ is a $\mathbf{k}$-subalgebra of $A[[t]]$ which is closed under the usual topology on $A[[t]]$. In particular, if g is an equigraded power series in $A[[t]]$ having constant term 1, then $\log g$ is equigraded. Applied to $A = \operatorname{NSym}$ and $g = \sum_{n\geq 0} H_n t^n$, this yields that the power series $\log\left(\sum_{n\geq 0} H_n t^n\right)$ is equigraded. Due to (13.127.1), this rewrites as follows: The power series $\sum_{n\geq 1} \xi_n t^n$ is equigraded. In other words, ξ_n is homogeneous of degree n for each $n \geq 1$. This completes the solution of Exercise 5.4.5(a).

For an alternative proof of Exercise 5.4.5(a), one can simply notice that it follows immediately from Exercise 5.4.5(c).

(b) The ring homomorphism $\pi : \operatorname{NSym} \to \Lambda$ induces a ring homomorphism $\operatorname{NSym}[[t]] \to \Lambda[[t]]$ which is continuous with respect to the usual topology on power series. Applying this latter homomorphism to the equality (13.127.1), we obtain

$$\sum_{n\geq 1} \pi(\xi_n) t^n = \log\left(\sum_{n\geq 0} \pi(H_n) t^n\right) = \log\left(\sum_{n\geq 0} h_n t^n\right) = \sum_{m=1}^{\infty} \frac{1}{m} p_m t^m$$

(where in the last step, we have used the equality (2.5.12)). Comparing coefficients, we obtain $\pi(\xi_n) = \frac{1}{n} p_n$ for all $n \geq 1$. Multiplying this by n, we obtain $\pi(n\xi_n) = p_n$ for all $n \geq 1$. This solves part (b) of the exercise.

Remark: Another way to solve Exercise 5.4.5(b) proceeds as follows: We know that $\pi(n\xi_n)$ is a primitive homogeneous element of Λ of degree n (by Exercise 5.4.5(a) and since π is a graded homomorphism of Hopf algebras). But Exercise 3.1.9 shows that all such elements are scalar multiples of p_n. Thus, $\pi(n\xi_n)$ is a scalar multiple of p_n. Finding the scalar is easy (e.g., it can be obtained by specializing at (1)).

(c) From (13.127.1), we have

$$
\begin{aligned}
\sum_{n\geq 1} \xi_n t^n &= \log\left(\underbrace{\sum_{n\geq 0} H_n t^n}_{=1+\sum_{n\geq 1} H_n t^n}\right) = \log\left(1+\sum_{n\geq 1} H_n t^n\right) \\
&= \sum_{i\geq 1} \frac{(-1)^{i-1}}{i} \underbrace{\left(\sum_{n\geq 1} H_n t^n\right)^i}_{=\sum_{n_1,n_2,\ldots,n_i\geq 1}\left(H_{n_1}t^{n_1}\right)\left(H_{n_2}t^{n_2}\right)\cdots\left(H_{n_i}t^{n_i}\right)} \\
&\qquad \text{(by the Mercator series for the logarithm)} \\
&= \sum_{i\geq 1} \frac{(-1)^{i-1}}{i} \underbrace{\sum_{n_1,n_2,\ldots,n_i\geq 1}}_{\substack{=\sum\limits_{\substack{(n_1,n_2,\ldots,n_i) \text{ is a composition} \\ \text{of length } i}}}} \underbrace{\left(H_{n_1}t^{n_1}\right)\left(H_{n_2}t^{n_2}\right)\cdots\left(H_{n_i}t^{n_i}\right)}_{=H_{n_1}H_{n_2}\cdots H_{n_i}t^{n_1+n_2+\cdots+n_i}} \\
&= \sum_{i\geq 1} \frac{(-1)^{i-1}}{i} \sum_{\substack{(n_1,n_2,\ldots,n_i) \text{ is a composition} \\ \text{of length } i}} \underbrace{H_{n_1}H_{n_2}\cdots H_{n_i}}_{=H_{(n_1,n_2,\ldots,n_i)}} \underbrace{t^{n_1+n_2+\cdots+n_i}}_{=t^{\left|(n_1,n_2,\ldots,n_i)\right|}} \\
&= \sum_{i\geq 1} \frac{(-1)^{i-1}}{i} \sum_{\substack{(n_1,n_2,\ldots,n_i) \text{ is a composition} \\ \text{of length } i}} H_{(n_1,n_2,\ldots,n_i)} t^{\left|(n_1,n_2,\ldots,n_i)\right|} \\
&= \sum_{i\geq 1} \frac{(-1)^{i-1}}{i} \sum_{\substack{\alpha \text{ is a composition} \\ \text{of length } i}} H_\alpha t^{|\alpha|} \\
&\qquad \text{(here, we renamed the summation index } (n_1, n_2, \ldots, n_i) \text{ as } \alpha \text{ in the inner sum)} \\
&= \sum_{i\geq 1} \sum_{\substack{\alpha \text{ is a composition} \\ \text{of length } i}} \underbrace{\frac{(-1)^{i-1}}{i}}_{\substack{=\frac{(-1)^{\ell(\alpha)-1}}{\ell(\alpha)} \\ (\text{since } i=\ell(\alpha))}} H_\alpha t^{|\alpha|} = \underbrace{\sum_{i\geq 1} \sum_{\substack{\alpha \text{ is a composition} \\ \text{of length } i}}}_{\substack{=\sum\limits_{\substack{\alpha \text{ is a nonempty} \\ \text{composition}}} \\ =\sum_{n\geq 1}\sum_{\alpha\in\operatorname{Comp}_n}}} \frac{(-1)^{\ell(\alpha)-1}}{\ell(\alpha)} H_\alpha t^{|\alpha|} \\
&= \sum_{n\geq 1} \sum_{\alpha\in\operatorname{Comp}_n} \frac{(-1)^{\ell(\alpha)-1}}{\ell(\alpha)} H_\alpha \underbrace{t^{|\alpha|}}_{\substack{=t^n \\ (\text{since } \alpha\in\operatorname{Comp}_n)}} = \sum_{n\geq 1} \sum_{\alpha\in\operatorname{Comp}_n} \frac{(-1)^{\ell(\alpha)-1}}{\ell(\alpha)} H_\alpha t^n \\
&= \sum_{n\geq 1} \left(\sum_{\alpha\in\operatorname{Comp}_n} \frac{(-1)^{\ell(\alpha)-1}}{\ell(\alpha)} H_\alpha\right) t^n .
\end{aligned}
$$

Comparing coefficients in this identity, we conclude that every $n \geq 1$ satisfies

$$
\xi_n = \sum_{\alpha\in\operatorname{Comp}_n} \frac{(-1)^{\ell(\alpha)-1}}{\ell(\alpha)} H_\alpha = \sum_{\alpha\in\operatorname{Comp}_n} (-1)^{\ell(\alpha)-1} \frac{1}{\ell(\alpha)} H_\alpha .
$$

This solves part (c) of the exercise.

(d) Applying the exponential to $\sum_{n\geq 1}\xi_n t^n = \log H(t)$, we obtain $\exp\left(\sum_{n\geq 1}\xi_n t^n\right) = H(t)$. Thus,

$$H(t) = \exp\left(\sum_{n\geq 1}\xi_n t^n\right) = \sum_{i\geq 0}\frac{1}{i!}\underbrace{\left(\sum_{n\geq 1}\xi_n t^n\right)^i}_{=\sum_{n_1,n_2,\ldots,n_i\geq 1}(\xi_{n_1}t^{n_1})(\xi_{n_2}t^{n_2})\cdots(\xi_{n_i}t^{n_i})}$$

$$= \sum_{i\geq 0}\frac{1}{i!}\underbrace{\sum_{n_1,n_2,\ldots,n_i\geq 1}}_{=\sum_{\substack{(n_1,n_2,\ldots,n_i)\text{ is a composition}\\ \text{of length } i}}}\underbrace{(\xi_{n_1}t^{n_1})(\xi_{n_2}t^{n_2})\cdots(\xi_{n_i}t^{n_i})}_{=\xi_{n_1}\xi_{n_2}\cdots\xi_{n_i}t^{n_1+n_2+\cdots+n_i}}$$

$$= \sum_{i\geq 0}\frac{1}{i!}\sum_{\substack{(n_1,n_2,\ldots,n_i)\text{ is a composition}\\ \text{of length } i}}\underbrace{\xi_{n_1}\xi_{n_2}\cdots\xi_{n_i}}_{\substack{=\xi_{(n_1,n_2,\ldots,n_i)}\\ \text{(because we defined }\xi_{(n_1,n_2,\ldots,n_i)}\\ \text{in such a way that this holds)}}}\underbrace{t^{n_1+n_2+\ldots+n_i}}_{=t^{|(n_1,n_2,\ldots,n_i)|}}$$

$$= \sum_{i\geq 0}\frac{1}{i!}\sum_{\substack{(n_1,n_2,\ldots,n_i)\text{ is a composition}\\ \text{of length } i}}\xi_{(n_1,n_2,\ldots,n_i)}t^{|(n_1,n_2,\ldots,n_i)|} = \sum_{i\geq 0}\frac{1}{i!}\sum_{\substack{\alpha\text{ is a composition}\\ \text{of length } i}}\xi_\alpha t^{|\alpha|}$$

$$\text{(here, we renamed the summation index } (n_1,n_2,\ldots,n_i) \text{ as } \alpha \text{ in the inner sum)}$$

$$= \sum_{i\geq 0}\sum_{\substack{\alpha\text{ is a composition}\\ \text{of length } i}}\underbrace{\frac{1}{i!}}_{\substack{=\frac{1}{\ell(\alpha)!}\\ \text{(since } i=\ell(\alpha))}}\xi_\alpha t^{|\alpha|} = \underbrace{\sum_{i\geq 0}\sum_{\substack{\alpha\text{ is a composition}\\ \text{of length } i}}}_{\substack{=\sum_{\alpha\text{ is a composition}}\\ =\sum_{n\geq 0}\sum_{\alpha\in\mathrm{Comp}_n}}}\frac{1}{\ell(\alpha)!}\xi_\alpha t^{|\alpha|}$$

$$= \sum_{n\geq 0}\sum_{\alpha\in\mathrm{Comp}_n}\frac{1}{\ell(\alpha)!}\xi_\alpha\underbrace{t^{|\alpha|}}_{\substack{=t^n\\ \text{(since } \alpha\in\mathrm{Comp}_n\text{)}}} = \sum_{n\geq 0}\sum_{\alpha\in\mathrm{Comp}_n}\frac{1}{\ell(\alpha)!}\xi_\alpha t^n = \sum_{n\geq 0}\left(\sum_{\alpha\in\mathrm{Comp}_n}\frac{1}{\ell(\alpha)!}\xi_\alpha\right)t^n.$$

Comparing coefficients in this identity, we conclude that every $n\geq 0$ satisfies

$$H_n = \sum_{\alpha\in\mathrm{Comp}_n}\frac{1}{\ell(\alpha)!}\xi_\alpha$$

(since the coefficient of t^n in $H(t)$ is H_n). This proves (5.4.8).

Notice that, for every $n\geq 1$, the element ξ_n of NSym is homogeneous of degree n (by Exercise 5.4.5(a)). Hence, for every composition α, the element ξ_α of NSym is homogeneous of degree $|\alpha|$. In particular, for every $n\in\mathbb{N}$, it is clear that $(\xi_\alpha)_{\alpha\in\mathrm{Comp}_n}$ is a family of elements of NSym_n. We now need to prove that this family is a $\mathbf{k}$-basis of NSym_n for every $n\in\mathbb{N}$.

Let $\mathfrak{A}$ be the $\mathbf{k}$-subalgebra of NSym generated by the elements ξ_1, ξ_2, ξ_3, ... Then, $\mathfrak{A}$ contains ξ_α for every composition α (by the definition of ξ_α). Therefore, $\mathfrak{A}$ contains H_n for every $n\geq 1$ (by (5.4.7)). Consequently, $\mathfrak{A} = \mathrm{NSym}$ (because NSym is generated as a $\mathbf{k}$-algebra by H_1, H_2, H_3, ...). In other words, the $\mathbf{k}$-algebra NSym is generated by the elements ξ_1, ξ_2, ξ_3, ... (since we defined $\mathfrak{A}$ as the $\mathbf{k}$-subalgebra of NSym generated by the elements ξ_1, ξ_2, ξ_3, ...). In other words, the $\mathbf{k}$-module NSym is spanned by all possible products of the elements ξ_1, ξ_2, ξ_3, ... In other words, the $\mathbf{k}$-module NSym is spanned by the elements ξ_α with $\alpha\in\mathrm{Comp}$ (because the elements ξ_α with $\alpha\in\mathrm{Comp}$ are precisely all possible products of the elements ξ_1, ξ_2, ξ_3, ...). In yet other words, the family $(\xi_\alpha)_{\alpha\in\mathrm{Comp}}$ spans the $\mathbf{k}$-module NSym.

Now, fix $n\in\mathbb{N}$. Every element of NSym_n can be written as a $\mathbf{k}$-linear combination of the elements ξ_α with $\alpha\in\mathrm{Comp}$ (since the family $(\xi_\alpha)_{\alpha\in\mathrm{Comp}}$ spans the $\mathbf{k}$-module NSym). In this $\mathbf{k}$-linear combination, we can remove all terms ξ_α with $\alpha\notin\mathrm{Comp}_n$ without changing the result (by gradedness, because ξ_α is homogeneous of degree $|\alpha|$), and so we conclude that every element of NSym_n can be written as a $\mathbf{k}$-linear combination of the elements ξ_α with $\alpha\in\mathrm{Comp}_n$. In other words, the family $(\xi_\alpha)_{\alpha\in\mathrm{Comp}_n}$ spans the $\mathbf{k}$-module NSym_n.

Now, we can apply Exercise 2.5.18(b) to $A = \operatorname{NSym}_n$, $I = \operatorname{Comp}_n$, $(\gamma_i)_{i\in I} = (H_\alpha)_{\alpha\in\operatorname{Comp}_n}$ and $(\beta_i)_{i\in I} = (\xi_\alpha)_{\alpha\in\operatorname{Comp}_n}$ (since we know that $(H_\alpha)_{\alpha\in\operatorname{Comp}_n}$ is a $\mathbf{k}$-basis of NSym_n, whereas $(\xi_\alpha)_{\alpha\in\operatorname{Comp}_n}$ spans the $\mathbf{k}$-module NSym_n). We conclude that $(\xi_\alpha)_{\alpha\in\operatorname{Comp}_n}$ is a $\mathbf{k}$-basis of NSym_n. Thus, Exercise 5.4.5(d) is solved.

13.128. **Solution to Exercise 5.4.6.** *Solution to Exercise 5.4.6.* We follow the hint.

Let f be the endomorphism $\operatorname{id}_A - u\epsilon$ of A. Then, $f = \sum_{n\geq 1}\pi_n$ (because the definition of the π_n yields that $\operatorname{id}_A = \sum_{n\geq 0}\pi_n = \underbrace{\pi_0}_{=u\epsilon} + \sum_{n\geq 1}\pi_n = u\epsilon + \sum_{n\geq 1}\pi_n$ and thus $\operatorname{id}_A - u\epsilon = \sum_{n\geq 1}\pi_n$). Notice that $f = \operatorname{id}_A - u\epsilon$, so that $\operatorname{id}_A = f + u\epsilon$. Now, $\mathfrak{e} = \log^\star\left(\underbrace{\operatorname{id}_A}_{=f+u\epsilon}\right) = \log^\star(f+u\epsilon) = \sum_{n\geq 1}(-1)^{n-1}\frac{1}{n}f^{\star n}$.

We let $\operatorname{End}_{\mathrm{gr}} A$ be the $\mathbf{k}$-submodule of $\operatorname{End} A$ consisting of all graded $\mathbf{k}$-linear maps $A \to A$. Then, it is easy to see that this $\mathbf{k}$-submodule $\operatorname{End}_{\mathrm{gr}} A$ is closed under convolution (i.e., if g_1 and g_2 are two graded $\mathbf{k}$-linear maps $A \to A$, then $g_1 \star g_2$ is also a graded $\mathbf{k}$-linear map $A \to A$) and contains the unity $u\epsilon$ of the algebra $(\operatorname{End} A, \star)$ (since $u\epsilon : A \to A$ is a graded $\mathbf{k}$-linear map). Hence, $(\operatorname{End}_{\mathrm{gr}} A, \star)$ is a $\mathbf{k}$-subalgebra of $(\operatorname{End} A, \star)$. Moreover, $(\operatorname{End}_{\mathrm{gr}} A, \star)$ contains π_n for every $n \geq 1$. Hence, the $\mathbf{k}$-algebra homomorphism $\mathfrak{W} : \operatorname{NSym} \to (\operatorname{End} A, \star)$ maps the generators H_n of NSym to elements of $(\operatorname{End}_{\mathrm{gr}} A, \star)$. Thus, the image of $\mathfrak{W}$ is contained in $(\operatorname{End}_{\mathrm{gr}} A, \star)$. In other words,

$$\mathfrak{W}(x) \in \operatorname{End}_{\mathrm{gr}} A \qquad \text{for every } x \in \operatorname{NSym}. \tag{13.128.1}$$

It is immediate to check that the $\mathbf{k}$-subalgebra $(\operatorname{End}_{\mathrm{gr}} A, \star)$ of $(\operatorname{End} A, \star)$ is closed under the topology of pointwise convergence. Hence, it is closed under taking logarithms. In other words, $\log^\star g \in \operatorname{End}_{\mathrm{gr}} A$ for every $g \in \operatorname{End}_{\mathrm{gr}} A$ for which $\log^\star g$ makes sense. Applied to $g = \operatorname{id}_A$, this yields that $\log^\star(\operatorname{id}_A) \in \operatorname{End}_{\mathrm{gr}} A$ (since $\operatorname{id}_A \in \operatorname{End}_{\mathrm{gr}} A$). Since $\log^\star(\operatorname{id}_A) = \mathfrak{e}$, this yields that $\mathfrak{e} \in \operatorname{End}_{\mathrm{gr}} A$. In other words, $\mathfrak{e}$ is a graded $\mathbf{k}$-linear map. The definition of $\mathfrak{e}_n$ is legitimate because the gradedness of $\mathfrak{e}$ yields $\pi_n \circ \mathfrak{e} = \mathfrak{e} \circ \pi_n$. This solves Exercise 5.4.6(a).

We notice for future use the fact that $\mathfrak{e}_0 = 0$ [914].

We also record the fact that

$$\Delta_A \circ \pi_n = \left(\sum_{k=0}^{n} \pi_k \otimes \pi_{n-k}\right) \circ \Delta_A \qquad \text{for all } n \in \mathbb{N}. \tag{13.128.2}$$

(This follows by checking that both sides of (13.128.2) are equal to Δ_A on the n-th homogeneous component A_n, while vanishing on all other components[915].)

[914] *Proof.* The definition of $\mathfrak{e}_0$ yields $\mathfrak{e}_0 = \pi_0 \circ \mathfrak{e} = \mathfrak{e} \circ \pi_0$. Thus, $\mathfrak{e}_0(A) = (\mathfrak{e} \circ \pi_0)(A) = \mathfrak{e}\left(\underbrace{\pi_0(A)}_{=\mathbf{k}\cdot 1_A}\right) = \mathfrak{e}(\mathbf{k}\cdot 1_A) = \mathbf{k}\cdot\mathfrak{e}(1_A)$. But every $n \geq 1$ satisfies $f^{\star n}(1_A) = 0$ (since $\Delta^{(n-1)}(1_A) = \underbrace{1_A \otimes 1_A \otimes \ldots \otimes 1_A}_{n \text{ times}}$ and $f(1_A) = 0$). Since $\mathfrak{e} = \sum_{n\geq 1}(-1)^{n-1}\frac{1}{n}f^{\star n}$, we have

$$\mathfrak{e}(1_A) = \left(\sum_{n\geq 1}(-1)^{n-1}\frac{1}{n}f^{\star n}\right)(1_A) = \sum_{n\geq 1}(-1)^{n-1}\frac{1}{n}\underbrace{f^{\star n}(1_A)}_{=0} = \sum_{n\geq 1}(-1)^{n-1}\frac{1}{n}0 = 0.$$

Now, $\mathfrak{e}_0(A) = \mathbf{k}\cdot\underbrace{\mathfrak{e}(1_A)}_{=0} = 0$, so that $\mathfrak{e}_0 = 0$, qed.

[915] And this is because Δ_A is graded.

(b) From (13.127.1), we have

$$\begin{aligned}\sum_{n\geq 1}\xi_n t^n &= \log\left(\sum_{n\geq 0}H_n t^n\right) = \log\left(1+\sum_{n\geq 1}H_n t^n\right)\\ &= \sum_{i\geq 1}(-1)^{i-1}\frac{1}{i}\underbrace{\left(\sum_{n\geq 1}H_n t^n\right)^i}_{=\sum_{n_1,n_2,\ldots,n_i\geq 1}H_{n_1}H_{n_2}\ldots H_{n_i}t^{n_1+n_2+\ldots+n_i}} \qquad \text{(by the Mercator series for the logarithm)}\\ &= \sum_{i\geq 1}(-1)^{i-1}\frac{1}{i}\sum_{n_1,n_2,\ldots,n_i\geq 1}H_{n_1}H_{n_2}\ldots H_{n_i}t^{n_1+n_2+\ldots+n_i}\\ &= \sum_{n\geq 1}\sum_{i\geq 1}(-1)^{i-1}\frac{1}{i}\sum_{\substack{n_1,n_2,\ldots,n_i\geq 1;\\ n_1+n_2+\ldots+n_i=n}}H_{n_1}H_{n_2}\ldots H_{n_i}t^n\end{aligned}$$

in the power series ring $\operatorname{NSym}[[t]]$. Comparing coefficients in this equality, we obtain

$$\xi_n = \sum_{i\geq 1}(-1)^{i-1}\frac{1}{i}\sum_{\substack{n_1,n_2,\ldots,n_i\geq 1;\\ n_1+n_2+\ldots+n_i=n}}H_{n_1}H_{n_2}\ldots H_{n_i} \tag{13.128.3}$$

for every $n\geq 1$. Hence, every $n\geq 1$ satisfies

$$\begin{aligned}\mathfrak{W}(\xi_n) &= \mathfrak{W}\left(\sum_{i\geq 1}(-1)^{i-1}\frac{1}{i}\sum_{\substack{n_1,n_2,\ldots,n_i\geq 1;\\ n_1+n_2+\ldots+n_i=n}}H_{n_1}H_{n_2}\ldots H_{n_i}\right)\\ &= \sum_{i\geq 1}(-1)^{i-1}\frac{1}{i}\sum_{\substack{n_1,n_2,\ldots,n_i\geq 1;\\ n_1+n_2+\ldots+n_i=n}}\mathfrak{W}(H_{n_1})\star\mathfrak{W}(H_{n_2})\star\ldots\star\mathfrak{W}(H_{n_i})\\ &\qquad \text{(since }\mathfrak{W}\text{ is a }\mathbf{k}\text{-algebra homomorphism)}\\ &= \sum_{i\geq 1}(-1)^{i-1}\frac{1}{i}\sum_{\substack{n_1,n_2,\ldots,n_i\geq 1;\\ n_1+n_2+\ldots+n_i=n}}\pi_{n_1}\star\pi_{n_2}\star\ldots\star\pi_{n_i} \qquad (13.128.4)\end{aligned}$$

(since $\mathfrak{W}$ maps every H_m to π_m).

On the other hand, $f=\sum_{n\geq 1}\pi_n$. Thus,

$$f^{\star i} = \left(\sum_{n\geq 1}\pi_n\right)^{\star i} = \sum_{n_1,n_2,\ldots,n_i\geq 1}\pi_{n_1}\star\pi_{n_2}\star\ldots\star\pi_{n_i}$$

for every $i\in\mathbb{N}$. Now,

$$\begin{aligned}\mathfrak{e} &= \sum_{n\geq 1}(-1)^{n-1}\frac{1}{n}f^{\star n} = \sum_{i\geq 1}(-1)^{i-1}\frac{1}{i}\underbrace{f^{\star i}}_{=\sum_{n_1,n_2,\ldots,n_i\geq 1}\pi_{n_1}\star\pi_{n_2}\star\ldots\star\pi_{n_i}}\\ &= \sum_{i\geq 1}(-1)^{i-1}\frac{1}{i}\sum_{n_1,n_2,\ldots,n_i\geq 1}\pi_{n_1}\star\pi_{n_2}\star\ldots\star\pi_{n_i}. \qquad (13.128.5)\end{aligned}$$

Now,

$$\begin{aligned}
\mathfrak{e}_n &= \pi_n \circ \mathfrak{e} \\
&= \pi_n \circ \left(\sum_{i\geq 1} (-1)^{i-1} \frac{1}{i} \sum_{n_1, n_2, \ldots, n_i \geq 1} \pi_{n_1} \star \pi_{n_2} \star \ldots \star \pi_{n_i} \right) \qquad \text{(by (13.128.5))} \\
&= \sum_{i\geq 1} (-1)^{i-1} \frac{1}{i} \sum_{n_1, n_2, \ldots, n_i \geq 1} \pi_n \circ \left(\pi_{n_1} \star \pi_{n_2} \star \ldots \star \pi_{n_i} \right) \\
&= \underbrace{\sum_{i\geq 1} (-1)^{i-1} \frac{1}{i} \sum_{\substack{n_1, n_2, \ldots, n_i \geq 1; \\ n_1+n_2+\ldots+n_i = n}} \pi_n \circ \left(\pi_{n_1} \star \pi_{n_2} \star \ldots \star \pi_{n_i} \right)}_{\substack{=\mathfrak{W}(\xi_n) \\ \text{(by (13.128.4))}}} \\
&\qquad + \sum_{i\geq 1} (-1)^{i-1} \frac{1}{i} \sum_{\substack{n_1, n_2, \ldots, n_i \geq 1; \\ n_1+n_2+\ldots+n_i \neq n}} \underbrace{\pi_n \circ \left(\pi_{n_1} \star \pi_{n_2} \star \ldots \star \pi_{n_i} \right)}_{\substack{=0 \\ \text{(because the image of} \\ \pi_{n_1} \star \pi_{n_2} \star \ldots \star \pi_{n_i} = m^{(i-1)} \circ \left(\pi_{n_1} \otimes \pi_{n_2} \otimes \ldots \otimes \pi_{n_i} \right) \circ \Delta^{(i-1)} \\ \text{is contained in} \\ m^{(i-1)} \left(\left(\pi_{n_1} \otimes \pi_{n_2} \otimes \ldots \otimes \pi_{n_i} \right) \left(A^{\otimes i} \right) \right) = A_{n_1} A_{n_2} \ldots A_{n_i} \subset A_{n_1+n_2+\ldots+n_i}, \\ \text{which is a different homogeneous component of } A \text{ than } A_n \text{)}}} \\
&= \mathfrak{W}\left(\xi_n\right) + \underbrace{\sum_{i\geq 1} (-1)^{i-1} \frac{1}{i} \sum_{\substack{n_1, n_2, \ldots, n_i \geq 1; \\ n_1+n_2+\ldots+n_i \neq n}} 0}_{=0} = \mathfrak{W}\left(\xi_n\right).
\end{aligned}$$

This solves Exercise 5.4.6(b).

(c) Exercise 1.5.6(a) (applied to $C = A$) yields that the comultiplication $\Delta_A : A \to A \otimes A$ is a **k**-coalgebra homomorphism (since A is cocommutative). Also, the comultiplications Δ_A and Δ_{NSym} are **k**-algebra homomorphisms (as is the comultiplication of any **k**-bialgebra).

Let R be the **k**-linear map $(\operatorname{End} A, \star) \otimes (\operatorname{End} A, \star) \to (\operatorname{End}(A \otimes A), \star)$ which sends every tensor $f \otimes g \in (\operatorname{End} A, \star) \otimes (\operatorname{End} A, \star)$ to the map $f \otimes g : A \otimes A \to A \otimes A$. This map R is a **k**-algebra homomorphism[916].

Let Ω_1 be the composition

$$\mathrm{NSym} \xrightarrow{\mathfrak{W}} (\operatorname{End} A, \star) \xrightarrow{\operatorname{post}(\Delta_A)} (\operatorname{Hom}(A, A \otimes A), \star),$$

where $\operatorname{post}(\Delta_A)$ denotes the **k**-linear map sending every $\gamma \in (\operatorname{End} A, \star)$ to $\Delta_A \circ \gamma \in (\operatorname{Hom}(A, A \otimes A), \star)$. Since the map $\mathfrak{W}$ is a **k**-algebra homomorphism, and since the map $\operatorname{post}(\Delta_A)$ is a **k**-algebra homomorphism[917], their composition Ω_1 also is a **k**-algebra homomorphism.

Let Ω_2 be the composition

$$\mathrm{NSym} \xrightarrow{\Delta_{\mathrm{NSym}}} \mathrm{NSym} \otimes \mathrm{NSym} \xrightarrow{\mathfrak{W} \otimes \mathfrak{W}} (\operatorname{End} A, \star) \otimes (\operatorname{End} A, \star) \xrightarrow{R} (\operatorname{End}(A \otimes A), \star) \xrightarrow{\operatorname{pre}(\Delta_A)} (\operatorname{Hom}(A, A \otimes A), \star),$$

where $\operatorname{pre}(\Delta_A)$ denotes the **k**-linear map sending every $\gamma \in (\operatorname{End}(A \otimes A), \star)$ to $\gamma \circ \Delta_A \in (\operatorname{Hom}(A, A \otimes A), \star)$. Since the maps Δ_{NSym}, $\mathfrak{W} \otimes \mathfrak{W}$ and R are **k**-algebra homomorphisms, and since the map $\operatorname{pre}(\Delta_A)$ is a **k**-algebra homomorphism[918], their composition Ω_2 also is a **k**-algebra homomorphism.

In order to solve Exercise 5.4.6(c), we need to show that every $w \in \mathrm{NSym}$ satisfies $\Delta \circ (\mathfrak{W}(w)) = \left(\sum_{(w)} \mathfrak{W}(w_1) \otimes \mathfrak{W}(w_2) \right) \circ \Delta$. Since

$$\Delta \circ (\mathfrak{W}(w)) = \Delta_A \circ (\mathfrak{W}(w)) = (\operatorname{post}(\Delta_A))(\mathfrak{W}(w)) = \underbrace{(\operatorname{post}(\Delta_A) \circ \mathfrak{W})}_{=\Omega_1}(w) = \Omega_1(w)$$

[916] In fact, this is a particular case of Exercise 1.4.4(b).

[917] This follows from Proposition 1.4.3 (applied to A, A, A, $A \otimes A$, id_A and Δ_A instead of C, C', A, A', γ and α) (because we know that Δ_A is a **k**-algebra homomorphism).

[918] This follows from Proposition 1.4.3 (applied to A, $A \otimes A$, $A \otimes A$, $A \otimes A$, Δ_A and $\mathrm{id}_{A \otimes A}$ instead of C, C', A, A', γ and α) (because we know that Δ_A is a **k**-coalgebra homomorphism).

and

$$\begin{aligned}\underbrace{\left(\sum_{(w)}\mathfrak{W}(w_1)\otimes\mathfrak{W}(w_2)\right)}_{=R((\mathfrak{W}\otimes\mathfrak{W})(\Delta_{\mathrm{NSym}}(w)))}\circ\underbrace{\Delta}_{=\Delta_A} &= \left(R\left((\mathfrak{W}\otimes\mathfrak{W})(\Delta_{\mathrm{NSym}}(w))\right)\right)\circ\Delta_A \\ &= (\operatorname{pre}(\Delta_A))\left(R\left((\mathfrak{W}\otimes\mathfrak{W})(\Delta_{\mathrm{NSym}}(w))\right)\right) \\ &= \underbrace{(\operatorname{pre}(\Delta_A)\circ R\circ(\mathfrak{W}\otimes\mathfrak{W})\circ\Delta_{\mathrm{NSym}})}_{=\Omega_2}(w)=\Omega_2(w),\end{aligned}$$

this is equivalent to showing that every $w\in\mathrm{NSym}$ satisfies $\Omega_1(w)=\Omega_2(w)$. In other words, we need to prove that $\Omega_1=\Omega_2$. Since Ω_1 and Ω_2 are $\mathbf{k}$-algebra homomorphisms, it will be enough to verify this on the generators $H_1,H_2,H_3,\ldots$ of the $\mathbf{k}$-algebra NSym. But on said generators, this is easily seen to hold, because every $n\geq 1$ satisfies

$$\begin{aligned}\Omega_1(H_n) &= (\operatorname{post}(\Delta_A)\circ\mathfrak{W})(H_n)=(\operatorname{post}(\Delta_A))\left(\underbrace{\mathfrak{W}(H_n)}_{=\pi_n}\right) \\ &= (\operatorname{post}(\Delta_A))(\pi_n)=\Delta_A\circ\pi_n=\left(\sum_{k=0}^{n}\pi_k\otimes\pi_{n-k}\right)\circ\Delta_A \qquad (\text{by } (13.128.2))\end{aligned}$$

and

$$\begin{aligned}\Omega_2(H_n) &= (\operatorname{pre}(\Delta_A)\circ R\circ(\mathfrak{W}\otimes\mathfrak{W})\circ\Delta_{\mathrm{NSym}})(H_n) \\ &= (\operatorname{pre}(\Delta_A)\circ R\circ(\mathfrak{W}\otimes\mathfrak{W}))\left(\underbrace{\Delta_{\mathrm{NSym}}(H_n)}_{=\sum_{k=0}^{n}H_k\otimes H_{n-k}}\right) \\ &= (\operatorname{pre}(\Delta_A)\circ R\circ(\mathfrak{W}\otimes\mathfrak{W}))\left(\sum_{k=0}^{n}H_k\otimes H_{n-k}\right) \\ &= (\operatorname{pre}(\Delta_A)\circ R)\left((\mathfrak{W}\otimes\mathfrak{W})\left(\sum_{k=0}^{n}H_k\otimes H_{n-k}\right)\right) \\ &= (\operatorname{pre}(\Delta_A)\circ R)\left(\sum_{k=0}^{n}\underbrace{\mathfrak{W}(H_k)}_{=\pi_k}\otimes\underbrace{\mathfrak{W}(H_{n-k})}_{=\pi_{n-k}}\right) \\ &= (\operatorname{pre}(\Delta_A)\circ R)\left(\sum_{k=0}^{n}\pi_k\otimes\pi_{n-k}\right) \\ &= (\operatorname{pre}(\Delta_A))\left(R\left(\sum_{k=0}^{n}\pi_k\otimes\pi_{n-k}\right)\right)=\underbrace{R\left(\sum_{k=0}^{n}\pi_k\otimes\pi_{n-k}\right)}_{=\sum_{k=0}^{n}\pi_k\otimes\pi_{n-k}}\circ\Delta_A \\ &= \left(\sum_{k=0}^{n}\pi_k\otimes\pi_{n-k}\right)\circ\Delta_A.\end{aligned}$$

Exercise 5.4.6(c) is proven.

(d) Let $n\geq 0$. We need to prove that $\mathfrak{e}_n(A)\subset\mathfrak{p}$. If $n=0$, then this is clear because $\mathfrak{e}_0=0$. Hence, we assume WLOG that we don't have $n=0$. Thus, $n\geq 1$. Hence, Exercise 5.4.5(a) yields that ξ_n is primitive, so that $\Delta(\xi_n)=\xi_n\otimes 1+1\otimes\xi_n$.

Applying Exercise 5.4.6(c) to $w = \xi_n$, we obtain

$$\begin{aligned}\Delta \circ (\mathfrak{W}(\xi_n)) &= \left(\sum_{(\xi_n)} \mathfrak{W}((\xi_n)_1) \otimes \mathfrak{W}((\xi_n)_2) \right) \circ \Delta \\ &= (\mathfrak{W}(\xi_n) \otimes \mathfrak{W}(1) + \mathfrak{W}(1) \otimes \mathfrak{W}(\xi_n)) \circ \Delta \\ &\qquad \left(\text{since } \sum_{(\xi_n)} (\xi_n)_1 \otimes (\xi_n)_2 = \Delta(\xi_n) = \xi_n \otimes 1 + 1 \otimes \xi_n \right).\end{aligned}$$

Since $\mathfrak{W}(1) = u\epsilon$ and $\mathfrak{W}(\xi_n) = \mathfrak{e}_n$, this rewrites as

$$\Delta \circ \mathfrak{e}_n = (\mathfrak{e}_n \otimes (u\epsilon) + (u\epsilon) \otimes \mathfrak{e}_n) \circ \Delta.$$

Now, let $x \in A$. Then,

$$\begin{aligned}\Delta(\mathfrak{e}_n(x)) &= \underbrace{(\Delta \circ \mathfrak{e}_n)}_{=(\mathfrak{e}_n \otimes (u\epsilon) + (u\epsilon) \otimes \mathfrak{e}_n) \circ \Delta} (x) = ((\mathfrak{e}_n \otimes (u\epsilon) + (u\epsilon) \otimes \mathfrak{e}_n) \circ \Delta)(x) \\ &= (\mathfrak{e}_n \otimes (u\epsilon))(\Delta(x)) + ((u\epsilon) \otimes \mathfrak{e}_n)(\Delta(x)) \\ &= \sum_{(x)} \mathfrak{e}_n(x_1) \otimes \underbrace{(u\epsilon)(x_2)}_{=1_A \epsilon(x_2)} + \sum_{(x)} \underbrace{(u\epsilon)(x_1)}_{=1_A \epsilon(x_1)} \otimes \mathfrak{e}_n(x_2) \\ &= \sum_{(x)} \mathfrak{e}_n(x_1) \otimes 1_A \epsilon(x_2) + \sum_{(x)} 1_A \epsilon(x_1) \otimes \mathfrak{e}_n(x_2) \\ &= \mathfrak{e}_n \left(\underbrace{\sum_{(x)} x_1 \epsilon(x_2)}_{=x} \right) \otimes 1_A + 1_A \otimes \mathfrak{e}_n \left(\underbrace{\sum_{(x)} \epsilon(x_1) x_2}_{=x} \right) \\ &= \mathfrak{e}_n(x) \otimes 1_A + 1_A \otimes \mathfrak{e}_n(x).\end{aligned}$$

Hence, $\mathfrak{e}_n(x)$ is a primitive element of A; thus, $\mathfrak{e}_n(x) \in \mathfrak{p}$. Forget now that we fixed x. We thus have seen that every $x \in A$ satisfies $\mathfrak{e}_n(x) \in \mathfrak{p}$. In other words, $\mathfrak{e}_n(A) \subset \mathfrak{p}$. This solves Exercise 5.4.6(d).

(e) Every $n \geq 0$ satisfies $\mathfrak{e}_n(A) \subset \mathfrak{p}$ (by Exercise 5.4.6(d)).

We have $\operatorname{id}_A = \sum_{n \geq 0} \pi_n$ (by the definition of π_n), and

$$\mathfrak{e} = \underbrace{\operatorname{id}_A}_{=\sum_{n \geq 0} \pi_n} \circ \mathfrak{e} = \sum_{n \geq 0} \underbrace{\pi_n \circ \mathfrak{e}}_{\substack{=\mathfrak{e}_n \\ \text{(since the definition of } \mathfrak{e}_n \\ \text{yields } \mathfrak{e}_n = \pi_n \circ \mathfrak{e})}} = \sum_{n \geq 0} \mathfrak{e}_n.$$

Hence, $\mathfrak{e}(A) = \left(\sum_{n \geq 0} \mathfrak{e}_n \right)(A) \subset \sum_{n \geq 0} \underbrace{\mathfrak{e}_n(A)}_{\subset \mathfrak{p}} \subset \sum_{n \geq 0} \mathfrak{p} \subset \mathfrak{p}$. This proves part (e) of the exercise.

(f) Let $x \in \mathfrak{p}$. Then, $\Delta(x) = x \otimes 1_A + 1_A \otimes x$.

Let $\mathfrak{g}(A, A)$ denote the $\mathbf{k}$-submodule of $\operatorname{End} A$ which consists of all $g \in \operatorname{End} A$ satisfying $g(1_A) = 0$. Then, it is easy to see that $\mathfrak{g}(A, A)$ is an ideal of the algebra $(\operatorname{End} A, \star)$ (indeed, any $g_1 \in \operatorname{End} A$ and $g_2 \in \mathfrak{g}(A, A)$ satisfy $g_1 \star g_2 \in \mathfrak{g}(A, A)$ and $g_2 \star g_1 \in \mathfrak{g}(A, A)$). We have $f \in \mathfrak{g}(A, A)$. Since $\mathfrak{g}(A, A)$ is an ideal of $(\operatorname{End} A, \star)$, this yields that every $n \geq 2$ satisfies $f^{\star n} \in (\mathfrak{g}(A, A))^{\star 2}$ (where $I^{\star 2}$ denotes the square of the ideal I for any ideal I of the $\mathbf{k}$-algebra $(\operatorname{End} A, \star)$). But it is easy to see that

$$g(x) = 0 \qquad \text{for every } g \in (\mathfrak{g}(A, A))^{\star 2}. \tag{13.128.6}$$

[919] In particular, this yields that

$$f^{\star n}(x) = 0 \qquad \text{for every } n \geq 2 \tag{13.128.7}$$

(since $f^{\star n} \in (\mathfrak{g}(A,A))^{\star 2}$ for every $n \geq 2$). Now, $\mathfrak{e} = \sum_{n \geq 1} (-1)^{n-1} \frac{1}{n} f^{\star n}$, so that

$$\begin{aligned}
\mathfrak{e}(x) &= \left(\sum_{n \geq 1} (-1)^{n-1} \frac{1}{n} f^{\star n} \right)(x) = \sum_{n \geq 1} (-1)^{n-1} \frac{1}{n} f^{\star n}(x) \\
&= \underbrace{(-1)^{1-1} \frac{1}{1}}_{=1} \underbrace{f^{\star 1}}_{=f = \operatorname{id}_A - u\epsilon} (x) + \sum_{n \geq 2} (-1)^{n-1} \frac{1}{n} \underbrace{f^{\star n}(x)}_{\substack{=0 \\ \text{(by (13.128.7))}}} \\
&= (\operatorname{id}_A - u\epsilon)(x) + \underbrace{\sum_{n \geq 2} (-1)^{n-1} \frac{1}{n} 0}_{=0} = (\operatorname{id}_A - u\epsilon)(x) = x - u \left(\underbrace{\epsilon(x)}_{\substack{=0 \\ \text{(by Proposition 1.4.17)}}} \right) = x - \underbrace{u(0)}_{=0} = x.
\end{aligned}$$

Now, forget that we fixed x. We thus have shown that $\mathfrak{e}(x) = x$ for every $x \in \mathfrak{p}$. In other words, the map $\mathfrak{e}$ fixes any element of $\mathfrak{p}$. This solves Exercise 5.4.6(f).

Combining the results of parts (e) and (f) of Exercise 5.4.6, we conclude that $\mathfrak{e}$ is a projection from A to the $\mathbf{k}$-submodule $\mathfrak{p}$. This completes the solution of the exercise.

13.129. **Solution to Exercise 5.4.8.** *Solution to Exercise 5.4.8.* Let us first show the following two lemmas, which have nothing to do with Hopf algebras:

Lemma 13.129.1. *Let V be any torsionfree abelian group (written additively). Let $N \in \mathbb{N}$. For every $k \in \{0, 1, ..., N\}$, let w_k be an element of V. Assume that*

$$\sum_{k=0}^{N} w_k n^k = 0 \qquad \text{for all } n \in \mathbb{N}. \tag{13.129.1}$$

Then, $w_k = 0$ for every $k \in \{0, 1, ..., N\}$.

Lemma 13.129.2. *Let V be any torsionfree abelian group (written additively). Let $N \in \mathbb{N}$. For every $(k, \ell) \in \{0, 1, ..., N\}^2$, let $v_{k,\ell}$ be an element of V. Assume that*

$$\sum_{k=0}^{N} \sum_{\ell=0}^{N} v_{k,\ell} n^k m^\ell = 0 \qquad \text{for all } n \in \mathbb{N} \text{ and } m \in \mathbb{N}. \tag{13.129.2}$$

Then, $v_{k,\ell} = 0$ for every $(k, \ell) \in \{0, 1, ..., N\}^2$.

Proof of Lemma 13.129.1. Lemma 13.129.1 has already appeared above (namely, as Lemma 1.7.24), and has already been proven (in the solution to Exercise 1.7.28). □

Proof of Lemma 13.129.2. Fix $m \in \mathbb{N}$. Every $n \in \mathbb{N}$ satisfies

$$\sum_{k=0}^{N} \left(\sum_{\ell=0}^{N} v_{k,\ell} m^\ell \right) n^k = \sum_{k=0}^{N} \sum_{\ell=0}^{N} v_{k,\ell} n^k m^\ell = 0.$$

Thus, Lemma 13.129.1 (applied to $w_k = \sum_{\ell=0}^{N} v_{k,\ell} m^\ell$) yields that $\sum_{\ell=0}^{N} v_{k,\ell} m^\ell = 0$ for every $k \in \{0, 1, ..., N\}$.

[919] *Proof of* (13.128.6)*:* It is clearly enough to check that $(g_1 \star g_2)(x) = 0$ for every $g_1 \in \mathfrak{g}(A,A)$ and $g_2 \in \mathfrak{g}(A,A)$. But this is easy: If $g_1 \in \mathfrak{g}(A,A)$ and $g_2 \in \mathfrak{g}(A,A)$, then $g_1(1_A) = 0$ and $g_2(1_A) = 0$, so that

$$\begin{aligned}
(g_1 \star g_2)(x) &= g_1(x) \underbrace{g_2(1_A)}_{=0} + \underbrace{g_1(1_A)}_{=0} g_2(x) \qquad (\text{since } \Delta(x) = x \otimes 1_A + 1_A \otimes x) \\
&= 0 + 0 = 0.
\end{aligned}$$

This proves (13.128.6).

Now, forget that we fixed m. We thus have proven that

$$\sum_{\ell=0}^{N} v_{k,\ell} m^{\ell} = 0 \qquad \text{for every } m \in \mathbb{N} \text{ and } k \in \{0,1,...,N\}. \tag{13.129.3}$$

Now, fix $g \in \{0,1,...,N\}$. For every $n \in \mathbb{N}$, we have

$$\begin{aligned} \sum_{k=0}^{N} v_{g,k} n^k &= \sum_{\ell=0}^{N} v_{g,\ell} n^{\ell} \qquad \text{(here, we renamed the summation index } k \text{ as } \ell\text{)} \\ &= 0 \qquad \text{(by (13.129.3), applied to } k = g \text{ and } m = n\text{)}. \end{aligned}$$

Hence, Lemma 13.129.1 (applied to $w_k = v_{g,k}$) yields that $v_{g,k} = 0$ for every $k \in \{0,1,...,N\}$.

Now, forget that we fixed g. We thus have shown that $v_{g,k} = 0$ for every $g \in \{0,1,...,N\}$ and $k \in \{0,1,...,N\}$. Renaming the indices g and k as k and ℓ in this statement, we obtain the following: We have $v_{k,\ell} = 0$ for every $k \in \{0,1,...,N\}$ and $\ell \in \{0,1,...,N\}$. In other words, $v_{k,\ell} = 0$ for every $(k,\ell) \in \{0,1,...,N\}^2$. This proves Lemma 13.129.2. □

Now, let us come to the solution of Exercise 5.4.8.

Define $\operatorname{End}_{\mathrm{gr}} A$ as in the solution of Exercise 5.4.6. We can prove (just as in the solution of Exercise 5.4.6) that $(\operatorname{End}_{\mathrm{gr}} A, \star)$ is a $\mathbf{k}$-subalgebra of $(\operatorname{End} A, \star)$. Since $\mathrm{id}_A \in \operatorname{End}_{\mathrm{gr}} A$, we thus have $\mathrm{id}_A^{\star \ell} \in \operatorname{End}_{\mathrm{gr}} A$ for every $\ell \in \mathbb{N}$. In other words,

$$\mathrm{id}_A^{\star \ell} \text{ is a graded } \mathbf{k}\text{-linear map for every } \ell \in \mathbb{N}. \tag{13.129.4}$$

We have $\mathfrak{e}(1_A) = 0$ (as was proven in a footnote in the solution of Exercise 5.4.6). Hence, $\mathfrak{e}\left(\underbrace{A_0}_{=\mathbf{k}\cdot 1_A}\right) = \mathfrak{e}(\mathbf{k} \cdot 1_A) = \mathbf{k} \cdot \underbrace{\mathfrak{e}(1_A)}_{=0} = 0$.

Notice that A is a $\mathbf{k}$-module, hence a $\mathbb{Q}$-module (since $\mathbb{Q}$ is a subring of $\mathbf{k}$), thus a torsionfree abelian group.

(a) For every $\mathbf{k}$-linear map $f : A \to A$ which annihilates A_0, we can define an endomorphism $\exp^\star f$ of A by setting $\exp^\star f = \sum_{\ell \geq 0} \frac{1}{\ell!} f^{\star \ell}$. (This follows from the same argument as the well-definedness of $\log^\star (f + u\epsilon)$. [920]) The usual rules for exponentials and logarithms apply:

- We have $\exp^\star(\log^\star(f + u\epsilon)) = f + u\epsilon$ for every $\mathbf{k}$-linear map $f : A \to A$ which annihilates A_0. [921]
- We have $\log^\star(\exp^\star f) = f$ for every $\mathbf{k}$-linear map $f : A \to A$ which annihilates A_0. [922]
- We have $\exp^\star(f + g) = (\exp^\star f) \star (\exp^\star g)$ for any two $\mathbf{k}$-linear maps $f : A \to A$ and $g : A \to A$ which annihilate A_0 and satisfy $f \star g = g \star f$. [923]

[920] This definition of $\exp^\star f$ is actually a particular case of Definition 1.7.10(d). This can be shown as follows: If $f : A \to A$ is a $\mathbf{k}$-linear map which annihilates A_0, then Proposition 1.7.11(h) (applied to $C = A$) yields $f \in \mathfrak{n}(A,A)$. Therefore, Definition 1.7.10(d) defines a map $\exp^\star f \in \mathfrak{n}(A,A)$. This map is identical to the map $\exp^\star f := \sum_{\ell \geq 0} \frac{1}{\ell!} f^{\star \ell}$ we have just defined, because the map $\exp^\star f$ defined using Definition 1.7.10(d) satisfies

$$\begin{aligned} \exp^\star f &= \sum_{n \geq 0} \frac{1}{n!} f^{\star n} \qquad \left(\text{since } \exp = \sum_{n \geq 0} \frac{1}{n!} T^n\right) \\ &= \sum_{\ell \geq 0} \frac{1}{\ell!} f^{\star \ell}. \end{aligned}$$

[921] Indeed, this follows from Proposition 1.7.18(b) (applied to $C = A$ and $g = f + u\epsilon$), after first observing that $f \in \mathfrak{n}(A,A)$ (by Proposition 1.7.11(h), applied to $C = A$).

[922] Indeed, this follows from Proposition 1.7.18(a) (applied to $C = A$), after first observing that $f \in \mathfrak{n}(A,A)$ (by Proposition 1.7.11(h), applied to $C = A$).

[923] Indeed, this follows from Proposition 1.7.18(c) (applied to $C = A$), after first observing that $f \in \mathfrak{n}(A,A)$ (by Proposition 1.7.11(h), applied to $C = A$) and $g \in \mathfrak{n}(A,A)$ (for similar reasons).

- We have $\exp^\star(nf) = (\exp^\star f)^{\star n}$ for every $n \in \mathbb{N}$ and any $\mathbf{k}$-linear map $f : A \to A$ which annihilates A_0. [924]

The map $\mathfrak{e}$ annihilates A_0 (since $\mathfrak{e}(A_0) = 0$), and thus an endomorphism $\exp^\star \mathfrak{e}$ of A is well-defined. We have $\mathfrak{e} = \log^\star(\mathrm{id}_A)$, so that $\exp^\star \mathfrak{e} = \exp^\star(\log^\star(\mathrm{id}_A)) = \mathrm{id}_A$ (since $\exp^\star(\log^\star(f + u\epsilon)) = f + u\epsilon$ for every $\mathbf{k}$-linear map $f : A \to A$ which annihilates A_0).

Let us recall that any f annihilating A_0 has the property that for each n one has that A_n is annihilated by $f^{\star m}$ for every $m > n$ (we saw this in the proof of Proposition 1.4.24). Applying this to $f = \mathfrak{e}$ (which, as we know, annihilates A_0), and renaming m and n as n and N, we obtain

$$\mathfrak{e}^{\star n}(A_N) = 0 \qquad \text{for every } N \in \mathbb{N} \text{ and every } n \in \mathbb{N} \text{ satisfying } n > N. \tag{13.129.5}$$

Recall that $\exp^\star(nf) = (\exp^\star f)^{\star n}$ for every $n \in \mathbb{N}$ and any $\mathbf{k}$-linear map $f : A \to A$ which annihilates A_0. Applying this to $f = \mathfrak{e}$, we see that every $n \in \mathbb{N}$ satisfies

$$\exp^\star(n\mathfrak{e}) = \Bigg(\underbrace{\exp^\star \mathfrak{e}}_{=\mathrm{id}_A}\Bigg)^{\star n} = \mathrm{id}_A^{\star n}. \tag{13.129.6}$$

Now, let us fix $M \in \mathbb{N}$. Let also N be any integer satisfying $N \geq M$. It is easy to see that

$$\mathrm{id}_A^{\star m}(v) = \sum_{\ell=0}^{N} \frac{1}{\ell!} m^\ell \mathfrak{e}^{\star\ell}(v) \qquad \text{for every } v \in A_M \text{ and every } m \in \mathbb{N}. \tag{13.129.7}$$

[925]

Now, let $v \in A_M$, $n \in \mathbb{N}$ and $m \in \mathbb{N}$ be arbitrary. Since $\mathrm{id}_A^{\star m}$ is a graded map (by (13.129.4), applied to $\ell = m$), we have $\mathrm{id}_A^{\star m}(v) \in A_M$ (since $v \in A_M$). Hence, (13.129.7) (applied to n and $\mathrm{id}_A^{\star m}(v)$ instead of m

[924] Indeed, this follows from Proposition 1.7.18(e) (applied to $C = A$), after first observing that $f \in \mathfrak{n}(A, A)$ (by Proposition 1.7.11(h), applied to $C = A$).

[925] *Proof of (13.129.7):* Let $v \in A_M$ and $m \in \mathbb{N}$. Every integer ℓ satisfying $\ell > M$ satisfies $\mathfrak{e}^{\star\ell}\left(\underbrace{v}_{\in A_M}\right) \in \mathfrak{e}^{\star\ell}(A_M) = 0$ (by (13.129.5), applied to ℓ and M instead of n and N). In other words,

$$\text{every integer } \ell \text{ satisfying } \ell > M \text{ satisfies } \mathfrak{e}^{\star\ell}(v) = 0. \tag{13.129.8}$$

From (13.129.6) (applied to m instead of n), we obtain $\exp^\star(m\mathfrak{e}) = \mathrm{id}_A^{\star m}$, so that

$$\mathrm{id}_A^{\star m} = \exp^\star(m\mathfrak{e}) = \sum_{\ell=0}^{\infty} \frac{1}{\ell!}(m\mathfrak{e})^{\star\ell} = \sum_{\ell=0}^{\infty} \frac{1}{\ell!} m^\ell \mathfrak{e}^{\star\ell}$$

and thus

$$\begin{aligned}\mathrm{id}_A^{\star m}(v) &= \sum_{\ell=0}^{\infty} \frac{1}{\ell!} m^\ell \mathfrak{e}^{\star\ell}(v) = \sum_{\ell=0}^{N} \frac{1}{\ell!} m^\ell \mathfrak{e}^{\star\ell}(v) + \sum_{\ell=N+1}^{\infty} \frac{1}{\ell!} m^\ell \underbrace{\mathfrak{e}^{\star\ell}(v)}_{\substack{=0 \\ (\text{by (13.129.8)}, \\ \text{since } \ell > N \geq M)}} \\ &= \sum_{\ell=0}^{N} \frac{1}{\ell!} m^\ell \mathfrak{e}^{\star\ell}(v) + \underbrace{\sum_{\ell=N+1}^{\infty} \frac{1}{\ell!} m^\ell 0}_{=0} = \sum_{\ell=0}^{N} \frac{1}{\ell!} m^\ell \mathfrak{e}^{\star\ell}(v).\end{aligned}$$

This proves (13.129.7).

and v) yields

$$\begin{aligned}
\operatorname{id}_A^{\star n}\left(\operatorname{id}_A^{\star m}(v)\right) &= \sum_{\ell=0}^{N} \frac{1}{\ell!} n^{\ell} \mathfrak{e}^{\star \ell}\left(\operatorname{id}_A^{\star m}(v)\right) = \sum_{k=0}^{N} \frac{1}{k!} n^{k} \mathfrak{e}^{\star k}\left(\underbrace{\operatorname{id}_A^{\star m}(v)}_{\substack{=\sum_{\ell=0}^{N} \frac{1}{\ell!} m^{\ell} \mathfrak{e}^{\star \ell}(v) \\ \text{(by (13.129.7))}}}\right) \\
&\qquad \left(\text{here, we have renamed the summation index } \ell \text{ as } k\right) \\
&= \sum_{k=0}^{N} \frac{1}{k!} n^{k} \underbrace{\mathfrak{e}^{\star k}\left(\sum_{\ell=0}^{N} \frac{1}{\ell!} m^{\ell} \mathfrak{e}^{\star \ell}(v)\right)}_{\substack{=\sum_{\ell=0}^{N} \frac{1}{\ell!} m^{\ell} \mathfrak{e}^{\star k}\left(\mathfrak{e}^{\star \ell}(v)\right) \\ \text{(since } \mathfrak{e}^{\star k} \text{ is a } \mathbf{k}\text{-linear map)}}} = \sum_{k=0}^{N} \frac{1}{k!} n^{k} \sum_{\ell=0}^{N} \frac{1}{\ell!} m^{\ell} \underbrace{\mathfrak{e}^{\star k}\left(\mathfrak{e}^{\star \ell}(v)\right)}_{=\left(\mathfrak{e}^{\star k} \circ \mathfrak{e}^{\star \ell}\right)(v)} \\
&= \sum_{k=0}^{N} \frac{1}{k!} n^{k} \sum_{\ell=0}^{N} \frac{1}{\ell!} m^{\ell} \left(\mathfrak{e}^{\star k} \circ \mathfrak{e}^{\star \ell}\right)(v) = \sum_{k=0}^{N} \sum_{\ell=0}^{N} \frac{\left(\mathfrak{e}^{\star k} \circ \mathfrak{e}^{\star \ell}\right)(v)}{k!\ell!} n^{k} m^{\ell}. \qquad (13.129.9)
\end{aligned}$$

But

$$\begin{aligned}
\operatorname{id}_A^{\star n}\left(\operatorname{id}_A^{\star m}(v)\right) &= \underbrace{\left(\operatorname{id}_A^{\star n} \circ \operatorname{id}_A^{\star m}\right)}_{\substack{=\operatorname{id}_A^{\star (nm)} \\ \text{(by the dual of Exercise 1.5.11(f),} \\ \text{applied to } k=m \text{ and } \ell=n)}}(v) = \operatorname{id}_A^{\star (nm)}(v) \\
&= \sum_{\ell=0}^{N} \frac{1}{\ell!} (nm)^{\ell} \mathfrak{e}^{\star \ell}(v) \qquad \left(\text{by (13.129.7), applied to } nm \text{ instead of } m\right), \qquad (13.129.10)
\end{aligned}$$

Now,

$$\begin{aligned}
&\sum_{k=0}^{N} \sum_{\ell=0}^{N} \left(\frac{\left(\mathfrak{e}^{\star k} \circ \mathfrak{e}^{\star \ell}\right)(v)}{k!\ell!} - \frac{\delta_{k,\ell} \mathfrak{e}^{\star k}(v)}{k!}\right) n^{k} m^{\ell} = \underbrace{\sum_{k=0}^{N} \sum_{\ell=0}^{N} \frac{\left(\mathfrak{e}^{\star k} \circ \mathfrak{e}^{\star \ell}\right)(v)}{k!\ell!} n^{k} m^{\ell}}_{\substack{=\operatorname{id}_A^{\star n}\left(\operatorname{id}_A^{\star m}(v)\right) \\ \text{(by (13.129.9))}}} - \underbrace{\sum_{k=0}^{N} \sum_{\ell=0}^{N} \frac{\delta_{k,\ell} \mathfrak{e}^{\star k}(v)}{k!} n^{k} m^{\ell}}_{\substack{=\sum_{\ell=0}^{N} \frac{\mathfrak{e}^{\star \ell}(v)}{\ell!} n^{\ell} m^{\ell} \\ =\sum_{\ell=0}^{N} \frac{1}{\ell!} (nm)^{\ell} \mathfrak{e}^{\star \ell}(v) \\ =\operatorname{id}_A^{\star n}\left(\operatorname{id}_A^{\star m}(v)\right) \\ \text{(by (13.129.10))}}} \\
&= \operatorname{id}_A^{\star n}\left(\operatorname{id}_A^{\star m}(v)\right) - \operatorname{id}_A^{\star n}\left(\operatorname{id}_A^{\star m}(v)\right) = 0.
\end{aligned}$$

Now, let us forget that we fixed n and m. We thus have proven that

$$\sum_{k=0}^{N} \sum_{\ell=0}^{N} \left(\frac{\left(\mathfrak{e}^{\star k} \circ \mathfrak{e}^{\star \ell}\right)(v)}{k!\ell!} - \frac{\delta_{k,\ell} \mathfrak{e}^{\star k}(v)}{k!}\right) n^{k} m^{\ell} = 0$$

for all $n \in \mathbb{N}$ and $m \in \mathbb{N}$. Hence, we can apply Lemma 13.129.2 to $V = A$ and $v_{k,\ell} = \frac{\left(\mathfrak{e}^{\star k} \circ \mathfrak{e}^{\star \ell}\right)(v)}{k!\ell!} - \frac{\delta_{k,\ell} \mathfrak{e}^{\star k}(v)}{k!}$ (because A is a torsionfree abelian group). As a result, we obtain $\frac{\left(\mathfrak{e}^{\star k} \circ \mathfrak{e}^{\star \ell}\right)(v)}{k!\ell!} - \frac{\delta_{k,\ell} \mathfrak{e}^{\star k}(v)}{k!} = 0$ for every $(k, \ell) \in \{0, 1, ..., N\}^2$. In other words,

$$\frac{\left(\mathfrak{e}^{\star k} \circ \mathfrak{e}^{\star \ell}\right)(v)}{k!\ell!} = \frac{\delta_{k,\ell} \mathfrak{e}^{\star k}(v)}{k!} \qquad \text{for every } (k, \ell) \in \{0, 1, ..., N\}^2. \qquad (13.129.11)$$

Now, forget that we fixed v, M and N. We thus have proven that every $M \in \mathbb{N}$, every integer N satisfying $N \geq M$, and every $v \in A_M$ satisfy (13.129.11).

Now, let us fix two elements $n \in \mathbb{N}$ and $m \in \mathbb{N}$. In order to finish the solution of Exercise 5.4.8(a), it remains to show that $\mathfrak{e}^{\star n} \circ \mathfrak{e}^{\star m} = n!\delta_{n,m}\mathfrak{e}^{\star n}$. In order to show this, it is clearly enough to prove that $\left(\mathfrak{e}^{\star n} \circ \mathfrak{e}^{\star m}\right)(v) = n!\delta_{n,m}\mathfrak{e}^{\star n}(v)$ for every $v \in A$. So let us fix some $v \in A$, and let us show that $\left(\mathfrak{e}^{\star n} \circ \mathfrak{e}^{\star m}\right)(v) = n!\delta_{n,m}\mathfrak{e}^{\star n}(v)$.

Since both sides of the identity $\left(\mathfrak{e}^{\star n} \circ \mathfrak{e}^{\star m}\right)(v) = n!\delta_{n,m}\mathfrak{e}^{\star n}(v)$ are $\mathbf{k}$-linear in v, we can WLOG assume that v is a homogeneous element of A (because every element of A is a $\mathbf{k}$-linear combination of homogeneous elements). Assume this. Thus, $v \in A_M$ for some $M \in \mathbb{N}$. Consider this M. Choose some $N \in \mathbb{N}$ satisfying $N \geq M$, $N \geq n$ and $N \geq m$. (Such an N clearly exists.) Then, $N \geq M$ and $(n,m) \in \{0,1,...,N\}^2$, and therefore we can apply (13.129.11) to $k = n$ and $\ell = m$. As a result, we obtain $\dfrac{\left(\mathfrak{e}^{\star n} \circ \mathfrak{e}^{\star m}\right)(v)}{n!m!} = \dfrac{\delta_{n,m}\mathfrak{e}^{\star n}(v)}{n!}$. Multiplying this identity with $n!m!$, we obtain

$$\left(\mathfrak{e}^{\star n} \circ \mathfrak{e}^{\star m}\right)(v) = \underbrace{m!\delta_{n,m}}_{\substack{=n!\delta_{n,m} \\ \text{(indeed, this is clear if } n=m\text{,} \\ \text{and otherwise follows from } \delta_{n,m}=0\text{)}}} \mathfrak{e}^{\star n}(v) = n!\delta_{n,m}\mathfrak{e}^{\star n}(v).$$

So we have proven $\left(\mathfrak{e}^{\star n} \circ \mathfrak{e}^{\star m}\right)(v) = n!\delta_{n,m}\mathfrak{e}^{\star n}(v)$. As we have seen, this completes the solution of Exercise 5.4.8(a).

(b) Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$. From (13.129.6) (applied to m instead of n), we obtain $\exp^{\star}(m\mathfrak{e}) = \operatorname{id}_A^{\star m}$, so that

$$\operatorname{id}_A^{\star m} = \exp^{\star}(m\mathfrak{e}) = \sum_{\ell=0}^{\infty} \frac{1}{\ell!}(m\mathfrak{e})^{\star \ell} = \sum_{\ell=0}^{\infty} \frac{1}{\ell!} m^{\ell} \mathfrak{e}^{\star \ell}.$$

Thus,

$$\begin{aligned} \mathfrak{e}^{\star n} \circ \underbrace{\operatorname{id}_A^{\star m}}_{=\sum_{\ell=0}^{\infty} \frac{1}{\ell!} m^{\ell} \mathfrak{e}^{\star \ell}} &= \mathfrak{e}^{\star n} \circ \left(\sum_{\ell=0}^{\infty} \frac{1}{\ell!} m^{\ell} \mathfrak{e}^{\star \ell} \right) = \underbrace{\sum_{\ell=0}^{\infty}}_{=\sum_{\ell \in \mathbb{N}}} \frac{1}{\ell!} m^{\ell} \underbrace{\mathfrak{e}^{\star n} \circ \mathfrak{e}^{\star \ell}}_{\substack{=n!\delta_{n,\ell}\mathfrak{e}^{\star n} \\ \text{(by Exercise 5.4.8(a),} \\ \text{applied to } \ell \text{ instead of } m\text{)}}} \\ &= \sum_{\ell \in \mathbb{N}} \frac{1}{\ell!} m^{\ell} n! \delta_{n,\ell} \mathfrak{e}^{\star n} = \frac{1}{n!} m^n n! \mathfrak{e}^{\star n} = m^n \mathfrak{e}^{\star n} \end{aligned}$$

and similarly $\operatorname{id}_A^{\star m} \circ \mathfrak{e}^{\star n} = m^n \mathfrak{e}^{\star n}$. This solves Exercise 5.4.8(b).

13.130. **Solution to Exercise 5.4.12.** *Solution to Exercise 5.4.12.* (a) This is proven by induction over i. The induction step relies on the observation that $R_{(1^i,n-i)} + R_{(1^{i-1},n-i+1)} = R_{(1^i)}H_{n-i}$ (where $R_{(1^{-1},n+1)}$ is to be understood to mean 0 in the $i = 0$ case). Let us prove this observation. We assume WLOG that $i > 0$ (the proof in the $i = 0$ case is analogous but simpler). The equality (5.4.9) yields $H_{n-i} = R_{(n-i)}$ (since only $(n-i)$ coarsens $(n-i)$), so that

$$\begin{aligned} R_{(1^i)}H_{n-i} &= R_{(1^i)}R_{(n-i)} = R_{(1^i)\cdot(n-i)} + R_{(1^i)\odot(n-i)} \qquad \text{(by (5.4.11))} \\ &= R_{(1^i,n-i)} + R_{(1^{i-1},n-i+1)}, \end{aligned}$$

qed.

(b) Part (a) yields

$$\begin{aligned} (-1)^i R_{(1^i,n-i)} &= (-1)^i \sum_{j=0}^{i} (-1)^{i-j} R_{(1^j)} H_{n-j} = \sum_{j=0}^{i} \underbrace{(-1)^j R_{(1^j)}}_{\substack{=S\left(R_{(j)}\right) \\ \text{(by (5.4.12), since } \omega((j))=\left(1^j\right)\text{)}}} H_{n-j} \\ &= \sum_{j=0}^{i} S\left(\underbrace{R_{(j)}}_{\substack{=H_j \\ \text{(by (5.4.9))}}} \right) H_{n-j} = \sum_{j=0}^{i} S(H_j) H_{n-j}. \end{aligned}$$

This proves (b).

(c) We have

$$\begin{aligned}
\Psi_n &= \sum_{i=0}^{n-1} \underbrace{(-1)^i R_{(1^i,n-i)}}_{\substack{=\sum_{j=0}^{i} S(H_j) H_{n-j} \\ \text{(by part (b))}}} = \sum_{i=0}^{n-1} \sum_{j=0}^{i} S(H_j) H_{n-j} = \sum_{j=0}^{n} \sum_{i=j}^{n-1} S(H_j) H_{n-j} \\
&= \sum_{j=0}^{n} (n-j) S(H_j) H_{n-j} = \sum_{j=0}^{n} S(H_j) \underbrace{(n-j) H_{n-j}}_{=\deg(H_{n-j}) H_{n-j} = E(H_{n-j})} \\
&= \sum_{j=0}^{n} S(H_j) E(H_{n-j}) = \sum_{(H_n)} S((H_n)_1) E((H_n)_2) \\
&\qquad \left(\text{using Sweedler notation, since } \Delta(H_n) = \sum_{j=0}^{n} H_j \otimes H_{n-j} \right) \\
&= (S \star E)(H_n),
\end{aligned}$$

and thus Ψ_n is primitive (by Exercise 1.5.14 (a)). (That said, there are other ways to prove the primitivity of Ψ_n.)

(d) Let $n \in \mathbb{N}$. Then,

$$\begin{aligned}
\sum_{k=0}^{n-1} H_k \underbrace{\Psi_{n-k}}_{\substack{=(S\star E)(H_{n-k}) \\ \text{(by part (c))}}} &= \sum_{k=0}^{n-1} H_k (S \star E)(H_{n-k}) = \sum_{k=0}^{n} H_k (S \star E)(H_{n-k}) \\
&\qquad \text{(we added a } k = n \text{ term, which does not matter since it vanishes)} \\
&= \sum_{(H_n)} (H_n)_1 (S \star E)((H_n)_2) \\
&\qquad \left(\text{using Sweedler notation, since } \Delta(H_n) = \sum_{k=0}^{n} H_k \otimes H_{n-k} \right) \\
&= \left(\underbrace{\operatorname{id} \star S}_{=u\epsilon} \star E \right)(H_n) = \underbrace{(u\epsilon \star E)}_{=E}(H_n) = E(H_n) = nH_n,
\end{aligned}$$

qed.

(e) Comparing coefficients reduces this to part (d).

(f) The ring homomorphism $\pi : \operatorname{NSym} \to \Lambda$ induces a ring homomorphism $\operatorname{NSym}[[t]] \to \Lambda[[t]]$. Applying this latter homomorphism to the equality $\dfrac{d}{dt}\widetilde{H}(t) = \widetilde{H}(t) \cdot \psi(t)$ of part (e), we obtain $\dfrac{d}{dt} H(t) = H(t) \cdot \overline{\psi}(t)$, where $H(t)$ is defined as in (2.4.1), whereas $\overline{\psi}(t) \in \Lambda[[t]]$ is defined by

$$\overline{\psi}(t) = \sum_{n \geq 1} \pi(\Psi_n) t^{n-1}.$$

Hence,

$$\begin{aligned}
\overline{\psi}(t) &= \frac{\frac{d}{dt} H(t)}{H(t)} = \frac{H'(t)}{H(t)} = \sum_{m \geq 0} p_{m+1} t^m \qquad \text{(by (2.5.13))} \\
&= \sum_{n \geq 1} p_n t^{n-1}.
\end{aligned}$$

Comparing coefficients in this equality yields $\pi(\Psi_n) = p_n$ for every positive integer n. This solves part (f).

Remark: Another way to solve Exercise 5.4.12(f) proceeds as follows: We know that Ψ_n is a primitive homogeneous element of NSym of degree n (by Exercise 5.4.12(c)). Thus, $\pi(\Psi_n)$ is a primitive homogeneous element of Λ of degree n (since π is a graded homomorphism of Hopf algebras). But Exercise 3.1.9 shows that all such elements are scalar multiples of p_n. Thus, $\pi(\Psi_n)$ is a scalar multiple of p_n. Finding the scalar is easy (e.g., it can be obtained by specializing at (1)).

(g) Let n be a positive integer. We know that $\Psi_n = \sum_{i=0}^{n-1} (-1)^i R_{(1^i, n-i)}$. Applying π to this equation, we obtain $\pi(\Psi_n) = \sum_{i=0}^{n-1} (-1)^i \pi\left(R_{(1^i,n-i)}\right)$. Since $\pi(\Psi_n) = p_n$ (by part (f)), this rewrites as $p_n = \sum_{i=0}^{n-1} (-1)^i \pi\left(R_{(1^i,n-i)}\right)$. But we need to show that $p_n = \sum_{i=0}^{n-1} (-1)^i s_{(n-i,1^i)}$. Hence, it is enough to prove that every $i \in \{0, 1, \ldots, n-1\}$ satisfies $\pi\left(R_{(1^i,n-i)}\right) = s_{(n-i,1^i)}$.

So let $i \in \{0, 1, \ldots, n-1\}$ be arbitrary. Theorem 5.4.10(b) (applied to $\alpha = \left(1^i, n-i\right)$) shows that $\pi\left(R_{(1^i,n-i)}\right) = s_\alpha$, where α is the ribbon diagram of the composition $\left(1^i, n-i\right)$. But since the ribbon diagram of the composition $\left(1^i, n-i\right)$ is the Ferrers diagram for the partition $\left(n-i, 1^i\right)$ (because its row lengths are $\underbrace{1, 1, \ldots, 1}_{i \text{ times}}, n-i$ going from bottom to top, with an overlap of 1 between every two adjacent rows), we have $s_\alpha = s_{(n-i,1^i)}$. Hence, $\pi\left(R_{(1^i,n-i)}\right) = s_\alpha = s_{(n-i,1^i)}$. As we have seen, this completes the solution of part (g).

(h) We will prove (5.4.13) by strong induction over n. So let N be an arbitrary positive integer, and let us assume that (5.4.13) has been proven for every $n < N$. We now need to prove (5.4.13) for $n = N$.

We notice that the family $(H_\alpha)_{\alpha \in \mathrm{NSym}}$ is multiplicative, in the sense that any two compositions β and γ satisfy $H_{\beta \cdot \gamma} = H_\beta \cdot H_\gamma$ (where, as we recall, $\beta \cdot \gamma$ denotes the concatenation of the compositions β and γ). This follows from the definition of the H_α.

We have $N > 0$. Hence, every $\alpha \in \mathrm{Comp}_N$ can be written uniquely in the form $\alpha = (q) \cdot \beta$ for some $q \in \{1, 2, ..., N\}$ and some $\beta \in \mathrm{Comp}_{N-q}$ (indeed, the q is just the first entry of the composition α, and β is the composition obtained by erasing this first entry). Hence,

$$\begin{aligned}
&\sum_{\alpha \in \mathrm{Comp}_N} (-1)^{\ell(\alpha)-1} \operatorname{lp}(\alpha) H_\alpha \\
&= \underbrace{\sum_{\substack{q \in \{1,2,\ldots,N\};\\ \beta \in \mathrm{Comp}_{N-q}}}}_{=\sum_{q=1}^{N} \sum_{\beta \in \mathrm{Comp}_{N-q}}} \underbrace{(-1)^{\ell((q)\cdot\beta)-1}}_{\substack{=(-1)^{\ell(\beta)}\\ \text{(since it is easy to see that}\\ \ell((q)\cdot\beta)-1=\ell(\beta))}} \operatorname{lp}((q)\cdot\beta) \underbrace{H_{(q)\cdot\beta}}_{\substack{=H_{(q)}\cdot H_\beta\\ \text{(since the family}\\ (H_\alpha)_{\alpha\in\mathrm{NSym}} \text{ is}\\ \text{multiplicative)}}} \\
&= \sum_{q=1}^{N} \sum_{\beta \in \mathrm{Comp}_{N-q}} (-1)^{\ell(\beta)} \operatorname{lp}((q)\cdot\beta) \cdot H_{(q)} \cdot H_\beta \\
&= \sum_{q=1}^{N-1} \sum_{\beta \in \mathrm{Comp}_{N-q}} (-1)^{\ell(\beta)} \underbrace{\operatorname{lp}((q)\cdot\beta)}_{\substack{=\operatorname{lp}(\beta)\\ \text{(since } \beta \text{ is}\\ \text{nonempty)}}} \cdot H_{(q)} \cdot H_\beta + \underbrace{\sum_{\beta \in \mathrm{Comp}_{N-N}} (-1)^{\ell(\beta)} \operatorname{lp}((N)\cdot\beta) \cdot H_{(N)} \cdot H_\beta}_{\substack{=(-1)^{\ell(\varnothing)} \operatorname{lp}((N)\cdot\varnothing)\cdot H_{(N)}\cdot H_\varnothing\\ \text{(since the only element of } \mathrm{Comp}_{N-N}\\ \text{is the empty composition } \varnothing)}} \\
&= \sum_{q=1}^{N-1} \underbrace{\sum_{\beta \in \mathrm{Comp}_{N-q}} (-1)^{\ell(\beta)} \operatorname{lp}(\beta) \cdot H_{(q)} \cdot H_\beta}_{=H_{(q)}\cdot\left(\sum_{\beta\in\mathrm{Comp}_{N-q}} (-1)^{\ell(\beta)} \operatorname{lp}(\beta) H_\beta\right)} + \underbrace{(-1)^{\ell(\varnothing)}}_{=1} \underbrace{\operatorname{lp}((N)\cdot\varnothing)}_{=\operatorname{lp}((N))=N} \cdot \underbrace{H_{(N)}}_{=H_N} \cdot \underbrace{H_\varnothing}_{=1} \\
&= \sum_{q=1}^{N-1} H_{(q)} \cdot \left(\sum_{\beta \in \mathrm{Comp}_{N-q}} (-1)^{\ell(\beta)} \operatorname{lp}(\beta) H_\beta \right) + N H_N. \qquad (13.130.1)
\end{aligned}$$

However, for every $q \in \{1, 2, ..., N-1\}$, we can apply (5.4.13) to $n = N - q$ (because $N - q < N$, and because (5.4.13) has been proven for every $n < N$). Thus, for every $q \in \{1, 2, ..., N-1\}$, we obtain

$$\Psi_{N-q} = \sum_{\alpha \in \operatorname{Comp}_{N-q}} \underbrace{(-1)^{\ell(\alpha)-1}}_{=-(-1)^{\ell(\alpha)}} \operatorname{lp}(\alpha) H_\alpha = - \sum_{\alpha \in \operatorname{Comp}_{N-q}} (-1)^{\ell(\alpha)} \operatorname{lp}(\alpha) H_\alpha = - \sum_{\beta \in \operatorname{Comp}_{N-q}} (-1)^{\ell(\beta)} \operatorname{lp}(\beta) H_\beta$$

(here, we renamed the summation index α as β), so that

$$-\Psi_{N-q} = \sum_{\beta \in \operatorname{Comp}_{N-q}} (-1)^{\ell(\beta)} \operatorname{lp}(\beta) H_\beta. \tag{13.130.2}$$

Thus, (13.130.1) becomes

$$\begin{aligned} \sum_{\alpha \in \operatorname{Comp}_N} (-1)^{\ell(\alpha)-1} \operatorname{lp}(\alpha) H_\alpha &= \sum_{q=1}^{N-1} \underbrace{H_{(q)}}_{=H_q} \cdot \underbrace{\left(\sum_{\beta \in \operatorname{Comp}_{N-q}} (-1)^{\ell(\beta)} \operatorname{lp}(\beta) H_\beta \right)}_{\substack{=-\Psi_{N-q} \\ \text{(this follows from (13.130.2))}}} + NH_N \\ &= \sum_{q=1}^{N-1} H_q \cdot (-\Psi_{N-q}) + NH_N = - \sum_{q=1}^{N-1} H_q \Psi_{N-q} + NH_N. \qquad (13.130.3) \end{aligned}$$

However, Exercise 5.4.12(d) (applied to $n = N$) yields $\sum_{k=0}^{N-1} H_k \Psi_{N-k} = NH_N$, so that

$$NH_N = \sum_{k=0}^{N-1} H_k \Psi_{N-k} = \sum_{q=0}^{N-1} H_q \Psi_{N-q} = \underbrace{H_0}_{=1} \underbrace{\Psi_{N-0}}_{=\Psi_N} + \sum_{q=1}^{N-1} H_q \Psi_{N-q} = \Psi_N + \sum_{q=1}^{N-1} H_q \Psi_{N-q}.$$

Now, (13.130.3) becomes

$$\begin{aligned} &\sum_{\alpha \in \operatorname{Comp}_N} (-1)^{\ell(\alpha)-1} \operatorname{lp}(\alpha) H_\alpha \\ &= - \sum_{q=1}^{N-1} H_q \Psi_{N-q} + \underbrace{NH_N}_{=\Psi_N + \sum_{q=1}^{N-1} H_q \Psi_{N-q}} = - \sum_{q=1}^{N-1} H_q \Psi_{N-q} + \Psi_N + \sum_{q=1}^{N-1} H_q \Psi_{N-q} = \Psi_N. \end{aligned}$$

In other words, $\Psi_N = \sum_{\alpha \in \operatorname{Comp}_N} (-1)^{\ell(\alpha)-1} \operatorname{lp}(\alpha) H_\alpha$. Thus, (5.4.13) is proven for $n = N$. This completes the induction step, and therefore the proof of (5.4.13) is complete. That is, part (h) of the exercise is solved.

(i) First of all, for every positive integer n, the element Ψ_n of NSym is homogeneous of degree n (this follows from the definition of Ψ_n or, alternatively, from Exercise 5.4.12(h)). Hence, for every composition α, the element Ψ_α of NSym is homogeneous of degree $|\alpha|$.

We notice that the family $(\Psi_\alpha)_{\alpha \in \mathrm{NSym}}$ is multiplicative, in the sense that any two compositions β and γ satisfy $\Psi_{\beta \cdot \gamma} = \Psi_\beta \cdot \Psi_\gamma$ (where, as we recall, $\beta \cdot \gamma$ denotes the concatenation of the compositions β and γ). This follows from the definition of the Ψ_α.

Let us now prove (5.4.14). Indeed, we will show (5.4.14) by strong induction over n. So let $N \in \mathbb{N}$ be arbitrary, and let us assume that (5.4.14) has been proven for every $n < N$. We now need to prove (5.4.14) for $n = N$.

If $N = 0$, then (5.4.14) obviously holds for $n = N$ (because both sides of (5.4.14) are 1_{NSym} in this case). We thus WLOG assume that $N \neq 0$. Thus, every $\alpha \in \operatorname{Comp}_N$ can be written uniquely in the form $\alpha = \beta \cdot (q)$ for some $q \in \{1, 2, ..., N\}$ and some $\beta \in \operatorname{Comp}_{N-q}$ (indeed, the q is just the last entry of the composition α, and β is the composition obtained by erasing this last entry). Hence,

$$\begin{aligned}
\sum_{\alpha\in\operatorname{Comp}_N} \frac{1}{\pi_u(\alpha)}\Psi_\alpha &= \underbrace{\sum_{\substack{q\in\{1,2,...,N\};\\ \beta\in\operatorname{Comp}_{N-q}}}}_{=\sum_{q=1}^{N}\sum_{\beta\in\operatorname{Comp}_{N-q}}} \underbrace{\frac{1}{\pi_u(\beta\cdot(q))}}_{\substack{=\frac{1}{\pi_u(\beta)\cdot N}\\ \text{(since it is easy to see that}\\ \pi_u(\beta\cdot(q))=\pi_u(\beta)\cdot N)}} \underbrace{\Psi_{\beta\cdot(q)}}_{\substack{=\Psi_\beta\cdot\Psi_{(q)}\\ \text{(since the family}\\ (\Psi_\alpha)_{\alpha\in\operatorname{NSym}}\text{ is}\\ \text{multiplicative)}}} \\
&= \sum_{q=1}^{N}\sum_{\beta\in\operatorname{Comp}_{N-q}} \underbrace{\frac{1}{\pi_u(\beta)\cdot N}}_{=\frac{1}{N}\cdot\frac{1}{\pi_u(\beta)}} \Psi_\beta\cdot\underbrace{\Psi_{(q)}}_{=\Psi_q} \\
&= \sum_{q=1}^{N}\sum_{\beta\in\operatorname{Comp}_{N-q}} \frac{1}{N}\cdot\frac{1}{\pi_u(\beta)}\Psi_\beta\cdot\Psi_q
\end{aligned}$$

$$= \frac{1}{N}\sum_{q=1}^{N}\left(\sum_{\beta\in\operatorname{Comp}_{N-q}}\frac{1}{\pi_u(\beta)}\Psi_\beta\right)\cdot\Psi_q. \tag{13.130.4}$$

However, for every $q\in\{1,2,...,N\}$, we can apply (5.4.14) to $n=N-q$ (because $N-q<N$, and because (5.4.14) has been proven for every $n<N$). Thus, for every $q\in\{1,2,...,N\}$, we obtain

$$H_{N-q} = \sum_{\alpha\in\operatorname{Comp}_{N-q}}\frac{1}{\pi_u(\alpha)}\Psi_\alpha = \sum_{\beta\in\operatorname{Comp}_{N-q}}\frac{1}{\pi_u(\beta)}\Psi_\beta \tag{13.130.5}$$

(here, we renamed the summation index α as β). Thus, (13.130.4) becomes

$$\begin{aligned}
\sum_{\alpha\in\operatorname{Comp}_N}\frac{1}{\pi_u(\alpha)}\Psi_\alpha &= \frac{1}{N}\sum_{q=1}^{N}\underbrace{\left(\sum_{\beta\in\operatorname{Comp}_{N-q}}\frac{1}{\pi_u(\beta)}\Psi_\beta\right)}_{\substack{=H_{N-q}\\ \text{(by (13.130.5))}}}\cdot\Psi_q \\
&= \frac{1}{N}\sum_{q=1}^{N}H_{N-q}\Psi_q = \frac{1}{N}\underbrace{\sum_{k=0}^{N-1}H_k\Psi_{N-k}}_{\substack{=NH_N\\ \text{(by Exercise 5.4.12(d),}\\ \text{applied to }n=N)}} \\
&\qquad \text{(here, we substituted } N-k \text{ for } q \text{ in the sum)} \\
&= \frac{1}{N}NH_N = H_N.
\end{aligned}$$

In other words, $H_N = \sum_{\alpha\in\operatorname{Comp}_N}\frac{1}{\pi_u(\alpha)}\Psi_\alpha$. Thus, (5.4.14) is proven for $n=N$. This completes the induction step, and therefore the proof of (5.4.14) is complete.

We now need to prove that $(\Psi_\alpha)_{\alpha\in\operatorname{Comp}_n}$ is a $\mathbf{k}$-basis of NSym_n for every $n\in\mathbb{N}$. It is clear that $(\Psi_\alpha)_{\alpha\in\operatorname{Comp}_n}$ is a family of elements of NSym_n (because for every composition α, the element Ψ_α of NSym is homogeneous of degree $|\alpha|$).

Let $\mathfrak{A}$ be the $\mathbf{k}$-subalgebra of NSym generated by the elements Ψ_1, Ψ_2, Ψ_3, ... Then, $\mathfrak{A}$ contains Ψ_α for every composition α (by the definition of Ψ_α). Therefore, $\mathfrak{A}$ contains H_n for every $n\geq 1$ (by (5.4.14)). Consequently, $\mathfrak{A}=\operatorname{NSym}$ (because NSym is generated as a $\mathbf{k}$-algebra by H_1, H_2, H_3, ...). In other words, the $\mathbf{k}$-algebra NSym is generated by the elements Ψ_1, Ψ_2, Ψ_3, ... (since we defined $\mathfrak{A}$ as the $\mathbf{k}$-subalgebra of NSym generated by the elements Ψ_1, Ψ_2, Ψ_3, ...). In other words, the $\mathbf{k}$-module NSym is spanned by all possible products of the elements Ψ_1, Ψ_2, Ψ_3, ... In other words, the $\mathbf{k}$-module NSym is spanned by the elements Ψ_α with $\alpha\in\operatorname{Comp}$ (because the elements Ψ_α with $\alpha\in\operatorname{Comp}$ are precisely all possible products of the elements Ψ_1, Ψ_2, Ψ_3, ...). In yet other words, the family $(\Psi_\alpha)_{\alpha\in\operatorname{Comp}}$ spans the $\mathbf{k}$-module NSym.

Now, fix $n \in \mathbb{N}$. Every element of NSym_n can be written as a $\mathbf{k}$-linear combination of the elements Ψ_α with $\alpha \in \mathrm{Comp}$ (since the family $(\Psi_\alpha)_{\alpha\in\mathrm{Comp}}$ spans the $\mathbf{k}$-module NSym). In this $\mathbf{k}$-linear combination, we can remove all terms Ψ_α with $\alpha \notin \mathrm{Comp}_n$ without changing its value (by gradedness, because Ψ_α is homogeneous of degree $|\alpha|$), and so we conclude that every element of NSym_n can be written as a $\mathbf{k}$-linear combination of the elements Ψ_α with $\alpha \in \mathrm{Comp}_n$. In other words, the family $(\Psi_\alpha)_{\alpha\in\mathrm{Comp}_n}$ spans the $\mathbf{k}$-module NSym_n.

Now, we can apply Exercise 2.5.18(b) to $A = \mathrm{NSym}_n$, $I = \mathrm{Comp}_n$, $(\gamma_i)_{i\in I} = (H_\alpha)_{\alpha\in\mathrm{Comp}_n}$ and $(\beta_i)_{i\in I} = (\Psi_\alpha)_{\alpha\in\mathrm{Comp}_n}$ (since we know that $(H_\alpha)_{\alpha\in\mathrm{Comp}_n}$ is a $\mathbf{k}$-basis of NSym_n, whereas $(\Psi_\alpha)_{\alpha\in\mathrm{Comp}_n}$ spans the $\mathbf{k}$-module NSym_n). We conclude that $(\Psi_\alpha)_{\alpha\in\mathrm{Comp}_n}$ is a $\mathbf{k}$-basis of NSym_n. This completes the solution of part (i).

(j) For every composition α, define an element $\mathfrak{b}_\alpha$ of $T(V)$ by $\mathfrak{b}_\alpha = \mathfrak{b}_{\alpha_1}\mathfrak{b}_{\alpha_2}\cdots\mathfrak{b}_{\alpha_\ell}$, where α is written in the form $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_\ell)$ with $\ell = \ell(\alpha)$. Then, $(\mathfrak{b}_\alpha)_{\alpha\in\mathrm{Comp}}$ is a $\mathbf{k}$-module basis of $T(V)$ (by the basic properties of tensor algebras, since $(\mathfrak{b}_n)_{n\in\{1,2,3,\ldots\}}$ is a $\mathbf{k}$-module basis of V). Notice that the family $(\mathfrak{b}_n)_{n\in\{1,2,3,\ldots\}}$ generates the $\mathbf{k}$-algebra $T(V)$ (since it is a basis of V).

For every composition α, define an element Ψ_α of NSym as in Exercise 5.4.12(i). We know from Exercise 5.4.12(i) that $(\Psi_\alpha)_{\alpha\in\mathrm{Comp}_n}$ is a $\mathbf{k}$-basis of NSym_n for every $n \in \mathbb{N}$. Hence, $(\Psi_\alpha)_{\alpha\in\mathrm{Comp}}$ is a $\mathbf{k}$-basis of NSym.

Every $\alpha \in \mathrm{Comp}$ satisfies $F(\mathfrak{b}_\alpha) = \Psi_\alpha$ [926]. The map F thus maps the basis $(\mathfrak{b}_\alpha)_{\alpha\in\mathrm{Comp}}$ of the $\mathbf{k}$-module $T(V)$ to the basis $(\Psi_\alpha)_{\alpha\in\mathrm{Comp}}$ of the $\mathbf{k}$-module NSym. Hence, F is a $\mathbf{k}$-module isomorphism (since any $\mathbf{k}$-linear map mapping a basis to a basis is a $\mathbf{k}$-module isomorphism).

Next, we are going to show the equality $\Delta_{\mathrm{NSym}} \circ F = (F \otimes F) \circ \Delta_{T(V)}$. Indeed, this is an equality between $\mathbf{k}$-algebra homomorphisms, and thus needs only to be verified on a generating set of the $\mathbf{k}$-algebra $T(V)$. Picking $(\mathfrak{b}_n)_{n\in\{1,2,3,\ldots\}}$ as this generating set, we thus only need to check that $(\Delta_{\mathrm{NSym}} \circ F)(\mathfrak{b}_n) = \left((F \otimes F) \circ \Delta_{T(V)}\right)(\mathfrak{b}_n)$ for every positive integer n. This is straightforward: If n is any positive integer, then comparing the equalities

$$(\Delta_{\mathrm{NSym}} \circ F)(\mathfrak{b}_n) = \Delta_{\mathrm{NSym}}\left(\underbrace{F(\mathfrak{b}_n)}_{\substack{=f(\mathfrak{b}_n)\\ \text{(since } F \text{ is induced by } f\text{)}}}\right) = \Delta_{\mathrm{NSym}}\left(\underbrace{f(\mathfrak{b}_n)}_{\substack{=\Psi_n\\ \text{(by the definition of } f\text{)}}}\right) = \Delta_{\mathrm{NSym}}(\Psi_n)$$
$$= 1 \otimes \Psi_n + \Psi_n \otimes 1 \qquad \text{(since Exercise 5.4.12(c) shows that } \Psi_n \text{ is primitive)}$$

and

$$\left((F \otimes F) \circ \Delta_{T(V)}\right)(\mathfrak{b}_n) = (F \otimes F)\left(\underbrace{\Delta_{T(V)}(\mathfrak{b}_n)}_{\substack{=1\otimes\mathfrak{b}_n+\mathfrak{b}_n\otimes 1\\ \text{(by the definition of the}\\ \text{comultiplication on } T(V)\text{)}}}\right) = (F \otimes F)(1 \otimes \mathfrak{b}_n + \mathfrak{b}_n \otimes 1)$$
$$= \underbrace{F(1)}_{\substack{=1\\ \text{(since } F \text{ is a } \mathbf{k}\text{-algebra}\\ \text{homomorphism)}}} \otimes \underbrace{F(\mathfrak{b}_n)}_{\substack{=f(\mathfrak{b}_n)\\ \text{(since } F \text{ is induced by } f\text{)}}} + \underbrace{F(\mathfrak{b}_n)}_{\substack{=f(\mathfrak{b}_n)\\ \text{(since } F \text{ is induced by } f\text{)}}} \otimes \underbrace{F(1)}_{\substack{=1\\ \text{(since } F \text{ is a } \mathbf{k}\text{-algebra}\\ \text{homomorphism)}}}$$
$$= 1 \otimes \underbrace{f(\mathfrak{b}_n)}_{\substack{=\Psi_n\\ \text{(by the definition of } f\text{)}}} + \underbrace{f(\mathfrak{b}_n)}_{\substack{=\Psi_n\\ \text{(by the definition of } f\text{)}}} \otimes 1 = 1 \otimes \Psi_n + \Psi_n \otimes 1$$

[926] *Proof.* Let $\alpha \in \mathrm{Comp}$. Write α in the form $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_\ell)$ with $\ell = \ell(\alpha)$. Then, $\mathfrak{b}_\alpha = \mathfrak{b}_{\alpha_1}\mathfrak{b}_{\alpha_2}\cdots\mathfrak{b}_{\alpha_\ell}$, so that

$$F(\mathfrak{b}_\alpha) = F(\mathfrak{b}_{\alpha_1}\mathfrak{b}_{\alpha_2}\cdots\mathfrak{b}_{\alpha_\ell}) = \underbrace{f(\mathfrak{b}_{\alpha_1})}_{\substack{=\Psi_{\alpha_1}\\ \text{(by the definition of } f\text{)}}} \underbrace{f(\mathfrak{b}_{\alpha_2})}_{\substack{=\Psi_{\alpha_2}\\ \text{(by the definition of } f\text{)}}} \cdots \underbrace{f(\mathfrak{b}_{\alpha_\ell})}_{\substack{=\Psi_{\alpha_\ell}\\ \text{(by the definition of } f\text{)}}}$$
$$\text{(by the definition of } F\text{)}$$
$$= \Psi_{\alpha_1}\Psi_{\alpha_2}\cdots\Psi_{\alpha_\ell} = \Psi_\alpha \qquad \text{(since } \Psi_\alpha \text{ was defined to be } \Psi_{\alpha_1}\Psi_{\alpha_2}\cdots\Psi_{\alpha_\ell}\text{)},$$

qed.

yields $\left(\Delta_{\mathrm{NSym}} \circ F\right)(\mathfrak{b}_n) = \left((F \otimes F) \circ \Delta_{T(V)}\right)(\mathfrak{b}_n)$. We thus have shown that $\Delta_{\mathrm{NSym}} \circ F = (F \otimes F) \circ \Delta_{T(V)}$. Combining this with the equality $\epsilon_{\mathrm{NSym}} \circ F = \epsilon_{T(V)}$ (whose proof is similar but yet simpler), we conclude that F is a **k**-coalgebra homomorphism, thus a **k**-bialgebra homomorphism, hence a **k**-Hopf algebra homomorphism (by Corollary 1.4.27), therefore a **k**-Hopf algebra isomorphism (since it is a **k**-module isomorphism). This solves Exercise 5.4.12(j).

(k) Define the map F as in Exercise 5.4.12(j). Then, Exercise 5.4.12(j) shows that F is a Hopf algebra isomorphism, hence a **k**-module isomorphism.

We can endow the **k**-module V with a grading by assigning to each basis vector $\mathfrak{b}_n$ the degree n. Then, V is of finite type and satisfies $V_0 = 0$, and thus $T(V)$ is a connected graded **k**-Hopf algebra.

The element Ψ_n of NSym is homogeneous of degree n for every positive integer n. (This follows from the definition of Ψ_n.)

The map f is graded (since it sends every basis vector $\mathfrak{b}_n$ of V to the vector $\Psi_n \in \mathrm{NSym}$, which is homogeneous of the same degree n as $\mathfrak{b}_n$). Hence, the map F (being the **k**-algebra homomorphism $T(V) \to \mathrm{NSym}$ induced by f) is also graded. It is well-known that if a **k**-module isomorphism is graded, then it is an isomorphism of graded **k**-modules. Thus, F is an isomorphism of graded **k**-modules (since F is a **k**-module isomorphism and is graded), hence an isomorphism of graded Hopf algebras (since F is a Hopf algebra isomorphism). Thus, $T(V) \cong \mathrm{NSym}$ as graded Hopf algebras. Therefore, $T(V)^o \cong \mathrm{NSym}^o \cong \mathrm{QSym}$ (since $\mathrm{NSym} = \mathrm{QSym}^o$) as graded Hopf algebras. Hence, $\mathrm{QSym} \cong T(V)^o$ as graded Hopf algebras.

Remark 1.6.9(b) shows that the Hopf algebra $T(V)^o$ is naturally isomorphic to the shuffle algebra $\mathrm{Sh}(V^o)$ as Hopf algebras. But $V^o \cong V$ as **k**-modules (since V is of finite type), and thus $\mathrm{Sh}(V^o) \cong \mathrm{Sh}(V)$ as Hopf algebras. Altogether, we obtain $\mathrm{QSym} \cong T(V)^o \cong \mathrm{Sh}(V^o) \cong \mathrm{Sh}(V)$ as Hopf algebras. This solves Exercise 5.4.12(k).

(l) We recall that

$$\pi\left(R_{(1^i,n-i)}\right) = s_{(n-i,1^i)} \qquad \text{for every positive integer } n \text{ and every } i \in \{0,1,...,n-1\}. \tag{13.130.6}$$

(This has been proven during the solution to Exercise 5.4.12(g).) Also, Theorem 5.4.10(b) (applied to $\alpha = (1^i)$) yields that

$$\pi\left(R_{(1^i)}\right) = s_{(1^i)} = e_i \qquad \text{for every } i \in \mathbb{N}. \tag{13.130.7}$$

In our solution to Exercise 5.4.12(a), we have shown that

$$R_{(1^i)}H_{n-i} = R_{(1^i,n-i)} + R_{(1^{i-1},n-i+1)} \qquad \text{for every positive integer } n \text{ and every } i \in \{1,2,...,n-1\}. \tag{13.130.8}$$

Now, we are ready to solve parts (a) and (b) of Exercise 2.9.14 anew.

Alternative solution to Exercise 2.9.14(a): Let n and m be positive integers. We need to prove that $e_n h_m = s_{(m+1,1^{n-1})} + s_{(m,1^n)}$.

Applying (13.130.8) to $n+m$ and n instead of n and i, we obtain $R_{(1^n)}H_{(n+m)-n} = R_{(1^n,(n+m)-n)} + R_{(1^{n-1},(n+m)-n+1)} = R_{(1^n,(n+m)-n)} + R_{(1^{n-1},(n+m)-(n-1))}$ (since $(n+m)-n+1 = (n+m)-(n-1)$). Applying the map π to both sides of this equality, we obtain

$$\begin{aligned}
\pi\left(R_{(1^n)}H_{(n+m)-n}\right) &= \pi\left(R_{(1^n,(n+m)-n)} + R_{(1^{n-1},(n+m)-(n-1))}\right) \\
&= \underbrace{\pi\left(R_{(1^n,(n+m)-n)}\right)}_{\substack{=s_{((n+m)-n,1^n)} \\ \text{(by (13.130.6), applied to} \\ n+m \text{ and } n \text{ instead of } n \text{ and } i)}} + \underbrace{\pi\left(R_{(1^{n-1},(n+m)-(n-1))}\right)}_{\substack{=s_{((n+m)-(n-1),1^{n-1})} \\ \text{(by (13.130.6), applied to} \\ n+m \text{ and } n-1 \text{ instead of } n \text{ and } i)}} \\
&= s_{((n+m)-n,1^n)} + s_{((n+m)-(n-1),1^{n-1})} \\
&= s_{(m,1^n)} + s_{(m+1,1^{n-1})} \qquad (\text{since } (n+m)-n = m \text{ and } (n+m)-(n-1) = m+1) \\
&= s_{(m+1,1^{n-1})} + s_{(m,1^n)}.
\end{aligned}$$

Compared with

$$\begin{aligned}\pi\left(R_{(1^n)}H_{(n+m)-n}\right) &= \underbrace{\pi\left(R_{(1^n)}\right)}_{\substack{=e_n\\ \text{(by (13.130.7), applied to}\\ n\text{ instead of }i)}} \cdot \pi\left(\underbrace{H_{(n+m)-n}}_{=H_m}\right) \qquad \text{(since }\pi\text{ is a }\mathbf{k}\text{-algebra morphism)}\\ &= e_n \cdot \underbrace{\pi\left(H_m\right)}_{\substack{=h_m\\ \text{(by the definition of }\pi)}} = e_n h_m,\end{aligned}$$

this yields $e_n h_m = s_{(m+1,1^{n-1})} + s_{(m,1^n)}$. Thus, Exercise 2.9.14(a) is solved again.

Alternative solution to Exercise 2.9.14(b): Let $a \in \mathbb{N}$ and $b \in \mathbb{N}$. Applying Exercise 5.4.12(b) to $n = a+b+1$ and $i = b$, we obtain

$$(-1)^b R_{(1^b,(a+b+1)-b)} = \sum_{j=0}^{b} S\left(H_j\right) H_{(a+b+1)-j} = \sum_{j=0}^{b} S\left(H_{b-j}\right) H_{(a+b+1)-(b-j)}$$

(here, we substituted $b-j$ for j in the sum). Multiplying both sides of this equality by $(-1)^b$, we obtain

$$\begin{aligned}R_{(1^b,(a+b+1)-b)} &= (-1)^b \sum_{j=0}^{b} S\left(H_{b-j}\right) \underbrace{H_{(a+b+1)-(b-j)}}_{\substack{=H_{a+j+1}\\ \text{(since }(a+b+1)-(b-j)=a+j+1)}} = (-1)^b \sum_{j=0}^{b} S\left(H_{b-j}\right) H_{a+j+1}\\ &= (-1)^b \sum_{i=0}^{b} S\left(H_{b-i}\right) H_{a+i+1} \qquad \text{(here, we renamed the summation index } j \text{ as } i).\end{aligned}$$

Applying the map π to both sides of this equality, we obtain

$$\begin{aligned}\pi\left(R_{(1^b,(a+b+1)-b)}\right) &= \pi\left((-1)^b \sum_{i=0}^{b} S\left(H_{b-i}\right) H_{a+i+1}\right) = (-1)^b \sum_{i=0}^{b} S\left(\underbrace{\pi\left(H_{b-i}\right)}_{\substack{=h_{b-i}\\ \text{(by the definition of }\pi)}}\right) \underbrace{\pi\left(H_{a+i+1}\right)}_{\substack{=h_{a+i+1}\\ \text{(by the definition of }\pi)}}\\ &\qquad \text{(since }\pi\text{ is a Hopf algebra homomorphism)}\\ &= (-1)^b \sum_{i=0}^{b} \underbrace{S\left(h_{b-i}\right)}_{\substack{=(-1)^{b-i}e_{b-i}\\ \text{(by Proposition 2.4.1(iii),}\\ \text{applied to }n=b-i)}} h_{a+i+1}\\ &= (-1)^b \sum_{i=0}^{b} (-1)^{b-i} e_{b-i} h_{a+i+1} = \sum_{i=0}^{b} (-1)^i h_{a+i+1} e_{b-i}.\end{aligned}$$

Thus,

$$\begin{aligned}\sum_{i=0}^{b} (-1)^i h_{a+i+1} e_{b-i} &= \pi\left(R_{(1^b,(a+b+1)-b)}\right) = s_{((a+b+1)-b,1^b)}\\ &\qquad \text{(by (13.130.6), applied to } n = a+b+1 \text{ and } i = b)\\ &= s_{(a+1,1^b)}.\end{aligned}$$

Thus, Exercise 2.9.14(b) is once again solved.

13.131. **Solution to Exercise 5.4.13.** *Solution to Exercise 5.4.13.*

(a) The solution of Exercise 5.4.13(a) can be obtained from the solution of Exercise 2.9.9(a) upon replacing $\mathbf{f}_n$ by $\mathbf{F}_n$.

(b) The solution of Exercise 5.4.13(b) can be obtained from the solution of Exercise 2.9.9(b) upon replacing $\mathbf{f}_n$, $\mathbf{f}_m$, $\mathbf{f}_{nm}$ and Λ by $\mathbf{F}_n$, $\mathbf{F}_m$, $\mathbf{F}_{nm}$ and QSym.

(c) The solution of Exercise 5.4.13(c) can be obtained from the solution of Exercise 2.9.9(c) upon replacing $\mathbf{f}_1$ and Λ by $\mathbf{F}_1$ and QSym.

(d) Let $n \in \{1, 2, 3, \ldots\}$ and $(\beta_1, \beta_2, \ldots, \beta_s) \in \operatorname{Comp}$. The definition of $M_{(\beta_1,\beta_2,\ldots,\beta_s)}$ yields that $M_{(\beta_1,\beta_2,\ldots,\beta_s)} = \sum_{i_1<i_2<\cdots<i_s} x_{i_1}^{\beta_1} x_{i_2}^{\beta_2} \cdots x_{i_s}^{\beta_s}$ (where the sum is over all s-tuples $(i_1, i_2, \ldots, i_s)$ of positive integers satisfying $i_1 < i_2 < \cdots < i_s$). Applying the map $\mathbf{F}_n$ to both sides of this equality, we obtain

$$\begin{aligned}
\mathbf{F}_n\left(M_{(\beta_1,\beta_2,\ldots,\beta_s)}\right) &= \mathbf{F}_n\left(\sum_{i_1<i_2<\cdots<i_s} x_{i_1}^{\beta_1} x_{i_2}^{\beta_2} \cdots x_{i_s}^{\beta_s}\right) \\
&= \left(\sum_{i_1<i_2<\cdots<i_s} x_{i_1}^{\beta_1} x_{i_2}^{\beta_2} \cdots x_{i_s}^{\beta_s}\right)(x_1^n, x_2^n, x_3^n, \ldots) \qquad \text{(by the definition of } \mathbf{F}_n\text{)} \\
&= \sum_{i_1<i_2<\cdots<i_s} \left(x_{i_1}^n\right)^{\beta_1} \left(x_{i_2}^n\right)^{\beta_2} \cdots \left(x_{i_s}^n\right)^{\beta_s} = \sum_{i_1<i_2<\cdots<i_s} x_{i_1}^{n\beta_1} x_{i_2}^{n\beta_2} \cdots x_{i_s}^{n\beta_s}.
\end{aligned}$$

Compared with

$$M_{(n\beta_1,n\beta_2,\ldots,n\beta_s)} = \sum_{i_1<i_2<\cdots<i_s} x_{i_1}^{n\beta_1} x_{i_2}^{n\beta_2} \cdots x_{i_s}^{n\beta_s} \qquad \left(\text{by the definition of } M_{(n\beta_1,n\beta_2,\ldots,n\beta_s)}\right),$$

this yields $\mathbf{F}_n\left(M_{(\beta_1,\beta_2,\ldots,\beta_s)}\right) = M_{(n\beta_1,n\beta_2,\ldots,n\beta_s)}$. This solves Exercise 5.4.13(d).

(e) Fix $n \in \{1, 2, 3, \ldots\}$. We now know that $\mathbf{F}_n$ is a $\mathbf{k}$-algebra homomorphism (due to Exercise 5.4.13(a)), thus a $\mathbf{k}$-linear map.

Let $\alpha \in \operatorname{Comp}$. Write α in the form $(\alpha_1, \alpha_2, \ldots, \alpha_\ell)$. Then,

$$\Delta M_\alpha = \sum_{k=0}^{\ell} M_{(\alpha_1,\alpha_2,\ldots,\alpha_k)} \otimes M_{(\alpha_{k+1},\alpha_{k+2},\ldots,\alpha_\ell)} \tag{13.131.1}$$

(according to Proposition 5.1.7). Applying the map $\mathbf{F}_n \otimes \mathbf{F}_n$ to both sides of this equality, we obtain

$$\begin{aligned}
(\mathbf{F}_n \otimes \mathbf{F}_n)(\Delta M_\alpha) &= (\mathbf{F}_n \otimes \mathbf{F}_n)\left(\sum_{k=0}^{\ell} M_{(\alpha_1,\alpha_2,\ldots,\alpha_k)} \otimes M_{(\alpha_{k+1},\alpha_{k+2},\ldots,\alpha_\ell)}\right) \\
&= \sum_{k=0}^{\ell} \underbrace{\mathbf{F}_n\left(M_{(\alpha_1,\alpha_2,\ldots,\alpha_k)}\right)}_{\substack{=M_{(n\alpha_1,n\alpha_2,\ldots,n\alpha_k)} \\ \text{(by Exercise 5.4.13(d), applied to} \\ (\alpha_1,\alpha_2,\ldots,\alpha_k) \text{ instead of } (\beta_1,\beta_2,\ldots,\beta_s))}} \otimes \underbrace{\mathbf{F}_n\left(M_{(\alpha_{k+1},\alpha_{k+2},\ldots,\alpha_\ell)}\right)}_{\substack{=M_{(n\alpha_{k+1},n\alpha_{k+2},\ldots,n\alpha_\ell)} \\ \text{(by Exercise 5.4.13(d), applied to} \\ (\alpha_{k+1},\alpha_{k+2},\ldots,\alpha_\ell) \text{ instead of } (\beta_1,\beta_2,\ldots,\beta_s))}} \\
&= \sum_{k=0}^{\ell} M_{(n\alpha_1,n\alpha_2,\ldots,n\alpha_k)} \otimes M_{(n\alpha_{k+1},n\alpha_{k+2},\ldots,n\alpha_\ell)}. \qquad (13.131.2)
\end{aligned}$$

But

$$\begin{aligned}(\Delta \circ \mathbf{F}_n)(M_\alpha) &= \Delta\left(\mathbf{F}_n\left(\underbrace{M_\alpha}_{\substack{=M_{(\alpha_1,\alpha_2,\ldots,\alpha_\ell)}\\ (\text{since } \alpha=(\alpha_1,\alpha_2,\ldots,\alpha_\ell))}}\right)\right) = \Delta\left(\underbrace{\mathbf{F}_n\left(M_{(\alpha_1,\alpha_2,\ldots,\alpha_\ell)}\right)}_{\substack{=M_{(n\alpha_1,n\alpha_2,\ldots,n\alpha_\ell)}\\ (\text{by Exercise 5.4.13(d), applied to}\\ (\alpha_1,\alpha_2,\ldots,\alpha_\ell) \text{ instead of } (\beta_1,\beta_2,\ldots,\beta_s))}}\right)\\
&= \Delta M_{(n\alpha_1,n\alpha_2,\ldots,n\alpha_\ell)} = \sum_{k=0}^{\ell} M_{(n\alpha_1,n\alpha_2,\ldots,n\alpha_k)} \otimes M_{(n\alpha_{k+1},n\alpha_{k+2},\ldots,n\alpha_\ell)}\\
&\qquad \left(\begin{array}{c}\text{by (13.131.1), applied to } (n\alpha_1,n\alpha_2,\ldots,n\alpha_\ell) \text{ and}\\ (n\alpha_1,n\alpha_2,\ldots,n\alpha_\ell) \text{ instead of } \alpha \text{ and } (\alpha_1,\alpha_2,\ldots,\alpha_\ell)\end{array}\right)\\
&= (\mathbf{F}_n \otimes \mathbf{F}_n)(\Delta M_\alpha) \qquad (\text{by (13.131.2)})\\
&= ((\mathbf{F}_n \otimes \mathbf{F}_n) \circ \Delta)(M_\alpha).\end{aligned}$$

Let us now forget that we fixed α. We thus have shown that

$$(\Delta \circ \mathbf{F}_n)(M_\alpha) = ((\mathbf{F}_n \otimes \mathbf{F}_n) \circ \Delta)(M_\alpha) \qquad \text{for every } \alpha \in \operatorname{Comp}. \tag{13.131.3}$$

Now, let us recall that the family $(M_\alpha)_{\alpha\in\operatorname{Comp}}$ is a basis of the $\mathbf{k}$-module QSym. The two $\mathbf{k}$-linear maps $\Delta \circ \mathbf{F}_n$ and $(\mathbf{F}_n \otimes \mathbf{F}_n) \circ \Delta$ are equal to each other on this basis (according to (13.131.3)), and therefore must be identical (because if two $\mathbf{k}$-linear maps from the same domain are equal to each other on a basis of their domain, then these two maps must be identical). In other words, $\Delta \circ \mathbf{F}_n = (\mathbf{F}_n \otimes \mathbf{F}_n) \circ \Delta$. We can prove (using a similar but simpler argument) that $\epsilon \circ \mathbf{F}_n = \epsilon$. Thus, the map $\mathbf{F}_n$ is a $\mathbf{k}$-coalgebra homomorphism (since it is $\mathbf{k}$-linear and satisfies $\Delta \circ \mathbf{F}_n = (\mathbf{F}_n \otimes \mathbf{F}_n) \circ \Delta$ and $\epsilon \circ \mathbf{F}_n = \epsilon$).

We now know that $\mathbf{F}_n$ is a $\mathbf{k}$-algebra homomorphism and a $\mathbf{k}$-coalgebra homomorphism. Hence, $\mathbf{F}_n$ is a $\mathbf{k}$-bialgebra homomorphism, thus a Hopf algebra homomorphism (due to Corollary 1.4.27). This solves Exercise 5.4.13(e).

(f) Let $n \in \{1,2,3,\ldots\}$. For every $a \in \Lambda$, we have

$$\begin{aligned}(\mathbf{F}_n\mid_\Lambda)(a) = \mathbf{F}_n(a) &= a(x_1^n, x_2^n, x_3^n, \ldots) \qquad (\text{by the definition of } \mathbf{F}_n)\\
&= \mathbf{f}_n(a) \qquad (\text{since } \mathbf{f}_n(a) = a(x_1^n, x_2^n, x_3^n, \ldots) \text{ (by the definition of } \mathbf{f}_n)).\end{aligned}$$

Thus, $\mathbf{F}_n\mid_\Lambda = \mathbf{f}_n$. This solves Exercise 5.4.13(f).

(g) The solution of Exercise 5.4.13(g) can be obtained from the solution of Exercise 2.9.9(f) upon replacing $\mathbf{f}_p$ and Λ by $\mathbf{F}_p$ and QSym, and replacing the word "symmetric" by "quasisymmetric".

(h) *Alternative solution of Exercise 2.9.9(d).* Fix $n \in \{1,2,3,\ldots\}$. Exercise 5.4.13(f) yields that $\mathbf{F}_n\mid_\Lambda = \mathbf{f}_n$. Hence, $\mathbf{f}_n = \mathbf{F}_n\mid_\Lambda$ is the restriction of $\mathbf{F}_n$ to the Hopf subalgebra Λ of QSym. Thus, $\mathbf{f}_n$ is the restriction of a Hopf algebra homomorphism to the Hopf subalgebra Λ of QSym (since $\mathbf{F}_n$ is a Hopf algebra homomorphism (by Exercise 5.4.13(e))). Consequently, $\mathbf{f}_n$ is a Hopf algebra homomorphism itself (since the restriction of a Hopf algebra homomorphism to a Hopf subalgebra must always be a Hopf algebra homomorphism). This gives a new solution to Exercise 2.9.9(d). Thus, Exercise 5.4.13(h) is solved.

13.132. **Solution to Exercise 5.4.14.** *Solution to Exercise 5.4.14.* Let us first notice that every positive integer n satisfies

$$\mathbf{V}_n(H_m) = \begin{cases} H_{m/n}, & \text{if } n \mid m;\\ 0, & \text{if } n \nmid m\end{cases} \qquad \text{for every } m \in \mathbb{N} \tag{13.132.1}$$

[927].

[927] *Proof.* To obtain a proof of (13.132.1), it is enough to repeat the proof of (13.87.1), making just the following changes:

- replacing every h_i by H_i;
- replacing $\mathbf{v}_n$ by $\mathbf{V}_n$.

Recall that NSym is (isomorphic to) the free associative algebra with generators $H_1, H_2, H_3, \ldots$ (according to (5.4.1)). Hence, the elements $H_1, H_2, H_3, \ldots$ generate the $\mathbf{k}$-algebra NSym. In other words,

$$\text{the family } \left(H_r\right)_{r\geq 1} \text{ generates the } \mathbf{k}\text{-algebra NSym}. \tag{13.132.2}$$

(e) A solution to Exercise 5.4.14(e) can be obtained by copying the solution of Exercise 2.9.10(e) and making the following changes:

- Replace every appearance of $\mathbf{v}_n$ by $\mathbf{V}_n$.
- Replace every appearance of Λ by NSym.
- Replace every appearance of h_j (for some $j \in \mathbb{N}$) by H_j.
- Replace the reference to Proposition 2.3.6(iii) by a reference to (5.4.2).
- Replace the reference to Proposition 2.4.1 by a reference to (13.132.2).
- Replace every reference to (13.87.1) by a reference to (13.132.1).

(c) A solution to Exercise 5.4.14(c) can be obtained by copying the solution of Exercise 2.9.10(c) and making the following changes:

- Replace every appearance of $\mathbf{v}_j$ (for some positive integer j) by $\mathbf{V}_j$.
- Replace every appearance of Λ by NSym.
- Replace every appearance of h_j (for some $j \in \mathbb{N}$) by H_j.
- Replace the reference to Proposition 2.4.1 by a reference to (13.132.2).

(d) A solution to Exercise 5.4.14(d) can be obtained by copying the solution of Exercise 2.9.10(d) and making the following changes:

- Replace every appearance of $\mathbf{v}_1$ by $\mathbf{V}_1$.
- Replace every appearance of Λ by NSym.
- Replace every appearance of h_j (for some $j \in \mathbb{N}$) by H_j.
- Replace the reference to Proposition 2.4.1 by a reference to (13.132.2).

(a) Here is one of several possible solutions of Exercise 5.4.14(a):[928] Define a $\mathbf{k}$-linear map $E : \text{NSym} \to \text{NSym}$ as in Exercise 1.5.14 (but with NSym instead of A). Every positive integer n satisfies

$$\Psi_n = \left(S \star E\right)\left(H_n\right) \tag{13.132.3}$$

(according to Exercise 5.4.12(c)).

Fix a positive integer n. Let us make some auxiliary observations first:

- Every composition $(\alpha_1, \alpha_2, \ldots, \alpha_\ell)$ satisfies

$$H_{(\alpha_1,\alpha_2,\ldots,\alpha_\ell)} = H_{\alpha_1} H_{\alpha_2} \cdots H_{\alpha_\ell}. \tag{13.132.4}$$

[929]

- If $(\alpha_1, \alpha_2, \ldots, \alpha_\ell)$ is a composition such that (not every $i \in \{1, 2, \ldots, \ell\}$ satisfies $n \mid \alpha_i$), then

$$\mathbf{V}_n\left(H_{(\alpha_1,\alpha_2,\ldots,\alpha_\ell)}\right) = 0. \tag{13.132.5}$$

[930]

[928] Another solution (which the reader can easily find) proceeds by making some relatively straightforward modifications to the solution of Exercise 2.9.10(a). (One needs to keep in mind that NSym, unlike Λ, is not commutative – but this does not prevent us from doing things such as substituting t^n for t in a power series over NSym (since t^n is a central element of NSym $[[t]]$).)

[929] This is precisely the equality (5.4.3), and has been proven before.

[930] *Proof of (13.132.5):* Let $(\alpha_1, \alpha_2, \ldots, \alpha_\ell)$ be a composition such that (not every $i \in \{1, 2, \ldots, \ell\}$ satisfies $n \mid \alpha_i$). Then, there exists an $i \in \{1, 2, \ldots, \ell\}$ satisfying $n \nmid \alpha_i$ (since not every $i \in \{1, 2, \ldots, \ell\}$ satisfies $n \mid \alpha_i$). Consider this i. We have

$$\begin{aligned}\mathbf{V}_n\left(H_{\alpha_i}\right) &= \begin{cases} H_{\alpha_i/n}, & \text{if } n \mid \alpha_i; \\ 0, & \text{if } n \nmid \alpha_i \end{cases} \qquad \left(\text{by the definition of } \mathbf{V}_n\left(H_{\alpha_i}\right)\right) \\ &= 0 \qquad \left(\text{since } n \nmid \alpha_i\right).\end{aligned}$$

But (13.132.4) yields $H_{(\alpha_1,\alpha_2,\ldots,\alpha_\ell)} = H_{\alpha_1} H_{\alpha_2} \cdots H_{\alpha_\ell}$. Applying $\mathbf{V}_n$ to both sides of this equality, we obtain

$$\begin{aligned}\mathbf{V}_n\left(H_{(\alpha_1,\alpha_2,\ldots,\alpha_\ell)}\right) &= \mathbf{V}_n\left(H_{\alpha_1} H_{\alpha_2} \cdots H_{\alpha_\ell}\right) = \mathbf{V}_n\left(H_{\alpha_1}\right) \mathbf{V}_n\left(H_{\alpha_2}\right) \cdots \mathbf{V}_n\left(H_{\alpha_\ell}\right) \\ &\qquad \left(\text{since } \mathbf{V}_n \text{ is a } \mathbf{k}\text{-algebra homomorphism}\right) \\ &= 0 \qquad \left(\begin{array}{c}\text{since at least one factor of the product} \\ \mathbf{V}_n\left(H_{\alpha_1}\right) \mathbf{V}_n\left(H_{\alpha_2}\right) \cdots \mathbf{V}_n\left(H_{\alpha_\ell}\right) \text{ is 0 (namely, the factor } \mathbf{V}_n\left(H_{\alpha_i}\right)\text{)}\end{array}\right).\end{aligned}$$

- If $(\alpha_1, \alpha_2, \ldots, \alpha_\ell)$ is a composition such that (every $i \in \{1, 2, \ldots, \ell\}$ satisfies $n \mid \alpha_i$), then

$$\mathbf{V}_n\left(H_{(\alpha_1,\alpha_2,\ldots,\alpha_\ell)}\right) = H_{(\alpha_1/n,\alpha_2/n,\ldots,\alpha_\ell/n)}. \tag{13.132.6}$$

[931]

- We have

$$\mathbf{V}_n \circ E = n \cdot (E \circ \mathbf{V}_n) \tag{13.132.7}$$

(as endomorphisms of NSym). [932]

- We have

$$\mathbf{V}_n \circ (S \star E) = n \cdot ((S \star E) \circ \mathbf{V}_n) \tag{13.132.10}$$

This proves (13.132.5).

[931] *Proof of (13.132.6):* Let $(\alpha_1, \alpha_2, \ldots, \alpha_\ell)$ be a composition such that (every $i \in \{1, 2, \ldots, \ell\}$ satisfies $n \mid \alpha_i$). Then, $(\alpha_1/n, \alpha_2/n, \ldots, \alpha_\ell/n)$ is a composition. Therefore, (13.132.4) (applied to $(\alpha_1/n, \alpha_2/n, \ldots, \alpha_\ell/n)$ instead of $(\alpha_1, \alpha_2, \ldots, \alpha_\ell)$) yields $H_{(\alpha_1/n,\alpha_2/n,\ldots,\alpha_\ell/n)} = H_{\alpha_1/n}H_{\alpha_2/n}\cdots H_{\alpha_\ell/n}$.

But every $i \in \{1, 2, \ldots, \ell\}$ satisfies

$$\begin{aligned}\mathbf{V}_n\left(H_{\alpha_i}\right) &= \begin{cases} H_{\alpha_i/n}, & \text{if } n \mid \alpha_i; \\ 0, & \text{if } n \nmid \alpha_i \end{cases} && \left(\text{by the definition of } \mathbf{V}_n\left(H_{\alpha_i}\right)\right) \\ &= H_{\alpha_i/n} && \left(\text{since } n \mid \alpha_i\right).\end{aligned}$$

Multiplying these equalities for all $i \in \{1, 2, \ldots, \ell\}$, we obtain

$$\mathbf{V}_n\left(H_{\alpha_1}\right)\mathbf{V}_n\left(H_{\alpha_2}\right)\cdots\mathbf{V}_n\left(H_{\alpha_\ell}\right) = H_{\alpha_1/n}H_{\alpha_2/n}\cdots H_{\alpha_\ell/n}.$$

But (13.132.4) yields $H_{(\alpha_1,\alpha_2,\ldots,\alpha_\ell)} = H_{\alpha_1}H_{\alpha_2}\cdots H_{\alpha_\ell}$. Applying $\mathbf{V}_n$ to both sides of this equality, we obtain

$$\begin{aligned}\mathbf{V}_n\left(H_{(\alpha_1,\alpha_2,\ldots,\alpha_\ell)}\right) &= \mathbf{V}_n\left(H_{\alpha_1}H_{\alpha_2}\cdots H_{\alpha_\ell}\right) = \mathbf{V}_n\left(H_{\alpha_1}\right)\mathbf{V}_n\left(H_{\alpha_2}\right)\cdots\mathbf{V}_n\left(H_{\alpha_\ell}\right) \\ &\qquad \left(\text{since } \mathbf{V}_n \text{ is a } \mathbf{k}\text{-algebra homomorphism}\right) \\ &= H_{\alpha_1/n}H_{\alpha_2/n}\cdots H_{\alpha_\ell/n} = H_{(\alpha_1/n,\alpha_2/n,\ldots,\alpha_\ell/n)}.\end{aligned}$$

This proves (13.132.6).

[932] *Proof of (13.132.7):* Let us show that

$$\left(\mathbf{V}_n \circ E\right)\left(H_\alpha\right) = \left(n \cdot \left(E \circ \mathbf{V}_n\right)\right)\left(H_\alpha\right) \qquad \text{for every } \alpha \in \operatorname{Comp}. \tag{13.132.8}$$

Proof of (13.132.8): Let $\alpha \in \operatorname{Comp}$. Then, H_α is a homogeneous element of NSym of degree $\deg\left(H_\alpha\right) = |\alpha|$. Thus, the definition of $E\left(H_\alpha\right)$ yields $E\left(H_\alpha\right) = \underbrace{\left(\deg\left(H_\alpha\right)\right)}_{=|\alpha|} \cdot H_\alpha = |\alpha| \cdot H_\alpha$. Now,

$$\left(\mathbf{V}_n \circ E\right)\left(H_\alpha\right) = \mathbf{V}_n\left(\underbrace{E\left(H_\alpha\right)}_{=|\alpha|\cdot H_\alpha}\right) = \mathbf{V}_n\left(|\alpha| \cdot H_\alpha\right) = |\alpha| \cdot \mathbf{V}_n\left(H_\alpha\right) \qquad \left(\text{since the map } \mathbf{V}_n \text{ is } \mathbf{k}\text{-linear}\right).$$

Now, let us write the composition α in the form $(\alpha_1, \alpha_2, \ldots, \alpha_\ell)$. We distinguish between two cases:

Case 1: Not every $i \in \{1, 2, \ldots, \ell\}$ satisfies $n \mid \alpha_i$.

Case 2: Every $i \in \{1, 2, \ldots, \ell\}$ satisfies $n \mid \alpha_i$.

Let us first consider Case 1. In this case, not every $i \in \{1, 2, \ldots, \ell\}$ satisfies $n \mid \alpha_i$. Thus, (13.132.5) yields $\mathbf{V}_n\left(H_{(\alpha_1,\alpha_2,\ldots,\alpha_\ell)}\right) = 0$. Since $(\alpha_1, \alpha_2, \ldots, \alpha_\ell) = \alpha$, this rewrites as $\mathbf{V}_n\left(H_\alpha\right) = 0$. Thus, $\left(\mathbf{V}_n \circ E\right)\left(H_\alpha\right) = |\alpha| \cdot \underbrace{\mathbf{V}_n\left(H_\alpha\right)}_{=0} = 0$. Compared with

$$\left(n \cdot \left(E \circ \mathbf{V}_n\right)\right)\left(H_\alpha\right) = n \cdot \underbrace{\left(E \circ \mathbf{V}_n\right)\left(H_\alpha\right)}_{=E\left(\mathbf{V}_n\left(H_\alpha\right)\right)} = n \cdot E\left(\underbrace{\mathbf{V}_n\left(H_\alpha\right)}_{=0}\right) = n \cdot \underbrace{E\left(0\right)}_{\substack{=0 \\ (\text{since } E \text{ is } \mathbf{k}\text{-linear})}} = 0,$$

this yields $\left(\mathbf{V}_n \circ E\right)\left(H_\alpha\right) = \left(n \cdot \left(E \circ \mathbf{V}_n\right)\right)\left(H_\alpha\right)$. Hence, (13.132.8) is proven in Case 1.

Let us now consider Case 2. In this case, every $i \in \{1, 2, \ldots, \ell\}$ satisfies $n \mid \alpha_i$. Thus, (13.132.6) yields $\mathbf{V}_n\left(H_{(\alpha_1,\alpha_2,\ldots,\alpha_\ell)}\right) = H_{(\alpha_1/n,\alpha_2/n,\ldots,\alpha_\ell/n)}$. Since $(\alpha_1, \alpha_2, \ldots, \alpha_\ell) = \alpha$, this rewrites as $\mathbf{V}_n\left(H_\alpha\right) = H_{(\alpha_1/n,\alpha_2/n,\ldots,\alpha_\ell/n)}$. Thus,

$$\left(\mathbf{V}_n \circ E\right)\left(H_\alpha\right) = |\alpha| \cdot \underbrace{\mathbf{V}_n\left(H_\alpha\right)}_{=H_{(\alpha_1/n,\alpha_2/n,\ldots,\alpha_\ell/n)}} = |\alpha| \cdot H_{(\alpha_1/n,\alpha_2/n,\ldots,\alpha_\ell/n)}. \tag{13.132.9}$$

On the other hand, $H_{(\alpha_1/n,\alpha_2/n,\ldots,\alpha_\ell/n)}$ is a homogeneous element of NSym of degree $\deg\left(H_{(\alpha_1/n,\alpha_2/n,\ldots,\alpha_\ell/n)}\right) = \left|(\alpha_1/n, \alpha_2/n, \ldots, \alpha_\ell/n)\right| = \left|\underbrace{(\alpha_1, \alpha_2, \ldots, \alpha_\ell)}_{=\alpha}\right|/n = |\alpha|/n$. Hence, the definition of $E\left(H_{(\alpha_1/n,\alpha_2/n,\ldots,\alpha_\ell/n)}\right)$ yields (where S, as usual, denotes the antipode of NSym). [933]

$E\left(H_{(\alpha_1/n,\alpha_2/n,\ldots,\alpha_\ell/n)}\right) = \underbrace{\left(\deg\left(H_{(\alpha_1/n,\alpha_2/n,\ldots,\alpha_\ell/n)}\right)\right)}_{=|\alpha|/n} \cdot H_{(\alpha_1/n,\alpha_2/n,\ldots,\alpha_\ell/n)} = (|\alpha|/n)\cdot H_{(\alpha_1/n,\alpha_2/n,\ldots,\alpha_\ell/n)}$. Now,

$$\begin{aligned}
\left(n\cdot\left(E\circ\mathbf{V}_n\right)\right)\left(H_\alpha\right) &= n\cdot\underbrace{\left(E\circ\mathbf{V}_n\right)\left(H_\alpha\right)}_{=E(\mathbf{V}_n(H_\alpha))} = n\cdot E\left(\underbrace{\mathbf{V}_n\left(H_\alpha\right)}_{=H_{(\alpha_1/n,\alpha_2/n,\ldots,\alpha_\ell/n)}}\right) \\
&= n\cdot\underbrace{E\left(H_{(\alpha_1/n,\alpha_2/n,\ldots,\alpha_\ell/n)}\right)}_{=(|\alpha|/n)\cdot H_{(\alpha_1/n,\alpha_2/n,\ldots,\alpha_\ell/n)}} = n\cdot(|\alpha|/n)\cdot H_{(\alpha_1/n,\alpha_2/n,\ldots,\alpha_\ell/n)} \\
&= |\alpha|\cdot H_{(\alpha_1/n,\alpha_2/n,\ldots,\alpha_\ell/n)}.
\end{aligned}$$

Compared with (13.132.9), this yields $\left(\mathbf{V}_n\circ E\right)\left(H_\alpha\right) = \left(n\cdot\left(E\circ\mathbf{V}_n\right)\right)\left(H_\alpha\right)$. Hence, (13.132.8) is proven in Case 2.

Thus, (13.132.8) is proven in both Cases 1 and 2. Since these two Cases cover all possibilities, this shows that (13.132.8) always holds.

Now, $\left(H_\alpha\right)_{\alpha\in\mathrm{Comp}}$ is a basis of the $\mathbf{k}$-module NSym. The equality (13.132.8) shows that the two $\mathbf{k}$-linear maps $\mathbf{V}_n\circ E$ and $n\cdot\left(E\circ\mathbf{V}_n\right)$ are equal to each other on every element of this basis. Hence, these two maps $\mathbf{V}_n\circ E$ and $n\cdot\left(E\circ\mathbf{V}_n\right)$ are identical (because if two $\mathbf{k}$-linear maps having the same domain are equal to each other on every element of some basis of this domain, then these two maps must be identical). In other words, $\mathbf{V}_n\circ E = n\cdot\left(E\circ\mathbf{V}_n\right)$. This proves (13.132.7).

[933] *Proof of (13.132.10):* We have $S\star E = m\circ(S\otimes E)\circ\Delta$ (by the definition of convolution). But we know (from Exercise 5.4.14(e)) that $\mathbf{V}_n$ is a Hopf algebra homomorphism; thus, $\mathbf{V}_n\circ S = S\circ\mathbf{V}_n$. Also, $\mathbf{V}_n$ is a $\mathbf{k}$-coalgebra homomorphism (since $\mathbf{V}_n$ is a Hopf algebra homomorphism); therefore, $\left(\mathbf{V}_n\otimes\mathbf{V}_n\right)\circ\Delta = \Delta\circ\mathbf{V}_n$. Finally, $\mathbf{V}_n$ is a $\mathbf{k}$-algebra homomorphism; thus, $\mathbf{V}_n\circ m = m\circ\left(\mathbf{V}_n\otimes\mathbf{V}_n\right)$. Now,

$$\begin{aligned}
\mathbf{V}_n\circ\left(\underbrace{S\star E}_{=m\circ(S\otimes E)\circ\Delta}\right) &= \underbrace{\mathbf{V}_n\circ m}_{=m\circ(\mathbf{V}_n\otimes\mathbf{V}_n)}\circ(S\otimes E)\circ\Delta = m\circ\underbrace{\left(\mathbf{V}_n\otimes\mathbf{V}_n\right)\circ(S\otimes E)}_{=(\mathbf{V}_n\circ S)\otimes(\mathbf{V}_n\circ E)}\circ\Delta \\
&= m\circ\left(\underbrace{\left(\mathbf{V}_n\circ S\right)}_{=S\circ\mathbf{V}_n}\otimes\underbrace{\left(\mathbf{V}_n\circ E\right)}_{\substack{=n\cdot(E\circ\mathbf{V}_n) \\ \text{(by (13.132.7))}}}\right)\circ\Delta = m\circ\left(\left(S\circ\mathbf{V}_n\right)\otimes\left(n\cdot\left(E\circ\mathbf{V}_n\right)\right)\right)\circ\Delta \\
&= n\cdot\left(m\circ\underbrace{\left(\left(S\circ\mathbf{V}_n\right)\otimes\left(E\circ\mathbf{V}_n\right)\right)}_{=(S\otimes E)\circ(\mathbf{V}_n\otimes\mathbf{V}_n)}\circ\Delta\right) \qquad \text{(since n is just a scalar factor)} \\
&= n\cdot\left(m\circ(S\otimes E)\circ\underbrace{\left(\mathbf{V}_n\otimes\mathbf{V}_n\right)\circ\Delta}_{=\Delta\circ\mathbf{V}_n}\right) = n\cdot\left(\underbrace{m\circ(S\otimes E)\circ\Delta}_{\substack{=S\star E \\ \text{(since } S\star E=m\circ(S\otimes E)\circ\Delta)}}\circ\mathbf{V}_n\right) = n\cdot\left(\left(S\star E\right)\circ\mathbf{V}_n\right).
\end{aligned}$$

This proves (13.132.10).

But fix a positive integer m. We have $\Psi_m = (S \star E)(H_m)$ (by (13.132.3), applied to m instead of n) and thus

$$\begin{aligned}
\mathbf{V}_n\left(\underbrace{\Psi_m}_{=(S\star E)(H_m)}\right) &= \mathbf{V}_n\left((S\star E)(H_m)\right) = \underbrace{\left(\mathbf{V}_n \circ (S\star E)\right)}_{\substack{=n\cdot((S\star E)\circ \mathbf{V}_n) \\ \text{(by (13.132.10))}}}(H_m) \\
&= \left(n\cdot\left((S\star E)\circ \mathbf{V}_n\right)\right)(H_m) = n\cdot \underbrace{\left((S\star E)\circ \mathbf{V}_n\right)(H_m)}_{=(S\star E)(\mathbf{V}_n(H_m))} \\
&= n\cdot (S\star E)\left(\underbrace{\mathbf{V}_n(H_m)}_{\substack{=\begin{cases} H_{m/n}, & \text{if } n\mid m; \\ 0, & \text{if } n\nmid m\end{cases} \\ \text{(by (13.132.1))}}}\right) = n\cdot (S\star E)\left(\begin{cases} H_{m/n}, & \text{if } n\mid m; \\ 0, & \text{if } n\nmid m\end{cases}\right) \\
&= \begin{cases} n\cdot (S\star E)\left(H_{m/n}\right), & \text{if } n\mid m; \\ n\cdot (S\star E)(0), & \text{if } n\nmid m\end{cases} \\
&= \begin{cases} n\cdot (S\star E)\left(H_{m/n}\right), & \text{if } n\mid m; \\ 0, & \text{if } n\nmid m\end{cases} \qquad \left(\text{since } n\cdot \underbrace{(S\star E)(0)}_{=0} = 0\right) \\
&= \begin{cases} n\Psi_{m/n}, & \text{if } n\mid m; \\ 0, & \text{if } n\nmid m\end{cases}
\end{aligned}$$

(because if $n \mid m$, then $(S\star E)\left(H_{m/n}\right) = \Psi_{m/n}$ [934]). This solves Exercise 5.4.14(a).

(b) Recall that a ring of formal power series $R[[t]]$ over a (not necessarily commutative) $\mathbf{k}$-algebra R has a canonical topology which makes it into a topological $\mathbf{k}$-algebra[935]. Thus, in particular, $\operatorname{NSym}[[t]]$ becomes a topological $\mathbf{k}$-algebra. We shall be considering this topology when we speak of continuity.

Assume that $\mathbb{Q}$ is a subring of $\mathbf{k}$.

Define two power series $\widetilde{H}(t)$ and $\xi(t)$ in $\operatorname{NSym}[[t]]$ by

$$\widetilde{H}(t) = \sum_{n\geq 0} H_n t^n;$$
$$\xi(t) = \sum_{n\geq 1} \xi_n t^n = \log\left(\widetilde{H}(t)\right).$$

(Here, the equality $\sum_{n\geq 1} \xi_n t^n = \log\left(\widetilde{H}(t)\right)$ follows from (5.4.6).)

We have $\widetilde{H}(t) = \sum_{n\geq 0} H_n t^n = \sum_{i\geq 0} H_i t^i$ (here, we renamed the summation index n as i) and $\xi(t) = \sum_{n\geq 1} \xi_n t^n = \sum_{i\geq 1} \xi_i t^i$ (here, we renamed the summation index n as i).

Now, let n be a positive integer. The $\mathbf{k}$-algebra homomorphism $\mathbf{V}_n : \operatorname{NSym} \to \operatorname{NSym}$ induces a continuous $\mathbf{k}$-algebra homomorphism $\mathbf{V}_n[[t]] : \operatorname{NSym}[[t]] \to \operatorname{NSym}[[t]]$ given by

$$\left(\mathbf{V}_n[[t]]\right)\left(\sum_{i\in\mathbb{N}} a_i t^i\right) = \sum_{i\in\mathbb{N}} \mathbf{V}_n(a_i)\, t^i \qquad \text{for all } (a_i)_{i\in\mathbb{N}} \in \operatorname{NSym}^{\mathbb{N}}.$$

[934] *Proof.* Assume that $n \mid m$. Then, m/n is a positive integer. Hence, (13.132.3) (applied to m/n instead of n) yields $\Psi_{m/n} = (S\star E)\left(H_{m/n}\right)$, so that $(S\star E)\left(H_{m/n}\right) = \Psi_{m/n}$, qed.

[935] This is proven just as in the case when R is commutative.

This $\mathbf{k}$-algebra homomorphism $\mathbf{V}_n[[t]]$ commutes with taking logarithms (since it is continuous and a $\mathbf{k}$-algebra homomorphism) – i.e., we have

$$(\mathbf{V}_n[[t]])(\log Q) = \log((\mathbf{V}_n[[t]])(Q)) \tag{13.132.11}$$

for every $Q \in \operatorname{NSym}[[t]]$ having constant coefficient 1.

We have $\xi(t) = \sum_{i\geq 1} \xi_i t^i$, thus

$$\sum_{i\geq 1} \xi_i t^i = \xi(t) = \log\left(\underbrace{\widetilde{H}(t)}_{=\sum_{i\geq 0} H_i t^i}\right) = \log\left(\sum_{i\geq 0} H_i t^i\right).$$

We can substitute t^n for t on both sides of this equality[936]. As a result, we obtain

$$\sum_{i\geq 1} \xi_i (t^n)^i = \log\left(\sum_{i\geq 0} H_i (t^n)^i\right). \tag{13.132.12}$$

But $\widetilde{H}(t) = \sum_{i\geq 0} H_i t^i$. Applying the map $\mathbf{V}_n[[t]]$ to both sides of this equality, we obtain

$$\begin{aligned}
(\mathbf{V}_n[[t]])\left(\widetilde{H}(t)\right) &= (\mathbf{V}_n[[t]])\left(\sum_{i\geq 0} H_i t^i\right) = \sum_{i\geq 0} \underbrace{\mathbf{V}_n(H_i)}_{\substack{= \begin{cases} H_{i/n}, & \text{if } n \mid i; \\ 0, & \text{if } n \nmid i \end{cases} \\ \text{(by (13.132.1), applied to } m=i)}} t^i \\
&= \sum_{i\geq 0} \begin{cases} H_{i/n}, & \text{if } n \mid i; \\ 0, & \text{if } n \nmid i \end{cases} \cdot t^i \\
&= \sum_{\substack{i\geq 0; \\ n\mid i}} \underbrace{\begin{cases} H_{i/n}, & \text{if } n \mid i; \\ 0, & \text{if } n \nmid i \end{cases}}_{\substack{=H_{i/n} \\ \text{(since } n\mid i)}} \cdot t^i + \sum_{\substack{i\geq 0; \\ n\nmid i}} \underbrace{\begin{cases} H_{i/n}, & \text{if } n \mid i; \\ 0, & \text{if } n \nmid i \end{cases}}_{\substack{=0 \\ \text{(since } n\nmid i)}} \cdot t^i \\
&= \sum_{\substack{i\geq 0; \\ n\mid i}} H_{i/n} t^i + \underbrace{\sum_{\substack{i\geq 0; \\ n\nmid i}} 0 t^i}_{=0} = \sum_{\substack{i\geq 0; \\ n\mid i}} H_{i/n} t^i \\
&= \sum_{i\geq 0} \underbrace{H_{ni/n}}_{=H_i} \underbrace{t^{ni}}_{=(t^n)^i} \qquad \text{(here, we have substituted } ni \text{ for } i \text{ in the sum)} \\
&= \sum_{i\geq 0} H_i (t^n)^i. \qquad (13.132.13)
\end{aligned}$$

[936] This is allowed because the element t^n of $\operatorname{NSym}[[t]]$ is central. (Generally, if R is a ring (possibly not commutative) and $Q \in R[[t]]$ is a power series, then every central element z of $R[[t]]$ can be substituted for t in Q as long as the constant coefficient of z is 0. The result of this substitution is denoted by $Q(z)$. The map $R[[t]] \to R[[t]]$ which sends every Q to $Q(z)$ (for a fixed z) is a ring homomorphism $R[[t]] \to R[[t]]$.)

Now,

$$\begin{aligned}(\mathbf{V}_n[[t]])\left(\underbrace{\xi(t)}_{=\log\left(\widetilde{H}(t)\right)}\right) &= (\mathbf{V}_n[[t]])\left(\log\left(\widetilde{H}(t)\right)\right) = \log\left(\underbrace{(\mathbf{V}_n[[t]])\left(\widetilde{H}(t)\right)}_{\substack{=\sum_{i\geq 0} H_i(t^n)^i \\ \text{(by (13.132.13))}}}\right) \\ &\qquad \left(\text{by (13.132.11), applied to } Q = \widetilde{H}(t)\right) \\ &= \log\left(\sum_{i\geq 0} H_i(t^n)^i\right) = \sum_{i\geq 1}\xi_i \underbrace{(t^n)^i}_{=t^{ni}} \qquad \text{(by (13.132.12))} \\ &= \sum_{i\geq 1}\xi_i t^{ni}. \end{aligned} \tag{13.132.14}$$

Comparing this with

$$(\mathbf{V}_n[[t]])\left(\underbrace{\xi(t)}_{=\sum_{i\geq 1}\xi_i t^i}\right) = (\mathbf{V}_n[[t]])\left(\sum_{i\geq 1}\xi_i t^i\right) = \sum_{i\geq 1}\mathbf{V}_n(\xi_i)\,t^i \qquad (\text{by the definition of } \mathbf{V}_n[[t]]),$$

we obtain

$$\sum_{i\geq 1}\mathbf{V}_n(\xi_i)\,t^i = \sum_{i\geq 1}\xi_i t^{ni}. \tag{13.132.15}$$

Now, let m be a positive integer. Comparing coefficients before t^m in the equality (13.132.15), we obtain

$$\mathbf{V}_n(\xi_m) = \begin{cases}\xi_{m/n}, & \text{if } n \mid m; \\ 0, & \text{if } n \nmid m\end{cases}.$$

This solves Exercise 5.4.14(b).

(f) Recall first that $(H_\alpha)_{\alpha\in\mathrm{Comp}}$ and $(M_\alpha)_{\alpha\in\mathrm{Comp}}$ are mutually dual bases with respect to the dual pairing $\mathrm{NSym}\otimes\mathrm{QSym}\xrightarrow{(\cdot,\cdot)}\mathbf{k}$. [937] Thus,

$$(H_\alpha, M_\beta) = \delta_{\alpha,\beta} \qquad \text{for any two compositions } \alpha \text{ and } \beta. \tag{13.132.16}$$

Let us introduce a notation: For every composition α and every positive integer s, let $\alpha\{s\}$ denote the ℓ-tuple $(s\alpha_1, s\alpha_2, \ldots, s\alpha_\ell)$, where α is written in the form $(\alpha_1,\alpha_2,\ldots,\alpha_\ell)$. Notice that if α is a composition and s is a positive integer, then $\alpha\{s\}$ is a composition again.

We have

$$\mathbf{F}_n M_\alpha = M_{\alpha\{n\}} \qquad \text{for every composition } \alpha. \tag{13.132.17}$$

[938].

[937] This follows from the definition of $(H_\alpha)_{\alpha\in\mathrm{Comp}}$.

[938] *Proof of (13.132.17):* Let α be a composition. Write α in the form $\alpha = (\alpha_1,\alpha_2,\ldots,\alpha_\ell)$. By the definition of M_α, we have $M_\alpha = \sum_{i_1<i_2<\cdots<i_\ell} x_{i_1}^{\alpha_1}x_{i_2}^{\alpha_2}\cdots x_{i_\ell}^{\alpha_\ell}$ (where the sum is over all ℓ-tuples $(i_1,i_2,\ldots,i_\ell)$ of positive integers satisfying $i_1<i_2<\cdots<i_\ell$).

On the other hand, the definition of $\alpha\{n\}$ yields $\alpha\{n\} = (n\alpha_1, n\alpha_2, \ldots, n\alpha_\ell)$ (since $\alpha = (\alpha_1,\alpha_2,\ldots,\alpha_\ell)$). Hence, the definition of $M_{\alpha\{n\}}$ yields

$$M_{\alpha\{n\}} = \sum_{i_1<i_2<\cdots<i_\ell} x_{i_1}^{n\alpha_1}x_{i_2}^{n\alpha_2}\cdots x_{i_\ell}^{n\alpha_\ell} \tag{13.132.18}$$

(where the sum is over all ℓ-tuples $(i_1,i_2,\ldots,i_\ell)$ of positive integers satisfying $i_1<i_2<\cdots<i_\ell$).

We need to prove that the maps $\mathbf{F}_n : \operatorname{QSym} \to \operatorname{QSym}$ and $\mathbf{V}_n : \operatorname{NSym} \to \operatorname{NSym}$ are adjoint with respect to the dual pairing $\operatorname{NSym} \otimes \operatorname{QSym} \xrightarrow{(\cdot,\cdot)} \mathbf{k}$. In other words, we need to prove that

$$(b, \mathbf{F}_n a) = (\mathbf{V}_n b, a) \qquad \text{for any } a \in \operatorname{QSym} \text{ and } b \in \operatorname{NSym}. \tag{13.132.19}$$

Proof of (13.132.19): Fix $a \in \operatorname{QSym}$ and $b \in \operatorname{NSym}$.

Recall that $(M_\alpha)_{\alpha \in \operatorname{Comp}}$ is a basis of the $\mathbf{k}$-module NSym. Hence, in proving (13.132.19), we can WLOG assume that a is an element of this basis $(M_\alpha)_{\alpha \in \operatorname{Comp}}$ (because the equality (13.132.19) is $\mathbf{k}$-linear in a). Assume this. Thus, a is an element of the basis $(M_\alpha)_{\alpha \in \operatorname{Comp}}$. In other words, there exists a $\beta \in \operatorname{Comp}$ such that $a = M_\beta$. Consider this β. We have $\mathbf{F}_n \underbrace{a}_{=M_\beta} = \mathbf{F}_n M_\beta = M_{\beta\{n\}}$ (by (13.132.17), applied to $\alpha = \beta$).

Recall that $(H_\alpha)_{\alpha \in \operatorname{Comp}}$ is a basis of the $\mathbf{k}$-module Λ. Hence, in proving (13.132.19), we can WLOG assume that b is an element of this basis $(H_\alpha)_{\alpha \in \operatorname{Comp}}$ (because the equality (13.132.19) is $\mathbf{k}$-linear in b). Assume this. Thus, b is an element of the basis $(H_\alpha)_{\alpha \in \operatorname{Comp}}$. In other words, there exists a $\gamma \in \operatorname{Comp}$ such that $b = H_\gamma$. Consider this γ. We have

$$\left(\underbrace{b}_{=H_\gamma}, \underbrace{\mathbf{F}_n a}_{=M_{\beta\{n\}}}\right) = \left(H_\gamma, M_{\beta\{n\}}\right) = \delta_{\gamma, \beta\{n\}} \tag{13.132.20}$$

(by (13.132.16), applied to γ and $\beta\{n\}$ instead of α and β).

Let us write the composition γ in the form $(\gamma_1, \gamma_2, \ldots, \gamma_\ell)$. Then, $\gamma = (\gamma_1, \gamma_2, \ldots, \gamma_\ell)$, so that $H_\gamma = H_{(\gamma_1, \gamma_2, \ldots, \gamma_\ell)} = H_{\gamma_1} H_{\gamma_2} \cdots H_{\gamma_\ell}$ (by (13.132.4), applied to $(\gamma_1, \gamma_2, \ldots, \gamma_\ell)$ instead of $(\alpha_1, \alpha_2, \ldots, \alpha_\ell)$).

Let us first assume that

$$(\text{not every } i \in \{1, 2, \ldots, \ell\} \text{ satisfies } n \mid \gamma_i). \tag{13.132.21}$$

Then, $\mathbf{V}_n \left(\underbrace{H_\gamma}_{=H_{(\gamma_1, \gamma_2, \ldots, \gamma_\ell)}} \right) = \mathbf{V}_n \left(H_{(\gamma_1, \gamma_2, \ldots, \gamma_\ell)} \right) = 0$ (by (13.132.5), applied to $(\gamma_1, \gamma_2, \ldots, \gamma_\ell)$ instead of $(\alpha_1, \alpha_2, \ldots, \alpha_\ell)$). Also, $\beta\{n\} \neq \gamma$ [939], so that $\delta_{\gamma, \beta\{n\}} = 0$. Thus, (13.132.20) becomes $(b, \mathbf{F}_n a) = \delta_{\gamma, \beta\{n\}} = 0$. Compared with $\left(\mathbf{V}_n \underbrace{b}_{=H_\gamma}, a \right) = \left(\underbrace{\mathbf{V}_n (H_\gamma)}_{=0}, a \right) = (0, a) = 0$ (since the dual pairing $\operatorname{NSym} \otimes \operatorname{QSym} \xrightarrow{(\cdot,\cdot)} \mathbf{k}$ is $\mathbf{k}$-bilinear), this yields $(b, \mathbf{F}_n a) = (\mathbf{V}_n b, a)$. Thus, (13.132.19) holds.

Now, let us forget that we assumed that (13.132.21) holds. We thus have proven (13.132.19) under the assumption that (13.132.21) holds. Hence, for the rest of our proof of (13.132.19), we can WLOG assume that (13.132.21) does not hold. Assume this.

We have assumed that (13.132.21) does not hold. In other words,

$$\text{every } i \in \{1, 2, \ldots, \ell\} \text{ satisfies } n \mid \gamma_i. \tag{13.132.22}$$

Thus, γ_i / n is a positive integer for every $i \in \{1, 2, \ldots, \ell\}$ (since γ_i is a positive integer for every $i \in \{1, 2, \ldots, \ell\}$). Thus, $(\gamma_1 / n, \gamma_2 / n, \ldots, \gamma_\ell / n)$ is a composition. Let us denote this composition by ζ. We have

Now, the definition of $\mathbf{F}_n$ yields

$$\begin{aligned} \mathbf{F}_n M_\alpha &= \underbrace{M_\alpha}_{= \sum_{i_1 < i_2 < \cdots < i_\ell} x_{i_1}^{\alpha_1} x_{i_2}^{\alpha_2} \cdots x_{i_\ell}^{\alpha_\ell}} (x_1^n, x_2^n, x_3^n, \ldots) = \left(\sum_{i_1 < i_2 < \cdots < i_\ell} x_{i_1}^{\alpha_1} x_{i_2}^{\alpha_2} \cdots x_{i_\ell}^{\alpha_\ell} \right) (x_1^n, x_2^n, x_3^n, \ldots) \\ &= \sum_{i_1 < i_2 < \cdots < i_\ell} x_{i_1}^{n\alpha_1} x_{i_2}^{n\alpha_2} \cdots x_{i_\ell}^{n\alpha_\ell} = M_{\alpha\{n\}} \qquad (\text{by } (13.132.18)). \end{aligned}$$

This proves (13.132.17).

[939] *Proof.* Assume the contrary. Thus, $\beta\{n\} = \gamma$. Let us write the composition β in the form $(\beta_1, \beta_2, \ldots, \beta_q)$. Then, $\beta\{n\} = (n\beta_1, n\beta_2, \ldots, n\beta_q)$ (by the definition of $\beta\{n\}$). Hence, $(\gamma_1, \gamma_2, \ldots, \gamma_\ell) = \gamma = \beta\{n\} = (n\beta_1, n\beta_2, \ldots, n\beta_q)$. As a consequence, $\ell = q$, so that $q = \ell$ and thus $(n\beta_1, n\beta_2, \ldots, n\beta_q) = (n\beta_1, n\beta_2, \ldots, n\beta_\ell)$. Hence, $(\gamma_1, \gamma_2, \ldots, \gamma_\ell) = (n\beta_1, n\beta_2, \ldots, n\beta_q) = (n\beta_1, n\beta_2, \ldots, n\beta_\ell)$. Consequently, every $i \in \{1, 2, \ldots, \ell\}$ satisfies $\gamma_i = n\beta_i$. Thus, every $i \in \{1, 2, \ldots, \ell\}$ satisfies $n \mid \gamma_i$. This contradicts (13.132.21). This contradiction proves that our assumption was wrong, qed.

$\zeta\{n\} = \gamma$ [940] and $\mathbf{V}_n(H_\gamma) = H_\zeta$ [941] and thus

$$\left(\mathbf{V}_n \underbrace{b}_{=H_\gamma}, \underbrace{a}_{=M_\beta}\right) = \left(\underbrace{\mathbf{V}_n(H_\gamma)}_{=H_\zeta}, M_\beta\right) = (H_\zeta, M_\beta) = \delta_{\zeta,\beta} \tag{13.132.23}$$

(by (13.132.16), applied to ζ instead of α).

Now, let us write the composition β in the form $(\beta_1, \beta_2, \ldots, \beta_s)$. Then, $\beta\{n\} = (n\beta_1, n\beta_2, \ldots, n\beta_s)$ (by the definition of $\beta\{n\}$).

Now, we have the following equivalence of assertions:

$$\begin{aligned}
&(\gamma = \beta\{n\}) \\
&\iff ((\gamma_1, \gamma_2, \ldots, \gamma_\ell) = (n\beta_1, n\beta_2, \ldots, n\beta_s)) \\
&\qquad (\text{since } \gamma = (\gamma_1, \gamma_2, \ldots, \gamma_\ell) \text{ and } \beta\{n\} = (n\beta_1, n\beta_2, \ldots, n\beta_s)) \\
&\iff \left(\text{we have } \ell = s \text{, and every } i \in \{1, 2, \ldots, \ell\} \text{ satisfies } \underbrace{\gamma_i = n\beta_i}_{\substack{\text{this is equivalent to} \\ \gamma_i/n=\beta_i}}\right) \\
&\iff (\text{we have } \ell = s \text{, and every } i \in \{1, 2, \ldots, \ell\} \text{ satisfies } \gamma_i/n = \beta_i) \\
&\iff ((\gamma_1/n, \gamma_2/n, \ldots, \gamma_\ell/n) = (\beta_1, \beta_2, \ldots, \beta_s)) \\
&\iff (\zeta = \beta) \qquad (\text{since } (\gamma_1/n, \gamma_2/n, \ldots, \gamma_\ell/n) = \zeta \text{ and } (\beta_1, \beta_2, \ldots, \beta_s) = \beta). \qquad (13.132.24)
\end{aligned}$$

This equivalence shows that $\delta_{\gamma,\beta\{n\}} = \delta_{\zeta,\beta}$. But (13.132.20) becomes

$$(b, \mathbf{F}_n a) = \delta_{\gamma,\beta\{n\}} = \delta_{\zeta,\beta} = (\mathbf{V}_n b, a) \qquad (\text{by (13.132.23)}).$$

Thus, (13.132.19) is proven. As we know, this completes the solution of Exercise 5.4.14(f).

(g) We know that $\pi : \mathrm{NSym} \to \Lambda$ is a $\mathbf{k}$-algebra homomorphism. Thus, $\pi(0) = 0$. Also, we know that

$$\pi(H_n) = h_n \qquad \text{for every positive integer } n. \tag{13.132.25}$$

Now, let n be a positive integer. The maps $\mathbf{v}_n \circ \pi$ and $\pi \circ \mathbf{V}_n$ are $\mathbf{k}$-algebra homomorphisms (since π and $\mathbf{v}_n$ are $\mathbf{k}$-algebra homomorphisms).

Let r be a positive integer. Applying (13.132.25) to r instead of n, we obtain $\pi(H_r) = h_r$.

If $n \mid r$, then r/n is a positive integer. Hence,

$$\text{if } n \mid r \text{, then } \pi\left(H_{r/n}\right) = h_{r/n}. \tag{13.132.26}$$

Now,

$$(\mathbf{v}_n \circ \pi)(H_r) = \mathbf{v}_n\left(\underbrace{\pi(H_r)}_{=h_r}\right) = \mathbf{v}_n(h_r) = \begin{cases} h_{r/n}, & \text{if } n \mid r; \\ 0, & \text{if } n \nmid r \end{cases} \qquad (\text{by the definition of } \mathbf{v}_n).$$

[940] *Proof.* We have $\zeta = (\gamma_1/n, \gamma_2/n, \ldots, \gamma_\ell/n)$. Hence, the definition of $\zeta\{n\}$ yields $\zeta\{n\} = (n\gamma_1/n, n\gamma_2/n, \ldots, n\gamma_\ell/n) = (\gamma_1, \gamma_2, \ldots, \gamma_\ell) = \gamma$, qed.

[941] *Proof.* We have $\gamma = (\gamma_1, \gamma_2, \ldots, \gamma_\ell)$, so that

$$\begin{aligned}
\mathbf{V}_n(H_\gamma) &= \mathbf{V}_n\left(H_{(\gamma_1, \gamma_2, \ldots, \gamma_\ell)}\right) \\
&= H_{(\gamma_1/n, \gamma_2/n, \ldots, \gamma_\ell/n)} \qquad (\text{by (13.132.6), applied to } (\gamma_1, \gamma_2, \ldots, \gamma_\ell) \text{ instead of } (\alpha_1, \alpha_2, \ldots, \alpha_\ell)) \\
&= H_\zeta \qquad (\text{since } (\gamma_1/n, \gamma_2/n, \ldots, \gamma_\ell/n) = \zeta),
\end{aligned}$$

qed.

Compared with

$$(\pi \circ \mathbf{V}_n)(H_r) = \pi \left(\underbrace{\mathbf{V}_n(H_r)}_{\substack{= \begin{cases} H_{r/n}, & \text{if } n \mid r; \\ 0, & \text{if } n \nmid r \end{cases} \\ \text{(by the definition of } \mathbf{V}_n)}} \right) = \pi \left(\begin{cases} H_{r/n}, & \text{if } n \mid r; \\ 0, & \text{if } n \nmid r \end{cases} \right)$$

$$= \begin{cases} \pi\left(H_{r/n}\right), & \text{if } n \mid r; \\ \pi(0), & \text{if } n \nmid r \end{cases} = \begin{cases} \pi\left(H_{r/n}\right), & \text{if } n \mid r; \\ 0, & \text{if } n \nmid r \end{cases}$$

$$(\text{since } \pi(0) = 0 \text{ in the case when } n \nmid r)$$

$$= \begin{cases} h_{r/n}, & \text{if } n \mid r; \\ 0, & \text{if } n \nmid r \end{cases}$$

$$\left(\text{since } \pi\left(H_{r/n}\right) = h_{r/n} \text{ in the case when } n \mid r \text{ (according to (13.132.26))}\right),$$

this yields $(\mathbf{v}_n \circ \pi)(H_r) = (\pi \circ \mathbf{V}_n)(H_r)$.

Let us now forget that we fixed r. We thus have proven that

$$(\mathbf{v}_n \circ \pi)(H_r) = (\pi \circ \mathbf{V}_n)(H_r) \qquad \text{for every positive integer } r. \tag{13.132.27}$$

Now, recall that the family $(H_r)_{r\geq 1}$ generates the $\mathbf{k}$-algebra NSym (according to (13.132.2)). In other words, $(H_r)_{r\geq 1}$ is a generating set of the $\mathbf{k}$-algebra NSym. The two $\mathbf{k}$-algebra homomorphisms $\mathbf{v}_n \circ \pi$ and $\pi \circ \mathbf{V}_n$ are equal to each other on this generating set (according to (13.86.8)), and therefore must be identical (because if two $\mathbf{k}$-algebra homomorphisms from the same domain are equal to each other on a generating set of their domain, then these two homomorphisms must be identical). In other words, $\mathbf{v}_n \circ \pi = \pi \circ \mathbf{V}_n$. This completes the solution of Exercise 5.4.14(g).

(h) *Alternative solution of Exercise 2.9.10(f).* Let i denote the inclusion map $\Lambda \to \text{QSym}$. Then, Corollary 5.4.3 yields that the map π is adjoint to the map i with respect to the dual pairing $\text{NSym} \otimes \text{QSym} \xrightarrow{(\cdot,\cdot)} \mathbf{k}$. In other words,

$$(\pi(b), a) = (b, i(a)) \qquad \text{for every } b \in \text{NSym and } a \in \Lambda. \tag{13.132.28}$$

But Exercise 5.4.14(f) yields that the maps $\mathbf{F}_n : \text{QSym} \to \text{QSym}$ and $\mathbf{V}_n : \text{NSym} \to \text{NSym}$ are adjoint with respect to the dual pairing $\text{NSym} \otimes \text{QSym} \xrightarrow{(\cdot,\cdot)} \mathbf{k}$. In other words,

$$(b, \mathbf{F}_n a) = (\mathbf{V}_n b, a) \qquad \text{for any } a \in \text{QSym and } b \in \text{NSym}. \tag{13.132.29}$$

Now, fix a positive integer n. Let $a \in \Lambda$ and $b \in \Lambda$ be arbitrary. Then, $a \in \Lambda \subset \text{QSym}$. Exercise 5.4.13(f) yields $\mathbf{F}_n \mid_\Lambda = \mathbf{f}_n$. Thus, $\mathbf{f}_n = \mathbf{F}_n \mid_\Lambda$, so that $\mathbf{f}_n a = (\mathbf{F}_n \mid_\Lambda) a = \mathbf{F}_n a$.

On the other hand, the projection $\pi : \text{NSym} \to \Lambda$ is surjective. Hence, there exists some $b' \in \text{NSym}$ satisfying $b = \pi(b')$. Consider this b'.

Since the Hall inner product on Λ is symmetric, we have

$$(\mathbf{f}_n a, b) = \left(\underbrace{b}_{=\pi(b')}, \mathbf{f}_n a \right) = (\pi(b'), \mathbf{f}_n a) = \left(b', \underbrace{i(\mathbf{f}_n a)}_{\substack{= \mathbf{f}_n a \\ \text{(since } i \text{ is an inclusion map)}}} \right)$$

$$\text{(by (13.132.28), applied to } b' \text{ and } \mathbf{f}_n a \text{ instead of } b \text{ and } a)$$

$$= \left(b', \underbrace{\mathbf{f}_n a}_{= \mathbf{F}_n a} \right) = (b', \mathbf{F}_n a) = (\mathbf{V}_n b', a) \qquad \text{(by (13.132.29), applied to } b' \text{ instead of } b).$$

Comparing this with

$$\begin{aligned}
(a, \mathbf{v}_n b) &= \left(\mathbf{v}_n \underbrace{b}_{=\pi(b')}, a\right) \qquad \text{(since the Hall inner product on Λ is symmetric)} \\
&= \left(\underbrace{\mathbf{v}_n (\pi (b'))}_{=(\mathbf{v}_n \circ \pi)(b')}, a\right) = \left(\underbrace{\underbrace{(\mathbf{v}_n \circ \pi)}_{=\pi \circ \mathbf{V}_n}}_{\text{(by Exercise 5.4.14(g))}} (b'), a\right) = \left(\underbrace{(\pi \circ \mathbf{V}_n)(b')}_{=\pi(\mathbf{V}_n b')}, a\right) \\
&= (\pi (\mathbf{V}_n b'), a) = \left(\mathbf{V}_n b', \underbrace{i(a)}_{\substack{=a \\ \text{(since i is an inclusion map)}}}\right) \qquad \text{(by (13.132.28), applied to $\mathbf{V}_n b'$ instead of b)} \\
&= (\mathbf{V}_n b', a),
\end{aligned}$$

we obtain $(\mathbf{f}_n a, b) = (a, \mathbf{v}_n b)$.

Let us now forget that we fixed a and b. We have thus shown that $(\mathbf{f}_n a, b) = (a, \mathbf{v}_n b)$ for every $a \in \Lambda$ and $b \in \Lambda$. In other words, the maps $\mathbf{f}_n : \Lambda \to \Lambda$ and $\mathbf{v}_n : \Lambda \to \Lambda$ are adjoint with respect to the Hall inner product on Λ. Thus, Exercise 2.9.10(f) is solved once again. Hence, Exercise 5.4.14(h) is solved.

13.133. **Solution to Exercise 6.1.3.** *Solution to Exercise 6.1.3.*

Proof of Proposition 6.1.2. (a) This can be easily verified by hand, but here is a slicker way to see it: Let $\mathfrak{B}$ be the set $\mathfrak{A} \sqcup \{-\infty\}$, where $-\infty$ is a symbol. We define a total order on $\mathfrak{B}$ by extending the given total order on $\mathfrak{A}$ in such a way that every $a \in \mathfrak{A}$ satisfies $-\infty < a$. Consider the set $\mathfrak{B}^\infty$ of all infinite sequences of elements of $\mathfrak{B}$. Then, $\mathfrak{A}^*$ embeds into $\mathfrak{B}^\infty$ by identifying every word $(a_1, a_2, \ldots, a_n) \in \mathfrak{A}^*$ with the sequence $(a_1, a_2, \ldots, a_n, -\infty, -\infty, -\infty, \ldots) \in \mathfrak{B}^\infty$. It is easy to check that our order relation $\leq$ on $\mathfrak{A}^*$ thus becomes the restriction to $\mathfrak{A}^*$ of the smaller-or-equal relation of the lexicographic order on $\mathfrak{B}^\infty$. Consequently, it is a total order. Proposition 6.1.2(a) is thus proven.

(h) follows immediately from the definition of $\leq$.

(i) follows from common sense.

(k) Let $a \in \mathfrak{A}^*$ and $b \in \mathfrak{A}^*$ be such that b is nonempty. Then, a is a prefix of ab. Thus, $a \leq ab$ (by Proposition 6.1.2(h), applied to ab instead of b). But $a \neq ab$ [942]. Combined with $a \leq ab$, this yields $a < ab$. Proposition 6.1.2(k) is thus proven.

(b) Proposition 6.1.2(b) is an almost immediate consequence of the definition of $\leq$; its proof is thus left to the reader.

(c) Let $a, c, d \in \mathfrak{A}^*$ satisfy $ac \leq ad$. We have $ac = \left(a_1, a_2, \ldots, a_{\ell(a)}, c_1, c_2, \ldots, c_{\ell(c)}\right)$ and $ad = \left(a_1, a_2, \ldots, a_{\ell(a)}, d_1, d_2, \ldots, d_{\ell(d)}\right)$, and we have $ac \leq ad$. Due to the definition of $\leq$, this means that we must be in one of the following two situations:

- There exists an $i \in \{1, 2, \ldots, \min\{\ell(a) + \ell(c), \ell(a) + \ell(d)\}\}$ such that

$$\left((ac)_i < (ad)_i, \text{ and every } j \in \{1, 2, \ldots, i-1\} \text{ satisfies } (ac)_j = (ad)_j\right). \tag{13.133.1}$$

- The word ac is a prefix of ad.

If we are in the first of these two situations, then we clearly must have $i > \ell(a)$ (since otherwise, both $(ac)_i$ and $(ad)_i$ would equal a_i, and then $(ac)_i < (ad)_i$ would contradict $(ac)_i = a_i = (ad)_i$), and therefore i has the form $i = \ell(a) + i'$ for some $i' \in \{1, 2, \ldots, \min\{\ell(c), \ell(d)\}\}$. Then, (13.133.1) yields $(c_{i'} < d_{i'}$, and every $j \in \{1, 2, \ldots, i'-1\}$ satisfies $c_j = d_j)$. But this shows that $c \leq d$, and so we are done

[942] *Proof.* Assume the contrary. Then, $a = ab$. Hence, $a\varnothing = a = ab$. Cancelling a from this equality, we obtain $\varnothing = b$, so that the word b is empty. This contradicts the fact that b is nonempty. This contradiction proves that our assumption was wrong, qed.

in the first situation. In the second situation, we also have $c \leq d$ [943], and again we are done. Thus, Proposition 6.1.2(c) is proven.

(d) Let $a, b, c, d \in \mathfrak{A}^*$ satisfy $a \leq c$. We have $a = \left(a_1, a_2, \ldots, a_{\ell(a)}\right)$ and $c = \left(c_1, c_2, \ldots, c_{\ell(c)}\right)$, and we have $a \leq c$. Due to the definition of $\leq$, this means that we must be in one of the following two situations:

- There exists an $i \in \{1, 2, \ldots, \min\{\ell(a), \ell(c)\}\}$ such that
$$(a_i < c_i, \text{ and every } j \in \{1, 2, \ldots, i-1\} \text{ satisfies } a_j = c_j).$$
- The word a is a prefix of c.

The first of these situations entails $ab \leq cd$ (because every $k \in \{1, 2, \ldots, \min\{\ell(a), \ell(c)\}\}$ satisfies $(ab)_k = a_k$ and $(cd)_k = c_k$). Hence, in both situations, Proposition 6.1.2(d) is proven.

(e) Let $a, b, c, d \in \mathfrak{A}^*$ satisfy $ab \leq cd$. We have $a = \left(a_1, a_2, \ldots, a_{\ell(a)}\right)$, $b = \left(b_1, b_2, \ldots, b_{\ell(b)}\right)$, $c = \left(c_1, c_2, \ldots, c_{\ell(c)}\right)$ and $d = \left(d_1, d_2, \ldots, d_{\ell(d)}\right)$, and we have $ab \leq cd$. Due to the definition of $\leq$, this means that we must be in one of the following two situations:

- There exists an $i \in \{1, 2, \ldots, \min\{\ell(a) + \ell(b), \ell(c) + \ell(d)\}\}$ such that
$$\left((ab)_i < (cd)_i, \text{ and every } j \in \{1, 2, \ldots, i-1\} \text{ satisfies } (ab)_j = (cd)_j\right). \tag{13.133.2}$$
- The word ab is a prefix of cd.

It is easy to see that if we are in the second situation, then either we have $a \leq c$ or the word c is a prefix of a [944]. We are therefore done in the second situation. We can thus WLOG assume that we are in the first situation. Assume this. We thus have an $i \in \{1, 2, \ldots, \min\{\ell(a) + \ell(b), \ell(c) + \ell(d)\}\}$ satisfying (13.133.2). If $i \leq \min\{\ell(a), \ell(c)\}$, then this yields $a \leq c$ (because every $k \in \{1, 2, \ldots, \min\{\ell(a), \ell(c)\}\}$ satisfies $(ab)_k = a_k$ and $(cd)_k = c_k$), and we are done. If $i > \min\{\ell(a), \ell(c)\}$, then every $j \in \{1, 2, \ldots, \min\{\ell(a), \ell(c)\}\}$ satisfies $a_j = c_j$ (because every $j \in \{1, 2, \ldots, \min\{\ell(a), \ell(c)\}\}$ satisfies $j \in \{1, 2, \ldots, i-1\}$ and thus
$$\begin{aligned} a_j = (ab)_j &= (cd)_j \qquad \text{(by (13.133.2))} \\ &= c_j \end{aligned}$$
), and thus we conclude that either a is a prefix of c, or c is a prefix of a. Again, this means that we are done. Thus, Proposition 6.1.2(e) is proven.

(f) follows from (e), because if c is a prefix of a satisfying $\ell(a) \leq \ell(c)$, then $c = a$.

(g) Let $a, b, c \in \mathfrak{A}^*$ satisfy $a \leq b \leq ac$. Then, $b\varnothing = b \leq ac$. Hence, Proposition 6.1.2(e) (applied to $b, \varnothing, a, c$ instead of a, b, c, d) yields that either we have $b \leq a$ or the word a is a prefix of b. Since $b \leq a$ leads to $a = b$ (in view of $a \leq b$), we get in both of these cases that a is a prefix of b. This proves Proposition 6.1.2(g).

(j) Let $a, b, c \in \mathfrak{A}^*$ satisfy $a \leq b$ and $\ell(a) \geq \ell(b)$. Proposition 6.1.2(d) (applied to a, c, b and c) yields that either we have $ac \leq bc$ or the word a is a prefix of b. In the first of these two cases, we are obviously done. Hence, we WLOG assume that we are in the second of these two cases. Thus, the word a is a prefix of b. Since the word a is at least as long as b (in fact, $\ell(a) \geq \ell(b)$), this yields that $a = b$, so that $\underbrace{a}_{=b} c = bc \leq bc$.

This proves Proposition 6.1.2(j). □

[943] *Proof.* Let us consider the second situation. In this situation, the word ac is a prefix of ad. In other words, there exists a word z such that $ad = acz$. Consider this z. Cancelling a from the equality $ad = acz$, we obtain $d = cz$. Hence, the word c is a prefix of d. Thus, $c \leq d$, qed.

[944] *Proof.* Assume that we are in the second situation. Then, the word ab is a prefix of cd. In other words, there exists a word $z \in \mathfrak{A}^*$ such that $cd = abz$. Consider this z.

Now, a is a prefix of the word cd (since $cd = abz = a(bz)$). Also, c is a prefix of the word cd. Hence, a and c are two prefixes of the word cd. Thus, Proposition 6.1.2(i) (applied to c and cd instead of b and c) yields that either a is a prefix of c, or c is a prefix of a. Hence, either we have $a \leq c$ or c is a prefix of a (because if a is a prefix of c, then $a \leq c$). Qed.

13.134. **Solution to Exercise 6.1.7.** *Solution to Exercise 6.1.7.*

Alternative proof of Corollary 6.1.6. We shall prove Corollary 6.1.6 by strong induction on $\ell(u)+\ell(v)+\ell(w)$. That is, we fix an $N \in \mathbb{N}$, and we assume (as the induction hypothesis) that Corollary 6.1.6 holds in the case when $\ell(u)+\ell(v)+\ell(w) < N$. We then need to prove that Corollary 6.1.6 holds in the case when $\ell(u)+\ell(v)+\ell(w) = N$.

So let u, v and w be words satisfying $uv \geq vu$ and $vw \geq wv$ and $\ell(u)+\ell(v)+\ell(w) = N$. Assume that v is nonempty. We want to prove that $uw \geq wu$.

If $u = \varnothing$, then $uw \geq wu$ holds obviously (because if $u = \varnothing$, then $\underbrace{u}_{=\varnothing} w = w = w \underbrace{\varnothing}_{=u} = wu$). Hence, for the rest of this proof, we can WLOG assume that $u \neq \varnothing$. Assume this. For a similar reason, we WLOG assume that $w \neq \varnothing$. Recall also that v is nonempty; that is, $v \neq \varnothing$.

We have $\ell(u) > 0$ (since $u \neq \varnothing$) and $\ell(v) > 0$ (since $v \neq \varnothing$) and $\ell(w) > 0$ (since $w \neq \varnothing$).

Notice that

$$uvc \geq vuc \qquad \text{for every } c \in \mathfrak{A}^*. \tag{13.134.1}$$

[945] Also,

$$vwc \geq wvc \qquad \text{for every } c \in \mathfrak{A}^*. \tag{13.134.2}$$

[946]

Let us first assume that u is a prefix of w. Then, there exists a $w' \in \mathfrak{A}^*$ satisfying $w = uw'$ (since u is a prefix of w). Consider this w'. Applying (13.134.1) to $c = w'$, we obtain $uvw' \geq v \underbrace{uw'}_{=w} = vw \geq \underbrace{w}_{=uw'} v = uw'v$. That is, $uw'v \leq uvw'$. Hence, Proposition 6.1.2(c) (applied to $a = u$, $c = w'v$ and $d = vw'$) yields $w'v \leq vw'$, so that $vw' \geq w'v$. Furthermore, $\ell\left(\underbrace{w}_{=uw'}\right) = \underbrace{\ell(u)}_{>0} + \ell(w') > \ell(w')$, thus $\ell(w') < \ell(w)$ and therefore $\ell(u)+\ell(v)+\underbrace{\ell(w')}_{<\ell(w)} < \ell(u)+\ell(v)+\ell(w) = N$. Hence, Corollary 6.1.6 holds for w' instead of w (by the induction hypothesis). Consequently, we obtain $uw' \geq w'u$ (since $vw' \geq w'v$). Thus, $w'u \leq uw'$, so that $uw'u \leq uuw'$ (by Proposition 6.1.2(b), applied to $a = u$, $c = w'u$ and $d = uw'$). Thus, $uuw' \geq \underbrace{uw'}_{=w} u = wu$, so that $u \underbrace{w}_{=uw'} = uuw' \geq wu$.

Now, let us forget that we assumed that u is a prefix of w. We thus have proven that $uw \geq wu$ under the assumption that u is a prefix of w. Hence, for the rest of this proof of $uw \geq wu$, we can WLOG assume that

$$u \text{ is not a prefix of } w.$$

Assume this.

Let us next assume that v is a prefix of w. Then, there exists a $w' \in \mathfrak{A}^*$ satisfying $w = vw'$ (since v is a prefix of w). Consider this w'. We have $v \underbrace{vw'}_{=w} = vw \geq \underbrace{w}_{=vw'} v = vw'v$, thus $vw'v \leq vvw'$. Thus, Proposition 6.1.2(c) (applied to $a = v$, $c = w'v$ and $d = vw'$) yields $w'v \leq vw'$, so that $vw' \geq w'v$. Furthermore, $\ell\left(\underbrace{w}_{=vw'}\right) = \underbrace{\ell(v)}_{>0} + \ell(w') > \ell(w')$, thus $\ell(w') < \ell(w)$ and therefore $\ell(u)+\ell(v)+\underbrace{\ell(w')}_{<\ell(w)} < \ell(u)+\ell(v)+\ell(w) = N$. Hence, Corollary 6.1.6 holds for w' instead of w (by the induction hypothesis). Consequently, we obtain $uw' \geq w'u$ (since $vw' \geq w'v$). Thus, $w'u \leq uw'$, so that $vw'u \leq vuw'$ (by

[945] *Proof of (13.134.1):* Let $c \in \mathfrak{A}^*$. We have $uv \geq vu$ and thus $vu \leq uv$. Also, $\ell(vu) = \ell(v)+\ell(u) = \ell(u)+\ell(v) = \ell(uv)$. Hence, Proposition 6.1.2(j) (applied to $a = vu$ and $b = uv$) yields $vuc \leq uvc$, so that $uvc \geq vuc$. This proves (13.134.1).

[946] *Proof of (13.134.2):* Let $c \in \mathfrak{A}^*$. We have $vw \geq wv$ and thus $wv \leq vw$. Also, $\ell(wv) = \ell(w)+\ell(v) = \ell(v)+\ell(w) = \ell(vw)$. Hence, Proposition 6.1.2(j) (applied to $a = wv$ and $b = vw$) yields $wvc \leq vwc$, so that $vwc \geq wvc$. This proves (13.134.2).

Proposition 6.1.2(b), applied to $a = v$, $c = w'u$ and $d = uw'$). Thus, $vuw' \geq \underbrace{vw'}_{=w} u = wu$. But (13.134.1) (applied to $c = w'$) yields $uvw' \geq vuw' \geq wu$, thus $u \underbrace{w}_{=vw'} = uvw' \geq wu$.

Now, let us forget that we assumed that v is a prefix of w. We thus have proven that $uw \geq wu$ under the assumption that v is a prefix of w. Hence, for the rest of this proof of $uw \geq wu$, we can WLOG assume that

$$v \text{ is not a prefix of } w.$$

Assume this.

We have $vw \geq wv$, thus $wv \leq vw$. Hence, Proposition 6.1.2(e) (applied to $a = w$, $b = v$, $c = v$ and $d = w$) yields that either we have $w \leq v$ or the word v is a prefix of w. Since the word v is not a prefix of w, this yields that we have $w \leq v$. Thus,

$$v \geq w.$$

Let us next assume that w is a prefix of u. Then, there exists a $u' \in \mathfrak{A}^*$ satisfying $u = wu'$ (since w is a prefix of u). Consider this u'. We have $uv \geq vu$ and thus $\underbrace{wu'}_{=u} v = uv \geq v \underbrace{u}_{=wu'} = vwu' \geq wvu'$ (by (13.134.2), applied to $c = u'$). In other words, $wvu' \leq wu'v$. Hence, Proposition 6.1.2(c) (applied to $a = w$, $c = vu'$ and $d = u'v$) yields $vu' \leq u'v$, so that $u'v \geq vu'$. Furthermore, $\ell \left(\underbrace{u}_{=wu'} \right) = \underbrace{\ell(w)}_{>0} + \ell(u') > \ell(u')$, thus $\ell(u') < \ell(u)$ and therefore $\underbrace{\ell(u')}_{<\ell(u)} + \ell(v) + \ell(w) < \ell(u) + \ell(v) + \ell(w) = N$. Hence, Corollary 6.1.6 holds for u' instead of u (by the induction hypothesis). Consequently, we obtain $u'w \geq wu'$ (since $u'v \geq vu'$). Thus, $wu' \leq u'w$, so that $wwu' \leq wu'w$ (by Proposition 6.1.2(b), applied to $a = w$, $c = wu'$ and $d = u'w$). Thus, $wu'w \geq w \underbrace{wu'}_{=u} = wu$, so that $\underbrace{u}_{=wu'} w = wu'w \geq wu$.

Now, let us forget that we assumed that w is a prefix of u. We thus have proven that $uw \geq wu$ under the assumption that w is a prefix of u. Hence, for the rest of this proof of $uw \geq wu$, we can WLOG assume that

$$w \text{ is not a prefix of } u.$$

Assume this.

Let us now assume that u is a prefix of v. Then, there exists a $v' \in \mathfrak{A}^*$ satisfying $v = uv'$ (since u is a prefix of v). Consider this v'. We have $uv' = v \geq w = w\varnothing$. Hence, $w\varnothing \leq uv'$. Thus, Proposition 6.1.2(e) (applied to $a = w$, $b = \varnothing$, $c = u$ and $d = v'$) yields that either we have $w \leq u$ or the word u is a prefix of w. Since the word u is not a prefix of w, we can conclude from this that $w \leq u$. Thus, Proposition 6.1.2(d) (applied to $a = w$, $b = u$, $c = u$ and $d = w$) shows that either we have $wu \leq uw$ or the word w is a prefix of u. Since the word w is not a prefix of u, this yields that $wu \leq uw$, so that $uw \geq wu$.

Now, let us forget that we assumed that u is a prefix of v. We thus have proven that $uw \geq wu$ under the assumption that u is a prefix of v. Hence, for the rest of this proof of $uw \geq wu$, we can WLOG assume that

$$u \text{ is not a prefix of } v.$$

Assume this.

We have $uv \geq vu$, thus $vu \leq uv$. Hence, Proposition 6.1.2(e) (applied to $a = v$, $b = u$, $c = u$ and $d = v$) yields that either we have $v \leq u$ or the word u is a prefix of v. Since the word u is not a prefix of v, this yields that we have $v \leq u$. Thus, $u \geq v$. Combined with $v \geq w$, this yields $u \geq w$, so that $w \leq u$. Thus, Proposition 6.1.2(d) (applied to $a = w$, $b = u$, $c = u$ and $d = w$) shows that either we have $wu \leq uw$ or the word w is a prefix of u. Since the word w is not a prefix of u, this yields that $wu \leq uw$, so that $uw \geq wu$. Our proof of $uw \geq wu$ is thus complete.

Now, let us forget that we fixed u, v and w. We have thus shown that if u, v and w are words satisfying $uv \geq vu$ and $vw \geq wv$ and $\ell(u) + \ell(v) + \ell(w) = N$, and if v is nonempty, then $uw \geq wu$. In other words, Corollary 6.1.6 holds in the case when $\ell(u) + \ell(v) + \ell(w) = N$. This completes the induction step. We thus have proven Corollary 6.1.6 (again). □

13.135. **Solution to Exercise 6.1.9.** *Solution to Exercise 6.1.9.* If the word u is empty, then Exercise 6.1.9 is easy to solve[947]. Hence, for the rest of this solution, we can WLOG assume that the word u is nonempty. Assume this. The word u^n is nonempty (since u is nonempty and n is a positive integer). In other words, the word v^m is nonempty (since $u^n = v^m$). Thus, v is nonempty.

We have $u \underbrace{v^m}_{=u^n} = uu^n = u^{n+1} = \underbrace{u^n}_{=v^m} u = v^m u$ and $v^m v = v^{m+1} = vv^m$. Hence, Corollary 6.1.6 (applied to u, v^m and v instead of u, v and w) yields $uv \geq vu$ (since v^m is nonempty).

But we also have $vv^m = v^m v$ (since $v^m v = vv^m$) and $v^m u = uv^m$ (since $uv^m = v^m u$). Thus, Corollary 6.1.6 (applied to v, v^m and u instead of u, v and w) yields $vu \geq uv$ (since v^m is nonempty). Combined with $uv \geq vu$, this yields $uv = vu$. Hence, Proposition 6.1.4 yields that there exist a $t \in \mathfrak{A}^*$ and two nonnegative integers i and j such that $u = t^i$ and $v = t^j$ [948]. Consider this t and these i and j. We have $i \neq 0$ (since $t^i = u$ is nonempty), so that i is a positive integer. Also, $j \neq 0$ (since $t^j = v$ is nonempty), and therefore j is a positive integer.

We have thus shown that there exists a word t and positive integers i and j such that $u = t^i$ and $v = t^j$. Exercise 6.1.9 is thus solved.

13.136. **Solution to Exercise 6.1.10.** *Solution to Exercise 6.1.10.* We WLOG assume that u is nonempty (because otherwise, both $uv \geq vu$ and $u^n v^m \geq v^m u^n$ hold for trivial reasons). Similarly, we WLOG assume that v is nonempty.

The exercise asks us to prove the equivalence $(uv \geq vu) \Longleftrightarrow (u^n v^m \geq v^m u^n)$. We shall verify its $\Longrightarrow$ and $\Longleftarrow$ parts separately:

$\Longrightarrow$: Assume that $uv \geq vu$ holds. Then, Corollary 6.1.6 (applied to $w = v^m$) yields $uv^m \geq v^m u$ (since $vv^m = v^{m+1} = v^m v$). Thus, Corollary 6.1.6 (applied to u^n, u and v^m instead of u, v and w) yields $u^n v^m \geq v^m u^n$ (since $u^n u = u^{n+1} = uu^n$). This proves the $\Longrightarrow$ part of the equivalence $(uv \geq vu) \Longleftrightarrow (u^n v^m \geq v^m u^n)$.

$\Longleftarrow$: Assume that $u^n v^m \geq v^m u^n$ holds. Then, Corollary 6.1.6 (applied to u, u^n and v^m instead of u, v and w) yields $uv^m \geq v^m u$ (since $uu^n = u^{n+1} = u^n u$, and since the word u^n is nonempty[949]). Hence, Corollary 6.1.6 (applied to u, v^m and v instead of u, v and w) yields $uv \geq vu$ (since $v^m v = v^{m+1} = vv^m$, and since the word v^m is nonempty[950]). This proves the $\Longleftarrow$ part of the equivalence $(uv \geq vu) \Longleftrightarrow (u^n v^m \geq v^m u^n)$.

Thus, both the $\Longrightarrow$ and the $\Longleftarrow$ part of the equivalence $(uv \geq vu) \Longleftrightarrow (u^n v^m \geq v^m u^n)$ are proven, and Exercise 6.1.10 is solved.

13.137. **Solution to Exercise 6.1.11.** *Solution to Exercise 6.1.11.*

Notice that $\ell(u^n) = n\ell(u) = m\ell(v) = \ell(v^m)$ (since $\ell(v^m) = m\ell(v)$), thus $\ell(v^m) = \ell(u^n)$. Now, we have the following two logical implications:

$$(u^n v^m \geq v^m u^n) \Longrightarrow (u^n \geq v^m)$$

[951] and

$$(u^n \geq v^m) \Longrightarrow (u^n v^m \geq v^m u^n)$$

[947] *Proof.* Assume that the word u is empty. Thus, $u = \varnothing$. Hence, $u^n = \varnothing^n = \varnothing$, so that $\varnothing = u^n = v^m$ and thus $v^m = \varnothing$. Consequently, $v = \varnothing$ (since m is positive). Thus, there exists a word t and positive integers i and j such that $u = t^i$ and $v = t^j$ (namely, we can take $t = \varnothing$ and $i = 1$ and $j = 1$). In other words, Exercise 6.1.9 is solved.

[948] The i and j here are the variables called n and m in Proposition 6.1.4.

[949] because u is nonempty and n is positive

[950] because v is nonempty and m is positive

[951] *Proof.* Assume that $u^n v^m \geq v^m u^n$. We need to prove that $u^n \geq v^m$.

We have $v^m u^n \leq u^n v^m$ (since $u^n v^m \geq v^m u^n$) and $\ell(v^m) \leq \ell(u^n)$ (since $\ell(v^m) = \ell(u^n)$). Hence, $v^m \leq u^n$ (by Proposition 6.1.2(f), applied to $a = v^m$, $b = u^n$, $c = u^n$ and $d = v^m$). Thus, $u^n \geq v^m$, qed.

[952]. Combining these two implications, we obtain the equivalence $(u^n v^m \geq v^m u^n) \iff (u^n \geq v^m)$. But Exercise 6.1.10 shows that we have the equivalence $(uv \geq vu) \iff (u^n v^m \geq v^m u^n)$. Altogether, we thus obtain the following chain of equivalences:

$$(uv \geq vu) \iff (u^n v^m \geq v^m u^n) \iff (u^n \geq v^m) .$$

This solves Exercise 6.1.11.

13.138. **Solution to Exercise 6.1.12.** *Solution to Exercise 6.1.12.* If w is a nonempty word, then let us denote by $\operatorname{rad}(w)$ the shortest word p such that w is a power of p. (This is clearly well-defined because w is a power of itself, and because for every given integer λ there exists at most one word v of length λ such that w is a power of v.)

Every nonempty word w and every positive integer n satisfy

$$\operatorname{rad}(w^n) = \operatorname{rad}(w) . \qquad (13.138.1)$$

Proof of (13.138.1): Let w be a nonempty word, and let n be a positive integer. Let $q = \operatorname{rad}(w^n)$. Thus, q is the shortest word p such that w^n is a power of p. Consequently, w^n is a power of q, so that there exists a positive integer N such that $w^n = q^N$. Consider this N.

The word w^n is nonempty (since w is nonempty and n is positive). That is, the word q^N is nonempty (since $w^n = q^N$). Hence, the word q is nonempty, so that $\ell(q)$ is nonzero.

Exercise 6.1.9 (applied to w, q, n and N instead of u, v, n and m) yields that there exists a word t and positive integers i and j such that $w = t^i$ and $q = t^j$. Consider these t, i and j. We have $q = t^j$ and thus $q^N = (t^j)^N = t^{jN}$, so that $t^{jN} = q^N = w^n$. Hence, w^n is a power of t. The word t thus cannot be shorter than q (since q is the shortest word p such that w^n is a power of p). Consequently, $j = 1$ (because otherwise, t would be shorter than $t^j = q$). Hence, $q = t^j = t$ (since $j = 1$) and $w = t^i = q^i$ (since $t = q$). Thus, w is a power of q. Consequently, $\ell(q) \geq \ell(\operatorname{rad}(w))$ (since $\operatorname{rad}(w)$ is the shortest word p such that w is a power of p).

On the other hand, $\operatorname{rad}(w)$ is the shortest word p such that w is a power of p. Thus, w is a power of $\operatorname{rad}(w)$. That is, there exists a $P \in \mathbb{N}$ such that $w = (\operatorname{rad}(w))^P$. Consider this P. We have $P > 0$ (since $(\operatorname{rad}(w))^P = w$ is nonempty), so that $Pn > 0$ (since $n > 0$). Also, taking both sides of the equality $w = (\operatorname{rad}(w))^P$ to the n-th power, we obtain $w^n = \left((\operatorname{rad}(w))^P\right)^n = (\operatorname{rad}(w))^{Pn}$, so that w^n is a power of $\operatorname{rad}(w)$.

Recall that q is the shortest word p such that w^n is a power of p. Since w^n is a power of $\operatorname{rad}(w)$, this yields that $\ell(\operatorname{rad}(w)) \geq \ell(q)$. Combined with $\ell(q) \geq \ell(\operatorname{rad}(w))$, this yields $\ell(\operatorname{rad}(w)) = \ell(q)$. Now,

$$\ell\left(\underbrace{w^n}_{=(\operatorname{rad}(w))^{Pn}}\right) = \ell\left((\operatorname{rad}(w))^{Pn}\right) = Pn \cdot \underbrace{\ell(\operatorname{rad}(w))}_{=\ell(q)} = Pn \cdot \ell(q) .$$

Compared with $\ell\left(\underbrace{w^n}_{=q^N}\right) = \ell\left(q^N\right) = N \cdot \ell(q)$, this yields $Pn \cdot \ell(q) = N \cdot \ell(q)$. Division by $\ell(q)$ (which is nonzero) yields $Pn = N$. Thus,

$$q^{Pn} = q^N = w^n = (\operatorname{rad}(w))^{Pn} .$$

[952] *Proof.* Assume that $u^n \geq v^m$. We need to prove that $u^n v^m \geq v^m u^n$.

If $v^m u^n \leq u^n v^m$, then $u^n v^m \geq v^m u^n$ is obviously true. Hence, for the rest of this proof of $u^n v^m \geq v^m u^n$, we can WLOG assume that we don't have $v^m u^n \leq u^n v^m$. Assume this.

We have $v^m \leq u^n$ (since $u^n \geq v^m$). Hence, Proposition 6.1.2(d) (applied to $a = v^m$, $b = u^n$, $c = u^n$ and $d = v^m$) yields that either we have $v^m u^n \leq u^n v^m$ or the word v^m is a prefix of u^n. Since we don't have $v^m u^n \leq u^n v^m$, we therefore conclude that the word v^m is a prefix of u^n. In other words, there exists a $t \in \mathfrak{A}^*$ such that $u^n = v^m t$. Consider this t.

We have $\ell(v^m) = \ell\left(\underbrace{u^n}_{=v^m t}\right) = \ell(v^m t) = \ell(v^m) + \ell(t)$, thus $0 = \ell(t)$. Hence, the word t is empty, i.e., we have $t = \varnothing$. Thus, $u^n = v^m \underbrace{t}_{=\varnothing} = v^m$, so that $\underbrace{u^n}_{=v^m} v^m = v^m \underbrace{v^m}_{=u^n} = v^m u^n \geq v^m u^n$, qed.

Since taking the Pn-th root of a word is unambiguous (when said root exists)[953], this yields $q = \operatorname{rad}(w)$. Because of $q = \operatorname{rad}(w^n)$, this rewrites as $\operatorname{rad}(w^n) = \operatorname{rad}(w)$. This proves (13.138.1).

Next, we notice that

$$\text{any two nonempty words } u \text{ and } v \text{ satisfying } uv = vu \text{ satisfy } \operatorname{rad}(u) = \operatorname{rad}(v). \qquad (13.138.2)$$

Proof of (13.138.2): Let u and v be two nonempty words satisfying $uv = vu$. Proposition 6.1.4 yields that there exist a $t \in \mathfrak{A}^*$ and two nonnegative integers n and m such that $u = t^n$ and $v = t^m$. Consider this t and these n and m. Since $u = t^n$, we have $\operatorname{rad}(u) = \operatorname{rad}(t^n) = \operatorname{rad}(t)$ (by (13.138.1), applied to t instead of w). Similarly, $\operatorname{rad}(v) = \operatorname{rad}(t)$. Thus, $\operatorname{rad}(u) = \operatorname{rad}(t) = \operatorname{rad}(v)$, and this proves (13.138.2).

Now, we can finally solve the exercise. Let us show that

$$u_1u_i \geq u_iu_1 \qquad \text{for every } i \in \{1, 2, \ldots, k\}. \qquad (13.138.3)$$

Indeed, (13.138.3) can be proven by induction over i: The *base case* ($i = 1$) is obvious, whereas the *induction step* (proving $u_1u_{i+1} \geq u_{i+1}u_1$ using $u_1u_i \geq u_iu_1$) results from applying Corollary 6.1.6 to $u = u_1$, $v = u_i$ and $w = u_{i+1}$ (because we have $u_iu_{i+1} \geq u_{i+1}u_i$ by assumption). We thus have shown (13.138.3).

Now, we can apply (13.138.3) to $i = k$, and obtain $u_1u_k \geq u_ku_1$. But on the other hand, we can apply $u_iu_{i+1} \geq u_{i+1}u_i$ to $i = k$, and obtain $u_ku_{k+1} \geq u_{k+1}u_k$. Since $u_{k+1} = u_1$, this rewrites as $u_ku_1 \geq u_1u_k$. Contrasting this with $u_1u_k \geq u_ku_1$, we obtain $u_1u_k = u_ku_1$. Hence, (13.138.2) (applied to $u = u_1$ and $v = u_k$) yields $\operatorname{rad}(u_1) = \operatorname{rad}(u_k)$. Similarly, we can show that $\operatorname{rad}(u_i) = \operatorname{rad}(u_{i-1})$ for every $i \in \{1, 2, \ldots, k\}$, where u_0 means u_k (in fact, our situation is invariant under cyclically shifting the k-tuple $(u_1, u_2, \ldots, u_k)$). Hence, $\operatorname{rad}(u_k) = \operatorname{rad}(u_{k-1}) = \cdots = \operatorname{rad}(u_1)$. Denote this common value $\operatorname{rad}(u_k) = \operatorname{rad}(u_{k-1}) = \cdots = \operatorname{rad}(u_1)$ by t. Then, for every $i \in \{1, 2, \ldots, k\}$, we have $t = \operatorname{rad}(u_i)$, whence u_i is a power of t (because $\operatorname{rad}(u_i)$ is defined as the shortest word p such that u_i is a power of p). This solves Exercise 6.1.12.

13.139. **Solution to Exercise 6.1.21.** *Solution to Exercise 6.1.21.* (a) We shall solve Exercise 6.1.21(a) by strong induction on $\ell(u) + \ell(v)$.

Induction step: Let $N \in \mathbb{N}$. Assume that Exercise 6.1.21(a) is already solved in the case when $\ell(u)+\ell(v) < N$. We now need to solve Exercise 6.1.21(a) in the case when $\ell(u) + \ell(v) = N$.

We have assumed that Exercise 6.1.21(a) is already solved in the case when $\ell(u) + \ell(v) < N$. In other words, we know that

$$\left(\begin{array}{c}\text{if } u \in \mathfrak{A}^* \text{ and } v \in \mathfrak{A}^* \text{ are two words satisfying } uv < vu \text{ and } \ell(u) + \ell(v) < N, \\ \text{then there exists a nonempty suffix } s \text{ of } u \text{ satisfying } sv < v\end{array}\right). \qquad (13.139.1)$$

So let $u \in \mathfrak{A}^*$ and $v \in \mathfrak{A}^*$ be two words satisfying $uv < vu$ and $\ell(u) + \ell(v) = N$. We are going to prove that there exists a nonempty suffix s of u satisfying $sv < v$.

Indeed, let us assume the contrary. Thus, there exists no nonempty suffix s of u satisfying $sv < v$. Hence,

$$\text{whenever } s \text{ is a nonempty suffix of } u, \text{ we do } \textbf{not} \text{ have } sv < v. \qquad (13.139.2)$$

We have $u \neq \varnothing$ (since otherwise, we would have $u = \varnothing$ and thus $\underbrace{u}_{=\varnothing} v = v = v \underbrace{\varnothing}_{=u} = vu$, which would contradict $uv < vu$). In other words, the word u is nonempty. Similarly, the word v is nonempty. We have $\ell(u) > 0$ (since u is nonempty) and $\ell(v) > 0$ (since v is nonempty).

We have $\ell(uv) = \underbrace{\ell(u)}_{>0} + \ell(v) > \ell(v)$, thus $\ell(uv) \neq \ell(v)$, hence $uv \neq v$.

Since u is a nonempty suffix of u, we can apply (13.139.2) to $s = u$. As a result, we conclude that we do **not** have $uv < v$. Hence, we do **not** have $uv \leq v$ [954]. In other words, we do **not** have $uv \leq v\varnothing$ (since $v\varnothing = v$).

But $uv\varnothing = uv < vu$. Thus, Proposition 6.1.2(e) (applied to $a = uv$, $b = \varnothing$, $c = v$ and $d = u$) yields that either we have $uv \leq v$ or the word v is a prefix of uv. Since we do not have $uv \leq v$, we can therefore

[953] because $Pn > 0$

[954] *Proof.* Assume the contrary. Then, $uv \leq v$. Since $uv \neq v$, this yields $uv < v$. This contradicts the fact that we do **not** have $uv < v$. This contradiction proves that our assumption was wrong, qed.

conclude that the word v is a prefix of the word uv. In other words, there exists a word $g \in \mathfrak{A}^*$ such that $uv = vg$. Consider this g.

Since u is a prefix of uv, we have $u \leq uv < vu$. Thus, $u\varnothing = u < vu$. Hence, Proposition 6.1.2(e) (applied to $a = u$, $b = \varnothing$, $c = v$ and $d = u$) yields that either we have $u \leq v$ or the word v is a prefix of u. In other words, we are in one of the following two cases:

Case 1: We have $u \leq v$.

Case 2: The word v is a prefix of u.

Let us consider Case 1 first. In this case, we have $u \leq v$. Thus, Proposition 6.1.2(d) (applied to $a = u$, $b = v$, $c = v$ and $d = \varnothing$) yields that either we have $uv \leq v\varnothing$ or the word u is a prefix of v. Since we do not have $uv \leq v\varnothing$, this yields that the word u is a prefix of v. In other words, there exists a word $q \in \mathfrak{A}^*$ such that $v = uq$. Consider this q. We have $\ell\left(\underbrace{v}_{=uq}\right) = \ell(uq) = \underbrace{\ell(u)}_{>0} + \ell(q) > \ell(q)$, so that $\ell(q) < \ell(v)$. Also, $u\underbrace{uq}_{=v} = uv < \underbrace{v}_{=uq}u = uqu$. Hence, Proposition 6.1.2(c) (applied to $a = u$, $c = uq$ and $d = qu$) yields that $uq \leq qu$. Hence, $uq < qu$ [955]. Since $\ell(u) + \underbrace{\ell(q)}_{<\ell(v)} < \ell(u) + \ell(v) = N$, we can therefore apply (13.139.1) to q instead of v. As a result, we obtain that there exists a nonempty suffix s of u satisfying $sq < q$. Let us denote this s by t. Thus, t is a nonempty suffix of u satisfying $tq < q$. Thus, Proposition 6.1.2(j) (applied to $a = tq$, $b = q$ and $c = g$) yields that $tqg \leq qg$ (because $\ell(tq) = \underbrace{\ell(t)}_{\substack{>0 \\ \text{(since } t \text{ is nonempty)}}} + \ell(q) > \ell(q)$).

But $\underbrace{uq}_{=v}g = vg = uv$ (since $uv = vg$). Cancelling u from this equality, we obtain $qg = v$. Hence, $t\underbrace{v}_{=qg} = tqg \leq qg = v$. But we have $\ell(tv) = \underbrace{\ell(t)}_{\substack{>0 \\ \text{(since } t \text{ is nonempty)}}} + \ell(v) > \ell(v)$, thus $\ell(tv) \neq \ell(v)$, hence $tv \neq v$. Combined with $tv \leq v$, this yields $tv < v$.

On the other hand, t is a nonempty suffix of u. Hence, (13.139.2) (applied to $s = t$) yields that we do **not** have $tv < v$. This contradicts the fact that $tv < v$. Hence, we have obtained a contradiction in Case 1.

Let us now consider Case 2. In this case, the word v is a prefix of u. Hence, there exists a word $r \in \mathfrak{A}^*$ satisfying $u = vr$. Consider this r. Clearly, r is a suffix of u (since $u = vr$).

We have $\underbrace{vr}_{=u}v = uv < v\underbrace{u}_{=vr} = vvr$. Thus, Proposition 6.1.2(c) (applied to $a = v$, $c = rv$ and $d = vr$) yields $rv \leq vr$. Hence, $rv < vr$ [956]. Also, $\ell\left(\underbrace{u}_{=vr}\right) = \ell(vr) = \underbrace{\ell(v)}_{>0} + \ell(r) > \ell(r)$, so that $\ell(r) < \ell(u)$ and thus $\underbrace{\ell(r)}_{<\ell(u)} + \ell(v) < \ell(u) + \ell(v) = N$. Therefore, (13.139.1) can be applied to r instead of u. As a result, we obtain that there exists a nonempty suffix s of r satisfying $sv < v$. Let us denote this by t. Thus, t is a nonempty suffix of r satisfying $tv < v$.

We know that t is a suffix of the word r, which (in turn) is a suffix of u. Hence, t is a suffix of u. Thus, (13.139.2) (applied to $s = t$) yields that we do **not** have $tv < v$. This contradicts the fact that $tv < v$. Hence, we have obtained a contradiction in Case 2.

We thus have obtained a contradiction in each of the two Cases 1 and 2. Since these two Cases cover all possibilities, this yields that we always have a contradiction. Therefore, our assumption was wrong. So we have proven that there exists a nonempty suffix s of u satisfying $sv < v$.

[955] *Proof.* If we had $uq = qu$, then we would have $u\underbrace{uq}_{=qu} = uqu$, which would contradict $uuq < uqu$. Hence, we cannot have $uq = qu$. Thus, we have $uq \neq qu$. Combined with $uq \leq qu$, this yields $uq < qu$, qed.

[956] *Proof.* If we had $rv = vr$, then we would have $v\underbrace{rv}_{=vr} = vvr$, which would contradict $vrv < vvr$. Hence, we cannot have $rv = vr$. Thus, we have $rv \neq vr$. Combined with $rv \leq vr$, this yields $rv < vr$, qed.

Now, let us forget that we fixed u and v. We thus have proven that if $u \in \mathfrak{A}^*$ and $v \in \mathfrak{A}^*$ are two words satisfying $uv < vu$ and $\ell(u) + \ell(v) = N$, then there exists a nonempty suffix s of u satisfying $sv < v$. In other words, we have solved Exercise 6.1.21(a) in the case when $\ell(u) + \ell(v) = N$. This completes the induction step. Hence, Exercise 6.1.21(a) is solved by strong induction.

(b)

Alternative proof of Theorem 6.1.20. Proof of the implication $\mathcal{A} \Longrightarrow \mathcal{B}$: Assume that Assertion $\mathcal{A}$ holds. Thus, the word w is Lyndon. Hence, every nonempty proper suffix of w is $> w$ (by the definition of a Lyndon word). Now, let u and v be any nonempty words satisfying $w = uv$. Then, v is a nonempty proper suffix of w, and therefore $> w$ (since every nonempty proper suffix of w is $> w$). That is, we have $v > w$.

Now, let us forget that we fixed u and v. We thus have shown that any nonempty words u and v satisfying $w = uv$ satisfy $v > w$. In other words, Assertion $\mathcal{B}$ holds. Thus, the implication $\mathcal{A} \Longrightarrow \mathcal{B}$ is proven.

Proof of the implication $\mathcal{A} \Longrightarrow \mathcal{C}$: This implication follows from Proposition 6.1.14(b).

Proof of the implication $\mathcal{A} \Longrightarrow \mathcal{D}$: This implication follows from Proposition 6.1.14(c).

Proof of the implication $\mathcal{B} \Longrightarrow \mathcal{A}$: Assume that Assertion $\mathcal{B}$ holds. Let v be a nonempty proper suffix of w. Then, there exists a nonempty $u \in \mathfrak{A}^*$ satisfying $w = uv$ (since v is a proper suffix of w). Consider this u. Assertion $\mathcal{B}$ yields $v > w$.

Now, let us forget that we fixed v. We thus have shown that every nonempty proper suffix v of w satisfies $v > w$. By the definition of a Lyndon word, this yields that w is Lyndon (since w is nonempty), so that Assertion $\mathcal{A}$ holds. Hence, the implication $\mathcal{B} \Longrightarrow \mathcal{A}$ is proven.

Proof of the implication $\mathcal{C} \Longrightarrow \mathcal{B}$: Assume that Assertion $\mathcal{C}$ holds. Let us prove that Assertion $\mathcal{B}$ holds.

Indeed, let u and v be two nonempty words satisfying $w = uv$. We will show that $v > w$.

Let us first notice that Assertion $\mathcal{C}$ yields $v > u$.

Let $\mathfrak{K}$ be the set $\left\{k \in \mathbb{N} \mid u^k \text{ is a prefix of } v\right\}$. The integer 0 belongs to this set $\mathfrak{K}$ (since $u^0 = \varnothing$ is a prefix of v), and therefore $\mathfrak{K}$ is nonempty. Also, this set $\mathfrak{K}$ is finite[957]. Hence, the set $\mathfrak{K}$ has a maximum element (since it is nonempty and finite). Let m be this maximum element. Then, $m \in \mathfrak{K}$ but $m + 1 \notin \mathfrak{K}$.

We have $m \in \mathfrak{K} = \left\{k \in \mathbb{N} \mid u^k \text{ is a prefix of } v\right\}$. Thus, m is an element of $\mathbb{N}$ such that u^m is a prefix of v. In other words, there exists a word $v' \in \mathfrak{A}^*$ such that $v = u^m v'$. Consider this v'. Using $m + 1 \notin \mathfrak{K}$, it is easy to see that u is not a prefix of v' [958].

It is easy to see that the word v' is nonempty[959]. Also, the word uu^m is nonempty (since u is nonempty). We have $w = u \underbrace{v}_{=u^m v'} = uu^m v'$. Therefore, Assertion $\mathcal{C}$ (applied to uu^m and v' instead of u and v) yields $v' > uu^m \geq u$. Thus, $u \leq v'$. Hence, Proposition 6.1.2(d) (applied to $a = u$, $b = v'$, $c = v'$ and $d = \varnothing$) yields that either we have $uv' \leq v'\varnothing$ or the word u is a prefix of v'. Since we know that the word u is not a prefix of v', we can thus conclude that $uv' \leq v'\varnothing$. Thus, $uv' \leq v'\varnothing = v'$. Hence, Proposition 6.1.2(b) (applied

[957] *Proof.* Let i be an element of $\mathfrak{K}$. Thus, $i \in \mathfrak{K} = \left\{k \in \mathbb{N} \mid u^k \text{ is a prefix of } v\right\}$. In other words i is an element of $\mathbb{N}$ such that u^i is a prefix of v. The word u^i is not longer than v (since it is a prefix of v); thus $\ell\left(u^i\right) \leq \ell(v)$. But u is nonempty, and thus $\ell(u) \geq 1$. Hence, $\ell\left(u^i\right) = i \underbrace{\ell(u)}_{\geq 1} \geq i$, so that $i \leq \ell\left(u^i\right) \leq \ell(v)$.

Now, let us forget that we fixed i. We thus have proven that every element i of $\mathfrak{K}$ satisfies $i \leq \ell(v)$. Thus, there are only finitely many elements of $\mathfrak{K}$ (since there are only finitely many $i \in \mathbb{N}$ satisfying $i \leq \ell(v)$). In other words, the set $\mathfrak{K}$ is finite.

[958] *Proof.* Assume the contrary. Thus, u is a prefix of v'. In other words, there exists a word $t \in \mathfrak{A}^*$ such that $v' = ut$. Consider this t. We have $v = u^m \underbrace{v'}_{=ut} = \underbrace{u^m u}_{=u^{m+1}} t = u^{m+1} t$, and thus the word u^{m+1} is a prefix of v. Hence, $m + 1$ is an element of $\mathbb{N}$ such that u^{m+1} is a prefix of v. In other words, $m + 1 \in \left\{k \in \mathbb{N} \mid u^k \text{ is a prefix of } v\right\} = \mathfrak{K}$. But this contradicts $m + 1 \notin \mathfrak{K}$. This contradiction proves that our assumption was wrong, qed.

[959] *Proof.* Assume the contrary. Thus, the word v' is empty; that is, we have $v' = \varnothing$. Now, $v = u^m \underbrace{v'}_{=\varnothing} = u^m$ and $w = u \underbrace{v}_{=u^m} = uu^m = u^{m+1} = \underbrace{u^m}_{=v} u = vu$. Hence, Assertion $\mathfrak{C}$ (applied to v and u instead of u and v) yields $u > v$. This contradicts $v > u$. This contradiction proves that our assumption was wrong, qed.

to $a = u^m$, $c = uv'$ and $d = v'$) yields $u^m uv' \leq u^m v'$. Thus, $u^m v' \geq \underbrace{u^m u}_{=u^{m+1}=uu^m} v' = uu^m v' = w$. Since $u^m v' = v$, this rewrites as $v \geq w$. Hence, $v > w$ [960].

Now, let us forget that we fixed u and v. We thus have proven that any nonempty words u and v satisfying $w = uv$ satisfy $v > w$. In other words, Assertion $\mathcal{B}$ holds. Hence, the implication $\mathcal{C} \Longrightarrow \mathcal{B}$ is proven.

Proof of the implication $\mathcal{D} \Longrightarrow \mathcal{B}$: Assume that Assertion $\mathcal{D}$ holds. Let us prove that Assertion $\mathcal{B}$ holds.

We shall prove Assertion $\mathcal{B}$ by strong induction over $\ell(u)$:

Induction step: Let $N \in \mathbb{N}$. Assume that Assertion $\mathcal{B}$ holds in the case when $\ell(u) < N$. We now need to prove Assertion $\mathcal{B}$ in the case when $\ell(u) = N$.

We have assumed that Assertion $\mathcal{B}$ holds in the case when $\ell(u) < N$. In other words,

$$\text{any nonempty words } u \text{ and } v \text{ satisfying } w = uv \text{ and } \ell(u) < N \text{ satisfy } v > w. \tag{13.139.3}$$

Now, let u and v be two nonempty words satisfying $w = uv$ and $\ell(u) = N$. We shall prove that $v > w$.

From Assertion $\mathcal{D}$, we obtain $vu > uv$. Thus, $uv < vu$. Hence, Exercise 6.1.21(a) yields that there exists a nonempty suffix s of u satisfying $sv < v$. Consider this s. There exists a $p \in \mathfrak{A}^*$ such that $u = ps$ (since s is a suffix of u). Consider this p. Since s is nonempty, we have $\ell(s) > 0$. Now, $\ell\left(\underbrace{u}_{=ps}\right) = \ell(ps) = \ell(p) + \underbrace{\ell(s)}_{>0} > \ell(p)$, so that $\ell(p) < \ell(u) = N$. Also, $w = \underbrace{u}_{=ps} v = psv$. Furthermore, if $p = \varnothing$, then $v > w$ is true[961]. Hence, for the rest of our proof of $v > w$, we can WLOG assume that we don't have $p = \varnothing$. Assume this. Thus, $p \neq \varnothing$, so that the word p is nonempty. Also, the word sv is nonempty (since v is nonempty).

Now, the words p and sv are nonempty and satisfy $w = psv$ and $\ell(p) < N$. Hence, we can apply (13.139.3) to p and sv instead of u and v. As a result, we obtain $sv > w$. But $sv < v$, so that $v > sv > w$. Hence, we have proven that $v > w$.

Now, let us forget that we fixed u and v. We thus have proven that any nonempty words u and v satisfying $w = uv$ and $\ell(u) = N$ satisfy $v > w$. In other words, we have proven Assertion $\mathcal{B}$ in the case when $\ell(u) = N$. Hence, the induction step is complete. Assertion $\mathcal{B}$ is thus proven by strong induction. And so, we have established the implication $\mathcal{D} \Longrightarrow \mathcal{B}$.

Combining the implications $\mathcal{A} \Longrightarrow \mathcal{B}$, $\mathcal{A} \Longrightarrow \mathcal{C}$, $\mathcal{A} \Longrightarrow \mathcal{D}$, $\mathcal{B} \Longrightarrow \mathcal{A}$, $\mathcal{C} \Longrightarrow \mathcal{B}$ and $\mathcal{D} \Longrightarrow \mathcal{B}$ that we have proven, we obtain the equivalence $\mathcal{A} \Longleftrightarrow \mathcal{B} \Longleftrightarrow \mathcal{C} \Longleftrightarrow \mathcal{D}$. Hence, Theorem 6.1.20 is proven again. □

13.140. **Solution to Exercise 6.1.22.** *Solution to Exercise 6.1.22.* Let us first assume that w is Lyndon. We shall prove that

$$\left(\text{every nonempty word } t \text{ and every positive integer } n \text{ satisfy } \left(\text{if } w \leq t^n, \text{ then } w \leq t\right)\right). \tag{13.140.1}$$

Let t be a nonempty word, and let n be a positive integer. Assume that $w \leq t^n$. We need to prove that $w \leq t$.

Assume the contrary. Thus, we don't have $w \leq t$. In other words, we don't have $w \leq t^1$.

Let m be the minimal $i \in \{1, 2, \ldots, n\}$ satisfying $w \leq t^i$. [962] Then, $w \leq t^m$. Hence, $m \neq 1$ (because $w \leq t^m$, but we don't have $w \leq t^1$). Thus, $m \geq 2$, so that $m - 1$ is also an element of $\{1, 2, \ldots, n\}$. If we had $w \leq t^{m-1}$, then $m - 1$ would be an $i \in \{1, 2, \ldots, n\}$ satisfying $w \leq t^i$, which would contradict the fact that m is the **minimal** such i. Thus, we cannot have $w \leq t^{m-1}$.

[960] *Proof.* We have $\ell\left(\underbrace{w}_{=uv}\right) = \ell(uv) = \underbrace{\ell(u)}_{\substack{>0 \\ (\text{since } u \text{ is nonempty})}} + \ell(v) > \ell(v)$, so that $\ell(w) \neq \ell(v)$ and thus $w \neq v$. Combined with $v \geq w$, this yields $v > w$, qed.

[961] *Proof.* Assume that $p = \varnothing$. Then, $u = \underbrace{p}_{=\varnothing} s = s$, so that $s = u$. Then, $\underbrace{s}_{=u} v = uv = w$, so that $w = sv < v$ and thus $v > w$, qed.

[962] Such an i exists, because $n \in \{1, 2, \ldots, n\}$ satisfies $w \leq t^n$.

But $w\varnothing = w \leq t^{m-1}t$. Hence, Proposition 6.1.2(e) (applied to $a = w$, $b = \varnothing$, $c = t^{m-1}$ and $d = t$) yields that either we have $w \leq t^{m-1}$ or the word t^{m-1} is a prefix of w. Since we cannot have $w \leq t^{m-1}$, this shows that the word t^{m-1} is a prefix of w. In other words, there exists a $v \in \mathfrak{A}^*$ such that $w = t^{m-1}v$. Consider this v. We have $v \neq \varnothing$ (because otherwise, we would have $v = \varnothing$ and thus $w = t^{m-1}\underbrace{v}_{=\varnothing} = t^{m-1} \leq t^{m-1}$, contradicting the fact that we cannot have $w \leq t^{m-1}$), so that v is nonempty. Hence, Proposition 6.1.14(b) (applied to $u = t^{m-1}$) now yields $v > t^{m-1}$. Hence, $t^{m-1} < v$. Thus, Proposition 6.1.2(b) (applied to $a = t^{m-1}$, $c = t^{m-1}$ and $d = v$) yields $t^{m-1}t^{m-1} \leq t^{m-1}v = w$. Hence,

$$w \geq t^{m-1}t^{m-1} = t^{2(m-1)}.$$

Combined with

$$\begin{aligned} w \leq t^m \leq t^m t^{m-2} \qquad & \text{(this makes sense since } m \geq 2) \\ = t^{m+(m-2)} = t^{2(m-1)}, & \end{aligned}$$

this yields $w = t^{2(m-1)} = t^{m-1}t^{m-1}$, so that $t^{m-1}t^{m-1} = w = t^{m-1}v$. Cancelling t^{m-1} in this, we obtain $t^{m-1} = v$, which contradicts $t^{m-1} < v$. This contradiction shows that our assumption (that we don't have $w \leq t$) was false. Hence, $w \leq t$.

Forget now that we assumed that $w \leq t^n$. We thus have proven that if $w \leq t^n$, then $w \leq t$.

Now, forget that we fixed t and assumed that w is Lyndon. We thus have shown that

$$\text{(if } w \text{ is Lyndon, then (13.140.1) holds)}. \tag{13.140.2}$$

Now, conversely, assume that (13.140.1) holds. We will prove that w is Lyndon.

In fact, assume the contrary. Then, w is not Lyndon. Let v be the (lexicographically) smallest nonempty suffix of w. Proposition 6.1.19(b) yields that there exists a nonempty $u \in \mathfrak{A}^*$ such that $w = uv$, $u \geq v$ and $uv \geq vu$. Consider this u. We have $u \geq v$, thus $v \leq u$, and therefore Proposition 6.1.2(b) (applied to $a = u$, $c = v$ and $d = u$) yields $uv \leq uu$. Now, $w = uv \leq uu = u^2$. Thus, (13.140.1) (applied to $t = u$ and $n = 2$) yields $w \leq u$. But this contradicts the fact that $w = uv > u$ (since v is nonempty). As this contradiction shows, our assumption was wrong. Thus, we have shown that w is Lyndon.

Now, forget that we fixed w. We thus have proven that

$$\text{(if (13.140.1) holds, then } w \text{ is Lyndon)}.$$

Combined with (13.140.2), this yields that w is Lyndon if and only if (13.140.1) holds. This solves the exercise.

13.141. **Solution to Exercise 6.1.23.** *Solution to Exercise 6.1.23.* We will solve Exercise 6.1.23 by induction over n:

The *induction base* is the case $n = 1$; this case is vacuously true (since $w_1 < w_n$ is impossible for $n = 1$).

For the *induction step*, we fix a positive integer $N > 1$, and we assume that Exercise 6.1.23 has been solved for $n = N - 1$. We now must solve Exercise 6.1.23 for $n = N$.

So let w_1, w_2, ..., w_N be N Lyndon words. Assume that $w_1 \leq w_2 \leq \cdots \leq w_N$ and $w_1 < w_N$. We need to show that $w_1 w_2 \cdots w_N$ is a Lyndon word.

Proposition 6.1.16(a) (applied to $u = w_1$ and $v = w_N$) yields that the word $w_1 w_N$ is Lyndon. If $N = 2$, then this yields that $w_1 w_2 \cdots w_N$ is Lyndon (because $w_1 w_2 \cdots w_N = w_1 w_N$ if $N = 2$). Thus, if $N = 2$, then we are done. We therefore WLOG assume that we don't have $N = 2$. Hence, $N \geq 3$. Thus, $w_3 w_4 \cdots w_N$ is a nonempty product. More precisely, $w_3 w_4 \cdots w_N$ is a nonempty product of nonempty words (since the words w_3, w_4, ..., w_N are nonempty (because they are Lyndon)), and therefore a nonempty word itself. Hence, $w_2 < w_2 (w_3 w_4 \cdots w_N) = w_2 w_3 \cdots w_N$.

We have $w_1 \leq w_2 \leq \cdots \leq w_N$. Hence, $w_2 \leq w_3 \leq \cdots \leq w_N$ and, in particular, $w_2 \leq w_N$. We distinguish between two cases:

Case 1: We have $w_2 = w_N$.

Case 2: We have $w_2 \neq w_N$.

Let us first consider Case 1. In this case, we have $w_2 = w_N$. Hence, $w_1 < w_N = w_2$. Thus, Proposition 6.1.16(a) (applied to $u = w_1$ and $v = w_2$) yields that the word $w_1 w_2$ is Lyndon. Moreover, $w_1 w_2 < w_2$ (by

Proposition 6.1.16(b), applied to $u = w_1$ and $v = w_2$). Hence, $w_1 w_2 \leq w_3 \leq w_4 \leq \cdots \leq w_N$ (this follows by combining $w_1 w_2 < w_2$ and $w_2 \leq w_3 \leq \cdots \leq w_N$) and $w_1 w_2 < w_2 = w_N$. Thus, we can apply Exercise 6.1.23 to $N-1$ and $(w_1 w_2, w_3, w_4, \ldots, w_N)$ instead of n and $(w_1, w_2, \ldots, w_n)$ (because we have assumed that Exercise 6.1.23 has been solved for $n = N-1$). We thus obtain that $w_1 w_2 w_3 w_4 \cdots w_N$ is a Lyndon word. In other words, $w_1 w_2 \cdots w_N$ is a Lyndon word. So we have shown that $w_1 w_2 \cdots w_N$ is a Lyndon word in Case 1.

Let us now consider Case 2. In this case, we have $w_2 \neq w_N$. Since $w_2 \leq w_N$, this yields $w_2 < w_N$. Thus, we can apply Exercise 6.1.23 to $N-1$ and $(w_2, w_3, \ldots, w_N)$ instead of n and $(w_1, w_2, \ldots, w_n)$ (because we have assumed that Exercise 6.1.23 has been solved for $n = N-1$). We thus obtain that $w_2 w_3 \cdots w_N$ is a Lyndon word. Now, w_1 and $w_2 w_3 \cdots w_N$ are two Lyndon words satisfying $w_1 \leq w_2 < w_2 w_3 \cdots w_N$. Therefore, Proposition 6.1.16(a) (applied to $u = w_1$ and $v = w_2 w_3 \cdots w_N$) yields that the word $w_1 w_2 w_3 \cdots w_N$ is Lyndon. In other words, the word $w_1 w_2 \cdots w_N$ is Lyndon. So we have shown that $w_1 w_2 \cdots w_N$ is a Lyndon word in Case 2.

Now, we have proven that $w_1 w_2 \cdots w_N$ is a Lyndon word in both possible Cases 1 and 2. Hence, $w_1 w_2 \cdots w_N$ always is a Lyndon word. In other words, Exercise 6.1.23 is solved for $n = N$. This completes the induction, and therefore Exercise 6.1.23 is solved.

13.142. **Solution to Exercise 6.1.24.**

Solution to Exercise 6.1.24. We have assumed that

$$w_i w_{i+1} \cdots w_n \geq w_1 w_2 \cdots w_n \qquad \text{for every } i \in \{1, 2, \ldots, n\}. \tag{13.142.1}$$

As a consequence,

$$w_i w_{i+1} \cdots w_n > w_1 w_2 \cdots w_n \qquad \text{for every } i \in \{2, 3, \ldots, n\}. \tag{13.142.2}$$

[963]

Now, we claim that if j is any element of $\{0, 1, \ldots, n\}$, then

$$\text{every nonempty proper suffix } v \text{ of } w_{n-j+1} w_{n-j+2} \cdots w_n \text{ satisfies } v > w_1 w_2 \cdots w_n. \tag{13.142.3}$$

Proof of (13.142.3): We will prove (13.142.3) by induction over j:

Induction base: For $j = 0$, we have $w_{n-j+1} w_{n-j+2} \cdots w_n = w_{n-0+1} w_{n-0+2} \cdots w_n = (\text{empty product}) = \varnothing$. Hence, for $j = 0$, the word $w_{n-j+1} w_{n-j+2} \cdots w_n$ has no nonempty proper suffix. Thus, (13.142.3) is vacuously true for $j = 0$. The induction base is thus complete.

Induction step: Let $J \in \{0, 1, \ldots, n-1\}$. We assume that (13.142.3) holds for $j = J$. We now need to prove that (13.142.3) holds for $j = J + 1$.

From $J \in \{0, 1, \ldots, n-1\}$, we obtain $n - J \in \{1, 2, \ldots, n\}$.

Let $g = w_{n-J+1} w_{n-J+2} \cdots w_n$. We have assumed that (13.142.3) holds for $j = J$. In other words, every nonempty proper suffix v of $w_{n-J+1} w_{n-J+2} \cdots w_n$ satisfies $v > w_1 w_2 \cdots w_n$. Since $w_{n-J+1} w_{n-J+2} \cdots w_n = g$, this rewrites as follows:

$$\text{Every nonempty proper suffix } v \text{ of } g \text{ satisfies } v > w_1 w_2 \cdots w_n. \tag{13.142.4}$$

Now, let us notice that

$$\begin{aligned} w_{n-(J+1)+1} w_{n-(J+1)+2} \cdots w_n = w_{n-J} w_{n-J+1} \cdots w_n &= w_{n-J} \underbrace{(w_{n-J+1} w_{n-J+2} \cdots w_n)}_{=g} \\ &= w_{n-J} g. \end{aligned} \tag{13.142.5}$$

Let now v be a nonempty proper suffix of $w_{n-J} g$. We are going to prove that $v > w_1 w_2 \cdots w_n$.

We have $v \neq \varnothing$ (since v is nonempty). Since v is a nonempty suffix of $w_{n-J} g$, we must be in one of the following two cases (depending on whether this suffix begins before the suffix g of $w_{n-J} g$ begins or afterwards):

[963] *Proof of (13.142.2):* Let $i \in \{2, 3, \ldots, n\}$. Then, $i - 1 \geq 1$. The words w_1, w_2, $\ldots$, w_{i-1} are Lyndon (since the words w_1, w_2, $\ldots$, w_n are Lyndon), and thus nonempty. Now, $i - 1 \geq 1 > 0$. Hence, the product $w_1 w_2 \cdots w_{i-1}$ is a nonempty product of nonempty words (since w_1, w_2, $\ldots$, w_{i-1} are nonempty words), and thus a nonempty word itself. Since $w_1 w_2 \cdots w_n = (w_1 w_2 \cdots w_{i-1})(w_i w_{i+1} \cdots w_n)$, this yields that $w_i w_{i+1} \cdots w_n$ is a **proper** suffix of the word $w_1 w_2 \cdots w_n$. As a consequence, $w_i w_{i+1} \cdots w_n \neq w_1 w_2 \cdots w_n$. Combining this with $w_i w_{i+1} \cdots w_n \geq w_1 w_2 \cdots w_n$ (by (13.142.1)), we obtain $w_i w_{i+1} \cdots w_n > w_1 w_2 \cdots w_n$. This proves (13.142.2).

Case 1: The word v is a nonempty suffix of g. (Note that v is allowed to be g.)

Case 2: The word v has the form hg where h is a nonempty suffix of w_{n-J}.

Let us first consider Case 1. In this case, the word v is a nonempty suffix of g. If v is a proper suffix of g, then we immediately obtain $v > w_1 w_2 \cdots w_n$ (from (13.142.4)). Thus, for the rest of the proof of $v > w_1 w_2 \cdots w_n$ in Case 1, we can WLOG assume that v is not a proper suffix of g. Assume this.

So we know that v is a suffix of g, but not a proper suffix of g. Hence, v must be g itself. That is, we have $v = g$. Hence, $v = g = w_{n-J+1} w_{n-J+2} \cdots w_n$. Consequently, $J \neq 0$ [964]. Combined with $J \in \{0, 1, \ldots, n-1\}$, this yields $J \in \{0, 1, \ldots, n-1\} \setminus \{0\} = \{1, 2, \ldots, n-1\}$, so that $n-J+1 \in \{2, 3, \ldots, n\}$. Now,

$$v = g = w_{n-J+1} w_{n-J+2} \cdots w_n = w_{n-J+1} w_{(n-J+1)+1} \cdots w_n > w_1 w_2 \cdots w_n$$

(by (13.142.2), applied to $i = n - J + 1$). Thus, $v > w_1 w_2 \cdots w_n$ is proven in Case 1.

Let us now consider Case 2. In this case, the word v has the form hg where h is a nonempty suffix of w_{n-J}. Consider this h. Since h is a suffix of w_{n-J}, we have $\ell(h) \leq \ell(w_{n-J})$, so that $\ell(w_{n-J}) \geq \ell(h)$. But w_{n-J} is a Lyndon word (since w_1, w_2, $\ldots$, w_n are Lyndon words), and thus $h \geq w_{n-J}$ (by Corollary 6.1.15, applied to w_{n-J} and h instead of w and v). Thus, $w_{n-J} \leq h$. Thus, Proposition 6.1.2(j) (applied to $a = w_{n-J}$, $b = h$ and $c = g$) yields $w_{n-J} g \leq hg = v$. Thus, $v \geq w_{n-J} g$.

But we also have $v \neq w_{n-J} g$ (since v is a proper suffix of $w_{n-J} g$). Combined with $v \geq w_{n-J} g$, this yields $v > w_{n-J} g$, so that

$$\begin{aligned} v > w_{n-J} g &= w_{n-(J+1)+1} w_{n-(J+1)+2} \cdots w_n && \text{(by (13.142.5))} \\ &= w_{n-J} w_{n-J+1} \cdots w_n \geq w_1 w_2 \cdots w_n && \text{(by (13.142.1), applied to } i = n - J) . \end{aligned}$$

Thus, $v > w_1 w_2 \cdots w_n$ is proven in Case 2.

Now, $v > w_1 w_2 \cdots w_n$ is proven in each of the two Cases 1 and 2. Since these two Cases cover all possibilities, this yields that $v > w_1 w_2 \cdots w_n$ always holds.

Now, let us forget that we fixed v. We thus have proven that every nonempty proper suffix v of $w_{n-J} g$ satisfies $v > w_1 w_2 \cdots w_n$. Since $w_{n-J} g = w_{n-(J+1)+1} w_{n-(J+1)+2} \cdots w_n$ (by (13.142.5)), this rewrites as follows: Every nonempty proper suffix v of $w_{n-(J+1)+1} w_{n-(J+1)+2} \cdots w_n$ satisfies $v > w_1 w_2 \cdots w_n$. In other words, (13.142.3) holds for $j = J + 1$. This completes the induction step. Thus, (13.142.3) is proven by induction.

Now, we can apply (13.142.3) to $j = n$. As a result, we obtain that every nonempty proper suffix v of $w_{n-n+1} w_{n-n+2} \cdots w_n$ satisfies $v > w_1 w_2 \cdots w_n$. In other words, every nonempty proper suffix v of $w_1 w_2 \cdots w_n$ satisfies $v > w_1 w_2 \cdots w_n$ (since $w_{n-n+1} w_{n-n+2} \cdots w_n = w_1 w_2 \cdots w_n$). Since the word $w_1 w_2 \cdots w_n$ is also nonempty[965], this shows that the word $w_1 w_2 \cdots w_n$ is Lyndon (by the definition of a Lyndon word). This solves Exercise 6.1.24.

13.143. **Solution to Exercise 6.1.29.**

Solution to Exercise 6.1.29. Let us first forget about the setting of Exercise 6.1.29 (so $\mathfrak{A}$ can be any alphabet, not necessarily finite).

Our solution to Exercise 6.1.29 will rely on two basic propositions:

Proposition 13.143.1. *Let $\mathfrak{M}$ denote the set of all finite multisets of Lyndon words.*

Define a map $\mathbf{m} : \mathfrak{M} \to \mathfrak{A}^*$ as follows: Given an $M \in \mathfrak{M}$, we set $\mathbf{m}(M) = a_1 a_2 \cdots a_k$, where $a_1, a_2, \ldots, a_k$ denote the elements of M listed in decreasing order[966]*. Thus, a map $\mathbf{m} : \mathfrak{M} \to \mathfrak{A}^*$ is defined.*

[964] *Proof.* Assume the contrary. Then, $J = 0$, so that

$$\begin{aligned} v = w_{n-J+1} w_{n-J+2} \cdots w_n &= w_{n-0+1} w_{n-0+2} \cdots w_n && \text{(since } J = 0) \\ &= (\text{empty product}) = \varnothing, \end{aligned}$$

contradicting the fact that $v \neq \varnothing$. This contradiction shows that our assumption was wrong, qed.

[965] *Proof.* The product $w_1 w_2 \cdots w_n$ is nonempty (since n is a positive integer), and the words w_1, w_2, $\ldots$, w_n are nonempty (since they are Lyndon words). Hence, $w_1 w_2 \cdots w_n$ is a nonempty product of nonempty words. Thus, $w_1 w_2 \cdots w_n$ is a nonempty word, qed.

[966] When we say "the elements of M listed in decreasing order", we mean that $(a_1, a_2, \ldots, a_k)$ should be the unique tuple satisfying $a_1 \geq a_2 \geq \cdots \geq a_k$ and $M = \{a_1, a_2, \ldots, a_k\}_{\text{multiset}}$. (The existence and the uniqueness of this tuple follow from basic properties of finite multisets, since $\geq$ is a total order.) Note that each element appears in the tuple $(a_1, a_2, \ldots, a_k)$ with the same multiplicity with which it appears in the multiset M.

Let $\mathbf{n} : \mathfrak{A}^* \to \mathfrak{M}$ *be the map that sends each word* $w \in \mathfrak{A}^*$ *to the multiset* $\{a_1, a_2, \ldots, a_k\}_{\text{multiset}}$, *where* $(a_1, a_2, \ldots, a_k)$ *is the CFL factorization of* w.

Then, these two maps $\mathbf{m}$ *and* $\mathbf{n}$ *are mutually inverse bijections.*

Note that $\mathfrak{A}$ is **not** assumed to be finite in Proposition 13.143.1.

Proof of Proposition 13.143.1. This follows from the definition of the CFL factorization (and from the fact that it exists and is unique).[967] □

Proposition 13.143.2. *Let* S *be any set.*

Let $\mathfrak{M}$ *be the set of all finite multisets of elements of* S.

Let $\mathfrak{N}$ *be the set of all families* $(k_w)_{w \in S} \in \mathbb{N}^S$ *of nonnegative integers (indexed by the elements of* S*) such that all but finitely many* $w \in S$ *satisfy* $k_w = 0$.

Then, the map mult $: \mathfrak{M} \to \mathfrak{N}$ *that sends each multiset* $M \in \mathfrak{M}$ *to the family*

$$\left(\textit{(multiplicity of } w \textit{ in the multiset } M)\right)_{w \in S} \in \mathfrak{N}$$

is well-defined and is a bijection.

Example 13.143.3. If $S = \mathbb{N}$, then the map mult defined in Proposition 13.143.2 sends the finite multiset $\{1, 4, 4, 5\}_{\text{multiset}} \in \mathfrak{M}$ to the family $(k_w)_{w \in S} \in \mathbb{N}^S$, where

$$k_1 = 1, \qquad k_4 = 2, \qquad k_5 = 1, \qquad \text{and } k_w = 0 \text{ for all } w \notin \{1, 4, 5\}.$$

Proof of Proposition 13.143.2. It is straightforward to see that the map mult is well-defined. On the other hand, the map

$$\mathfrak{N} \to \mathfrak{M},$$
$$(k_w)_{w \in S} \mapsto (\text{the multiset that contains each } w \in S \text{ with multiplicity } k_w)$$

is also well-defined. These two maps are clearly mutually inverse. Thus, the map mult is invertible, i.e., is a bijection. This proves Proposition 13.143.2. □

Let us now solve Exercise 6.1.29. Let $\mathfrak{A}$ and q be as in Exercise 6.1.29.

For every positive integer n, let $\operatorname{lyn} n$ denote the number of Lyndon words of length n. We need to prove that

$$(13.143.1) \qquad \operatorname{lyn} n = \frac{1}{n} \sum_{d \mid n} \mu(d)\, q^{n/d} \qquad \text{for every positive integer } n.$$

Let $\mathfrak{M}$ denote the set of all finite multisets of Lyndon words. Define two maps $\mathbf{m} : \mathfrak{M} \to \mathfrak{A}^*$ and $\mathbf{n} : \mathfrak{A}^* \to \mathfrak{M}$ as in Proposition 13.143.1. Then, Proposition 13.143.1 shows that these two maps $\mathbf{m}$ and $\mathbf{n}$ are mutually inverse bijections. Hence, the map $\mathbf{m}$ is a bijection.

On the other hand, let $\mathfrak{L}$ be the set of all Lyndon words. Thus, the Lyndon words are precisely the elements of $\mathfrak{L}$. But the definition of $\mathfrak{M}$ says that $\mathfrak{M}$ is the set of all finite multisets of Lyndon words. In other words, $\mathfrak{M}$ is the set of all finite multisets of elements of $\mathfrak{L}$ (since the Lyndon words are precisely the elements of $\mathfrak{L}$).

The definition of $\operatorname{lyn} n$ now rewrites as

$$(13.143.2) \qquad \operatorname{lyn} n = |\{w \in \mathfrak{L} \mid \ell(w) = n\}| \qquad \text{for every positive integer } n$$

For example, if $M = \{2, 2, 3536, 24\}_{\text{multiset}}$, then $(a_1, a_2, \ldots, a_k) = (3536, 24, 2, 2)$.

[967] Here are a few details:

- The equality $\mathbf{m} \circ \mathbf{n} = \operatorname{id}$ follows immediately from the definition of the CFL factorization.
- In order to prove the equality $\mathbf{n} \circ \mathbf{m} = \operatorname{id}$, we fix some $M \in \mathfrak{M}$; our goal is thus to show that $(\mathbf{n} \circ \mathbf{m})(M) = M$. Let $a_1, a_2, \ldots, a_k$ denote the elements of M listed in decreasing order (so that $M = \{a_1, a_2, \ldots, a_k\}_{\text{multiset}}$ and $a_1 \geq a_2 \geq \cdots \geq a_k$). The definition of $\mathbf{m}$ then yields $\mathbf{m}(M) = a_1 a_2 \cdots a_k$. Hence, $a_1, a_2, \ldots, a_k$ are the elements of M, and thus are Lyndon words (since M is a multiset of Lyndon words). Therefore, $(a_1, a_2, \ldots, a_k)$ is a tuple of Lyndon words satisfying $\mathbf{m}(M) = a_1 a_2 \cdots a_k$ and $a_1 \geq a_2 \geq \cdots \geq a_k$. In other words, $(a_1, a_2, \ldots, a_k)$ is a CFL factorization of $\mathbf{m}(M)$ (by the definition of a CFL factorization). Since the CFL factorization of a word is unique, this shows that $(a_1, a_2, \ldots, a_k)$ is **the** CFL factorization of $\mathbf{m}(M)$. Hence, the definition of $\mathbf{n}$ yields $\mathbf{n}(\mathbf{m}(M)) = \{a_1, a_2, \ldots, a_k\}_{\text{multiset}} = M$. In other words, $(\mathbf{n} \circ \mathbf{m})(M) = M$. This concludes the proof of $\mathbf{n} \circ \mathbf{m} = \operatorname{id}$.

Combining the equalities $\mathbf{m} \circ \mathbf{n} = \operatorname{id}$ and $\mathbf{n} \circ \mathbf{m} = \operatorname{id}$, we conclude that the maps $\mathbf{m}$ and $\mathbf{n}$ are mutually inverse, hence bijections.

(since the Lyndon words are precisely the elements of $\mathfrak{L}$).

Let $\mathfrak{N}$ be the set of all families $(k_w)_{w\in\mathfrak{L}} \in \mathbb{N}^{\mathfrak{L}}$ of nonnegative integers (indexed by the Lyndon words) such that all but finitely many $w \in \mathfrak{L}$ satisfy $k_w = 0$. Proposition 13.143.2 (applied to $S = \mathfrak{L}$) shows that the map $\operatorname{mult} : \mathfrak{M} \to \mathfrak{N}$ that sends each multiset $M \in \mathfrak{M}$ to the family

$$((\text{multiplicity of } w \text{ in the multiset } M))_{w\in S} \in \mathfrak{N}$$

is well-defined and is a bijection. Consider this map mult.

The composition $\mathbf{m} \circ \operatorname{mult}^{-1} : \mathfrak{N} \to \mathfrak{A}^*$ of the bijections $\mathbf{m}$ and mult^{-1} is clearly a bijection. It can easily be seen to satisfy

$$\ell\left(\left(\mathbf{m} \circ \operatorname{mult}^{-1}\right)\left((k_w)_{w\in\mathfrak{L}}\right)\right) = \sum_{w\in\mathfrak{L}} k_w \cdot \ell(w) \qquad \text{for every } (k_w)_{w\in\mathfrak{L}} \in \mathfrak{N}. \tag{13.143.3}$$

[968] Now, in the ring $\mathbb{Q}[[t]]$ of formal power series, we have

$$\sum_{w\in\mathfrak{A}^*} t^{\ell(w)} = \sum_{n\in\mathbb{N}} \underbrace{\left|\{w \in \mathfrak{A}^* \mid \ell(w) = n\}\right|}_{\substack{=|\mathfrak{A}^n|=|\mathfrak{A}|^n=q^n \\ (\text{since } |\mathfrak{A}|=q)}} t^n = \sum_{n\in\mathbb{N}} q^n t^n = \frac{1}{1-qt}.$$

[968] *Proof of (13.143.3):* Let $(k_w)_{w\in\mathfrak{L}} \in \mathfrak{N}$. Let $M = \operatorname{mult}^{-1}\left((k_w)_{w\in\mathfrak{L}}\right)$. Then, M is a multiset of elements of $\mathfrak{L}$ and satisfies $(k_w)_{w\in\mathfrak{L}} = \operatorname{mult} M = ((\text{multiplicity of } w \text{ in the multiset } M))_{w\in\mathfrak{L}}$. In other words, every $w \in \mathfrak{L}$ satisfies

$$k_w = (\text{multiplicity of } w \text{ in the multiset } M). \tag{13.143.4}$$

Let $a_1, a_2, \ldots, a_k$ denote the elements of this multiset M listed in decreasing order. Then, the definition of $\mathbf{m}$ yields $\mathbf{m}(M) = a_1 a_2 \cdots a_k$, so that

$$\begin{aligned}
\ell(\mathbf{m}(M)) &= \ell(a_1 a_2 \cdots a_k) = \sum_{i\in\{1,2,\ldots,k\}} \ell(a_i) = \sum_{w\in\mathfrak{L}} \sum_{\substack{i\in\{1,2,\ldots,k\};\\ a_i=w}} \ell\left(\underbrace{a_i}_{\substack{=w \\ (\text{since } a_i=w)}}\right) \qquad (\text{since every } a_i \text{ belongs to } \mathfrak{L}) \\
&= \sum_{w\in\mathfrak{L}} \underbrace{\sum_{\substack{i\in\{1,2,\ldots,k\};\\ a_i=w}} \ell(w)}_{=(\text{number of } i\in\{1,2,\ldots,k\} \text{ satisfying } a_i=w)\cdot\ell(w)} = \sum_{w\in\mathfrak{L}} \underbrace{(\text{number of } i \in \{1,2,\ldots,k\} \text{ satisfying } a_i = w)}_{\substack{=(\text{multiplicity of } w \text{ in the multiset } M)=k_w \\ (\text{by } (13.143.4))}} \cdot \ell(w) \\
&= \sum_{w\in\mathfrak{L}} k_w \cdot \ell(w).
\end{aligned}$$

Now,

$$\ell\left(\left(\mathbf{m} \circ \operatorname{mult}^{-1}\right)\left((k_w)_{w\in\mathfrak{L}}\right)\right) = \ell\left(\mathbf{m}\left(\underbrace{\operatorname{mult}^{-1}\left((k_w)_{w\in\mathfrak{L}}\right)}_{=M}\right)\right) = \ell(\mathbf{m}(M)) = \sum_{w\in\mathfrak{L}} k_w \cdot \ell(w),$$

which proves (13.143.3).

Hence,

$$\begin{aligned}
\frac{1}{1-qt} &= \sum_{w\in\mathfrak{A}^*} t^{\ell(w)} = \sum_{(k_w)_{w\in\mathfrak{L}}\in\mathfrak{N}} t^{\ell\left(\left(\mathbf{m}\circ\operatorname{mult}^{-1}\right)\left((k_w)_{w\in\mathfrak{L}}\right)\right)} \\
&\qquad \left(\begin{array}{c}\text{here, we substituted } \left(\mathbf{m}\circ\operatorname{mult}^{-1}\right)\left((k_w)_{w\in\mathfrak{L}}\right) \text{ for } w \text{ in the sum,}\\ \text{since the map } \mathbf{m}\circ\operatorname{mult}^{-1} : \mathfrak{N}\to\mathfrak{A}^* \text{ is a bijection}\end{array}\right) \\
&= \sum_{(k_w)_{w\in\mathfrak{L}}\in\mathfrak{N}} t^{\sum_{w\in\mathfrak{L}} k_w\cdot\ell(w)} \qquad \text{(by (13.143.3))} \\
&= \sum_{(k_w)_{w\in\mathfrak{L}}\in\mathfrak{N}} \prod_{w\in\mathfrak{L}} t^{k_w\cdot\ell(w)} = \prod_{w\in\mathfrak{L}} \underbrace{\sum_{k\in\mathbb{N}} t^{k\cdot\ell(w)}}_{=\dfrac{1}{1-t^{\ell(w)}}} \qquad \text{(by the product rule)} \\
&= \prod_{w\in\mathfrak{L}} \frac{1}{1-t^{\ell(w)}} = \prod_{n\geq 1} \prod_{\substack{w\in\mathfrak{L};\\ \ell(w)=n}} \underbrace{\frac{1}{1-t^{\ell(w)}}}_{\substack{=\dfrac{1}{1-t^n} \\ \text{(since } \ell(w)=n)}} \qquad \text{(since } \ell(w)\geq 1 \text{ for every } w\in\mathfrak{L}) \\
&= \prod_{n\geq 1} \underbrace{\prod_{\substack{w\in\mathfrak{L};\\ \ell(w)=n}} \frac{1}{1-t^n}}_{\substack{=\left(\dfrac{1}{1-t^n}\right)^{\operatorname{lyn} n} \\ \text{(by (13.143.2))}}} = \prod_{n\geq 1} \left(\frac{1}{1-t^n}\right)^{\operatorname{lyn} n}.
\end{aligned}$$

Taking the logarithm of both sides of this identity, we obtain

$$\begin{aligned}
\log\frac{1}{1-qt} &= \log\left(\prod_{n\geq 1}\left(\frac{1}{1-t^n}\right)^{\operatorname{lyn} n}\right) = \sum_{n\geq 1} (\operatorname{lyn} n)\cdot \underbrace{\log\left(\frac{1}{1-t^n}\right)}_{\substack{=-\log(1-t^n)=\sum_{u\geq 1}\frac{1}{u}(t^n)^u \\ \text{(by the Mercator series for the logarithm)}}} \\
&= \sum_{n\geq 1} (\operatorname{lyn} n)\cdot\sum_{u\geq 1}\frac{1}{u}(t^n)^u = \sum_{n\geq 1}\sum_{u\geq 1}(\operatorname{lyn} n)\frac{1}{u}\underbrace{(t^n)^u}_{=t^{nu}} = \sum_{n\geq 1}\sum_{u\geq 1}(\operatorname{lyn} n)\frac{1}{u}t^{nu} \\
&= \underbrace{\sum_{n\geq 1}\sum_{\substack{v\geq 1;\\ n\mid v}}}_{=\sum_{v\geq 1}\sum_{n\mid v}} (\operatorname{lyn} n)\underbrace{\frac{1}{v/n}}_{=\dfrac{n}{v}} t^v \qquad \text{(here, we substituted } v/n \text{ for } u \text{ in the second sum)} \\
&= \sum_{v\geq 1}\sum_{n\mid v}(\operatorname{lyn} n)\frac{n}{v}t^v = \sum_{n\geq 1}\sum_{d\mid n}(\operatorname{lyn} d)\frac{d}{n}t^n
\end{aligned}$$

(here, we renamed the summation indices v and n as n and d). Since

$$\begin{aligned}
\log\frac{1}{1-qt} &= -\log(1-qt) = \sum_{n\geq 1}\frac{1}{n}(qt)^n \qquad \text{(by the Mercator series for the logarithm)} \\
&= \sum_{n\geq 1}\frac{1}{n}q^n t^n,
\end{aligned}$$

this rewrites as

$$\sum_{n\geq 1}\frac{1}{n}q^n t^n = \sum_{n\geq 1}\sum_{d\mid n}(\operatorname{lyn} d)\frac{d}{n}t^n.$$

Comparing coefficients, we conclude that every positive integer n satisfies

$$\frac{1}{n}q^n = \sum_{d\mid n} (\operatorname{lyn} d) \frac{d}{n}.$$

Multiplying this with n, we obtain

$$q^n = \sum_{d\mid n} (\operatorname{lyn} d) \, d. \tag{13.143.5}$$

Now, recall that every positive integer N satisfies

$$\sum_{d\mid N} \mu(d) = \delta_{N,1}. \tag{13.143.6}$$

[969] Now, every positive integer n satisfies

$$\begin{aligned}\sum_{d\mid n} \mu(d)\, q^{n/d} &= \sum_{e\mid n} \mu(e) \underbrace{q^{n/e}}_{\substack{=\sum_{d\mid n/e}(\operatorname{lyn} d)d \\ \text{(by (13.143.5), applied} \\ \text{to } n/e \text{ instead of } n)}} = \sum_{e\mid n} \mu(e) \sum_{d\mid n/e} (\operatorname{lyn} d)\, d \\ &= \underbrace{\sum_{e\mid n} \sum_{d\mid n/e}}_{=\sum_{d\mid n}\sum_{e\mid n/d}} \mu(e)\, (\operatorname{lyn} d)\, d = \sum_{d\mid n} \underbrace{\sum_{e\mid n/d} \mu(e)}_{\substack{=\delta_{n/d,1} \\ \text{(by (13.143.6), applied} \\ \text{to } N=n/d)}} (\operatorname{lyn} d)\, d \\ &= \sum_{d\mid n} \underbrace{\delta_{n/d,1}}_{=\delta_{n,d}} (\operatorname{lyn} d)\, d = \sum_{d\mid n} \delta_{n,d} (\operatorname{lyn} d)\, d = (\operatorname{lyn} n)\, n.\end{aligned}$$

Dividing this by n, we obtain $\frac{1}{n}\sum_{d\mid n} \mu(d)\, q^{n/d} = \operatorname{lyn} n$. This proves (13.143.1). Thus, Exercise 6.1.29 is solved.

13.144. **Solution to Exercise 6.1.31.** *Solution to Exercise 6.1.31.* The word v is nonempty (since it is Lyndon). Thus, $u \neq w$ [970]. Combined with $u \leq uv = w$, this yields $u < w$.

(b) We have $w = uv$; thus, v is a proper suffix of w (since u is nonempty). Also, the word w is Lyndon, and therefore

$$\text{every nonempty proper suffix of } w \text{ is } > w \tag{13.144.1}$$

(by the definition of a Lyndon word). Applying (13.144.1) to the nonempty proper suffix v of w, we obtain $v > w$. Thus, $w < v$. Hence, $u < w < v$, and this solves Exercise 6.1.31(b).

(a) The word v is Lyndon (by its definition). It remains to prove that the word u is Lyndon.

Recall that a word x is Lyndon if and only if it is nonempty and satisfies the property that every nonempty proper suffix y of x satisfies $y > x$. [971] Applied to $x = u$, this shows that the word u is Lyndon if and only if it is nonempty and satisfies the property that every nonempty proper suffix y of u satisfies $y > u$. Let us now prove that

$$\text{every nonempty proper suffix } y \text{ of } u \text{ satisfies } y > u. \tag{13.144.2}$$

Proof of (13.144.2): Assume the contrary. Then, there exists a nonempty proper suffix y of u which does not satisfy $y > u$. Let p be the **shortest** such suffix. Thus, p is a nonempty proper suffix of u which does not satisfy $p > u$.

Notice that $p \neq u$ (since p is a **proper** suffix of u).

[969] This is one of the most fundamental properties of the number-theoretic Möbius function. For a proof of (13.143.6), see the solution of Exercise 2.9.6. (More precisely, the equality (13.143.6) is obtained from (13.84.3) by renaming n as N.)

[970] *Proof.* Assume the contrary. Then, $u = w$. Hence, $u\varnothing = w = uv$. Cancelling u from this equality, we obtain $\varnothing = v$. Thus, the word v is empty; this contradicts the fact that v is nonempty. This contradiction proves that our assumption was wrong, qed.

[971] This is just a restatement of the definition of a Lyndon word.

There exists a nonempty $t \in \mathfrak{A}^*$ such that $u = tp$ (since p is a proper suffix of u). Consider this t. We have $w = \underbrace{u}_{=tp} v = tpv = t(pv)$, and thus the word pv is a proper suffix of w (since t is nonempty). Also, the word pv is nonempty (since v is nonempty). Hence, pv is a nonempty proper suffix of w. But applying (13.144.1) to the nonempty proper suffix pv of w, we obtain $pv > w = uv$. Thus, $uv < pv$. Hence, Proposition 6.1.2(e) (applied to $a = u$, $b = v$, $c = p$ and $d = v$) yields that either we have $u \leq p$ or the word p is a prefix of u. Since we don't have $u \leq p$ [972], we therefore conclude that the word p is a prefix of u. Hence, $p \leq u$. Combined with $p \neq u$, this yields $p < u$. Combined with $u < v$ (by Exercise 6.1.31(b)), this yields $p < u < v$.

Let us now show that the word p is Lyndon. Indeed, let r be a nonempty proper suffix of p. Then, r is shorter than p (being a proper suffix of p); in other words, $\ell(r) < \ell(p)$. On the other hand, there exists some nonempty $q \in \mathfrak{A}^*$ satisfying $p = qr$ (since r is a proper suffix of p). Consider this q. We have $u = t \underbrace{p}_{=qr} = tqr = (tq)r$. Thus, r is a proper suffix of u (since tq is nonempty (because q is nonempty)). Hence, $r > u$ [973]. Hence, $r > u > p$ (since $p < u$). Now, let us forget that we fixed r. We thus have shown that

$$\text{every nonempty proper suffix } r \text{ of } p \text{ satisfies } r > p.$$

But recall that the word p is Lyndon if and only if it is nonempty and satisfies the property that every nonempty proper suffix r of p satisfies $r > p$. [974] Hence, we conclude that the word p is Lyndon (since we have shown that p is nonempty and satisfies the property that every nonempty proper suffix r of p satisfies $r > p$).

Now, the words p and v are Lyndon and satisfy $p < v$. Thus, Proposition 6.1.16(a) (applied to p instead of u) yields that the word pv is Lyndon. Notice also that $w = \underbrace{u}_{=tp} v = tpv = t(pv)$, and thus pv is a proper suffix of w (since t is nonempty). Hence, pv is a proper suffix of w such that pv is Lyndon. In other words, pv is a proper suffix z of w such that z is Lyndon. Since v is the longest such suffix[975], this yields that pv is not longer than v. In other words, $\ell(pv) \leq \ell(v)$. This contradicts $\ell(pv) = \underbrace{\ell(p)}_{\substack{>0 \\ \text{(since } p \text{ is nonempty)}}} + \ell(v) > \ell(v)$. This contradiction proves that our assumption was wrong. Hence, (13.144.2) is proven.

Now, recall that the word u is Lyndon if and only if it is nonempty and satisfies the property that every nonempty proper suffix y of u satisfies $y > u$. Hence, the word u is Lyndon (because we have shown that u is nonempty and satisfies the property that every nonempty proper suffix y of u satisfies $y > u$). The solution of Exercise 6.1.31(a) is thus complete.

(c) Let us denote by u' and v' the words u and v constructed in Theorem 6.1.30 (to avoid confusing them with the words u and v defined in Exercise 6.1.31). We must then show that $u = u'$ and $v = v'$.

From Theorem 6.1.30, we see that v' is the (lexicographically) smallest nonempty **proper** suffix of w, and that u' is a nonempty word such that $w = u'v'$.

The word v is a nonempty proper suffix of w. Since v' is the smallest such suffix, we thus conclude that $v' \leq v$. We shall now prove that $v' = v$.

Indeed, v is a proper suffix of w. In other words, v is a proper suffix of $u'v'$ (since $w = u'v'$). Hence, we must be in one of the following two cases (depending on whether this suffix begins before the suffix v' of $u'v'$ begins or afterwards):

Case 1: The word v is a proper suffix of v'.

Case 2: The word v has the form qv' where q is a proper suffix of u'. (This suffix q may be empty.)

[972] *Proof.* Assume the contrary. Then, $u \leq p$. Thus, $p \geq u$. Combined with $p \neq u$, this yields $p > u$. This contradicts the fact that p does not satisfy $p > u$. This contradiction proves that our assumption was wrong, qed.

[973] *Proof.* Assume the contrary. Then, we don't have $r > u$. Thus, r is a nonempty proper suffix of u which does not satisfy $r > u$. In other words, r is a nonempty proper suffix y of u which does not satisfy $y > u$. Since the shortest such suffix is p (by the definition of p), this yields that r is not shorter than p. That is, $\ell(r) \geq \ell(p)$. This contradicts $\ell(r) < \ell(p)$. This contradiction proves that our assumption was wrong, qed.

[974] This is just a restatement of the definition of a Lyndon word (applied to p).

[975] *Proof.* We defined v as the longest proper suffix of w such that v is Lyndon. In other words, v is the longest proper suffix z of w such that z is Lyndon, qed.

Let us first consider Case 1. In this case, the word v is a proper suffix of v'. Now, every nonempty proper suffix s of v' satisfies $s > v'$ [976]. Thus, the word v' is nonempty and satisfies the property that every nonempty proper suffix s of v' satisfies $s > v'$. In other words, the word v' is Lyndon (according to the definition of a Lyndon word).

Now, v is the longest proper suffix of w such that v is Lyndon. In other words, v is the longest proper suffix z of w such that z is Lyndon. But v' also is a proper suffix z of w such that z is Lyndon (since v' is Lyndon). Since v is the longest such suffix, this yields that v' is not longer than v. In other words, $\ell(v') \leq \ell(v)$. But since v is a proper suffix of v', we have $\ell(v) < \ell(v') \leq \ell(v)$. This is absurd. Hence, $v' = v$ (since ex falso quodlibet). Thus, $v' = v$ is proven in Case 1.

Let us now consider Case 2. In this case, the word v has the form qv' where q is a proper suffix of u'. Consider this q.

We need to prove that $v' = v$. If $q = \varnothing$, then this is obvious (because if $q = \varnothing$, then $v = \underbrace{q}_{=\varnothing} v' = v'$ and thus $v' = v$). Hence, for the rest of this proof, we can WLOG assume that we don't have $q = \varnothing$. Assume this. The word q is nonempty (since we don't have $q = \varnothing$), and thus v' is a proper suffix of v (since $v = qv'$).

The word v is Lyndon. Hence, every nonempty proper suffix of v is $> v$ (by the definition of a Lyndon word). Applying this to the nonempty proper suffix v' of v, we conclude that $v' > v$. This contradicts $v' \leq v$. This contradiction shows that we have $v' = v$ (since ex falso quodlibet). Thus, $v' = v$ is proven in Case 2.

Now, $v' = v$ is proven in each of the two Cases 1 and 2. Since these two Cases cover all possibilities, this yields that $v' = v$ always holds.

Now we have proven that $v' = v$. Cancelling v from the equality $uv = w = u' \underbrace{v'}_{=v} = u'v$, we obtain $u = u'$. Thus, we have $u = u'$ and $v = v'$. This solves Exercise 6.1.31(c).

[*Remark:* We have not used any statement of Theorem 6.1.30 in our above solution. Thus, parts (a) and (b) of Exercise 6.1.31 provide an alternative proof of Theorem 6.1.30 (because Exercise 6.1.31(c) shows that the words u and v defined in Exercise 6.1.31 are precisely the words u and v constructed in Theorem 6.1.30).]

13.145. **Solution to Exercise 6.1.32.** *Solution to Exercise 6.1.32.* (a) Let us prove the implications $\mathcal{A}' \Longrightarrow \mathcal{D}'$ and $\mathcal{D}' \Longrightarrow \mathcal{A}'$.

Proof of the implication $\mathcal{A}' \Longrightarrow \mathcal{D}'$: Assume that Assertion $\mathcal{A}'$ holds. Thus, w is a power of a Lyndon word. Let said Lyndon word be t. Thus, $w = t^n$ for some $n \in \mathbb{N}$. Consider this n. Since w is nonempty, we have $n \geq 1$. Hence, t^{n-1} is well-defined.

We will first show the following simple lemma:

Lemma A: Let u', v' and p be three words, and $N \in \mathbb{N}$ be such that $u'v' = p^N$. Assume that p is not a prefix of u', and that p is not a suffix of v'. Then, $N \leq 1$.

Proof of Lemma A: Assume the contrary. Thus, $N > 1$, so that $N \geq 2$. The word v' is a suffix of $p^{N-1}p$ (since $u'v' = p^N = p^{N-1}p$). Thus, we must be in one of the following two cases (depending on whether this suffix begins before the suffix p of $p^{N-1}p$ begins, or afterwards):

Case 1: The word v' has the form rp where r is a suffix of p^{N-1}.

Case 2: The word v' is a suffix of p.

Let us first consider Case 1. In this case, the word v' has the form rp where r is a suffix of p^{N-1}. Hence, p is a suffix of v', contradicting the fact that p is not a suffix of v'. Hence, Case 1 leads to a contradiction.

Let us now consider Case 2. In this case, the word v' is a suffix of p. In other words, there exists a $q \in \mathfrak{A}^*$ such that $p = qv'$. Consider this q. We have $u'v' = p^N = p^{N-1} \underbrace{p}_{=qv'} = p^{N-1}qv'$. Cancelling v' from this equality, we obtain $u' = p^{N-1}q$. Since $p^{N-1} = pp^{N-2}$ (this is well-defined because $N \geq 2$), this further becomes $u' = \underbrace{p^{N-1}}_{=pp^{N-2}} q = pp^{N-2}q = p\left(p^{N-2}q\right)$. Hence, p is a prefix of u', contradicting the fact that p is not a prefix of u'. We have thus obtained a contradiction in Case 2.

[976] *Proof.* Let s be a nonempty proper suffix of v'. The word s is a proper suffix of v', which (in turn) is a suffix of w. Hence, s is a proper suffix of w. Thus, s is a nonempty proper suffix of w. Since the smallest such suffix is v' (by the definition of v'), this yields $s \geq v'$. Since $s \neq v'$ (because s is a **proper** suffix of v'), this yields $s > v'$, qed.

We have thus obtained a contradiction in each of the two Cases 1 and 2. Since these two Cases cover all possibilities, this shows that we always get a contradiction. This completes the proof of Lemma A.

Now, let us return to the proof of the implication $\mathcal{A}' \Longrightarrow \mathcal{D}'$. Let u and v be nonempty words satisfying $w = uv$. We want to show that $vu \geq uv$.

The word t is nonempty (since it is Lyndon); thus, $\ell(t) \geq 1$. Hence, for every sufficiently large $a \in \mathbb{N}$, we have that the word t^a is not a prefix of u [977]. Similarly, for every sufficiently large $b \in \mathbb{N}$, we have that the word t^b is not a suffix of v.

Let A be the largest nonnegative integer a such that t^a is a prefix of u [978]. Thus, t^A is a prefix of u, but t^{A+1} is not a prefix of u. Let B be the largest nonnegative integer b such that t^b is a suffix of v [979]. Thus, t^B is a suffix of v, but t^{B+1} is not a suffix of v.

There exists a $u' \in \mathfrak{A}^*$ such that $u = t^A u'$ (since t^A is a prefix of u). Consider this u'. Recall that t^{A+1} is not a prefix of u. In other words, $t^A t$ is not a prefix of $t^A u'$ (since $t^{A+1} = t^A t$ and $u = t^A u'$).

There exists a $v' \in \mathfrak{A}^*$ such that $v = v' t^B$ (since t^B is a suffix of v). Consider this v'. Recall that t^{B+1} is not a suffix of v. In other words, tt^B is not a suffix of $v' t^B$ (since $t^{B+1} = tt^B$ and $v = v' t^B$).

If the word t was a prefix of u', then $t^A t$ would be a prefix of $t^A u'$, which would contradict the fact that $t^A t$ is not a prefix of $t^A u'$. Thus, t is not a prefix of u'. In particular, $t \neq u'$.

If the word t was a suffix of v', then tt^B would be a suffix of $v' t^B$, which would contradict the fact that tt^B is not a suffix of $v' t^B$. Thus, t is not a suffix of v'. In particular, $t \neq v'$.

But we have $t^n = w = \underbrace{u}_{=t^A u'} \underbrace{v}_{=v' t^B} = t^A u' v' t^B = t^A \left(u' v' t^B\right)$. Hence, t^A is a prefix of t^n; therefore, $A \leq n$ and thus $n - A \geq 0$. Hence, $t^A t^{n-A} = t^n = t^A \left(u' v' t^B\right)$. Cancelling t^A in this equality, we obtain $t^{n-A} = u' v' t^B = (u' v') t^B$, which shows that t^B is a suffix of t^{n-A}. Thus, $B \leq n - A$ and therefore $n - A - B \geq 0$. Hence, $t^{n-A-B} t^B = t^{n-A} = (u' v') t^B$. Cancelling t^B in this yields $t^{n-A-B} = u' v'$. Now, Lemma A (applied to $p = t$ and $N = n - A - B$) yields $n - A - B \leq 1$. Since $n - A - B \geq 0$, this shows that we have either $n - A - B = 0$ or $n - A - B = 1$. We thus must be in one of the following two cases:

Case 1: We have $n - A - B = 0$.

Case 2: We have $n - A - B = 1$.

Let us first consider Case 1. In this case, we have $n - A - B = 0$. Thus, $t^{n-A-B} = u' v'$ rewrites as $t^0 = u' v'$, so that $u' v' = t^0 = \varnothing$. Consequently, $u' = \varnothing$ and $v' = \varnothing$. Now, $u = t^A \underbrace{u'}_{=\varnothing} = t^A$ and $v = \underbrace{v'}_{=\varnothing} t^B = t^B$, so that $\underbrace{v}_{=t^B} \underbrace{u}_{=t^A} = t^B t^A = t^{B+A} = t^{A+B} = \underbrace{t^A}_{=u} \underbrace{t^B}_{=v} = uv \geq uv$. Hence, $vu \geq uv$ is proven in Case 1.

Let us now consider Case 2. In this case, we have $n - A - B = 1$. Thus, $t^{n-A-B} = u' v'$ rewrites as $t^1 = u' v'$, so that $u' v' = t^1 = t$. As a consequence, u' is nonempty (because if u' was empty, then we would have $u' v' = v'$ and thus $v' = u' v' = t$, contradicting $t \neq v'$), so that $\ell(u') > 0$. Now, $\ell\left(\underbrace{t}_{=u'v'}\right) = \underbrace{\ell(u')}_{>0} + \ell(v') > \ell(v')$, and therefore t is not a prefix of v'. Also, v' is nonempty (because if v' was empty, then we would have $u' v' = u'$ and thus $u' = u' v' = t$, contradicting $t \neq u'$). But t is Lyndon, and therefore Proposition 6.1.14(a) (applied to t, u' and v' instead of w, u and v) yields $v' \geq t$. That is, $t \leq v'$. Thus, Proposition 6.1.2(d) (applied to $a = t$, $b = t^{n-1}$, $c = v'$ and $d = t^B u$) yields that either we have $tt^{n-1} \leq v' t^B u$ or the word t is a prefix of

[977] *Proof.* Let $a \in \mathbb{N}$ be such that $a > \ell(u)$. Then, $\ell(t^a) = \underbrace{a}_{>\ell(u)} \underbrace{\ell(t)}_{\geq 1} > \ell(u)$, so that the word t^a is longer than u. Hence, the word t^a is not a prefix of u.

Now, let us forget that we have fixed a. We thus have shown that for every $a \in \mathbb{N}$ satisfying $a > \ell(u)$, we have that the word t^a is not a prefix of u. Consequently, for every sufficiently large $a \in \mathbb{N}$, we have that the word t^a is not a prefix of u, qed.

[978] This is well-defined because of the following two facts:

- There exists an $a \in \mathbb{N}$ such that the word t^a is a prefix of u (namely, $a = 0$).
- For every sufficiently large $a \in \mathbb{N}$, we have that the word t^a is not a prefix of u.

[979] This is well-defined because of the following two facts:

- There exists a $b \in \mathbb{N}$ such that the word t^b is a suffix of v (namely, $b = 0$).
- For every sufficiently large $b \in \mathbb{N}$, we have that the word t^b is not a suffix of v.

v'. Since t is not a prefix of v', this yields $tt^{n-1} \leq v't^B u$. Since $tt^{n-1} = t^n = w = uv$ and $\underbrace{v't^B}_{=v} u = vu$, this rewrites as $uv \leq vu$. Hence, $vu \geq uv$ is proven in Case 2.

We thus have shown that $vu \geq uv$ in each of the two Cases 1 and 2. Since these two Cases cover all possibilities, this yields that we always have $vu \geq uv$. Thus, Assertion $\mathcal{D}'$ is satisfied, so we have proven the implication $\mathcal{A}' \Longrightarrow \mathcal{D}'$.

Proof of the implication $\mathcal{D}' \Longrightarrow \mathcal{A}'$: Assume that Assertion $\mathcal{D}'$ holds. Thus, if u and v are nonempty words satisfying $w = uv$, then we have $vu \geq uv$.

We need to prove that Assertion $\mathcal{A}'$ holds, i.e., that w is a power of a Lyndon word. Assume the contrary. Thus, w is not a power of a Lyndon word; hence, w is not a Lyndon word itself. Consequently, there exist nonempty words u and v such that $w = uv$ and $vu \leq uv$ [980]. Consider such a pair of nonempty words u and v with **minimum** $\ell(u)$. The minimality of $\ell(u)$ shows that

$$\text{(13.145.1)} \qquad (\text{if } u' \text{ and } v' \text{ are nonempty words such that } w = u'v' \text{ and } v'u' \leq u'v', \text{ then } \ell(u') \geq \ell(u)).$$

We have $vu \leq uv$. In combination with $vu \geq uv$ (which follows from Assertion $\mathcal{D}'$), this yields $vu = uv$. Therefore, Proposition 6.1.4 yields that there exist a $t \in \mathfrak{A}^*$ and two nonnegative integers n and m such that $u = t^n$ and $v = t^m$. Consider this t and these n and m. We have $n \neq 0$ (since $t^n = u$ is nonempty) and $m \neq 0$ (since $t^m = v$ is nonempty), and the word t is nonempty (since $t^n = u$ is nonempty). Moreover, we have $n = 1$ [981]. Hence, $t^n = t^1 = t$, so that $u = t^n = t$ and $w = \underbrace{u}_{=t}\underbrace{v}_{=t^m} = tt^m = t^{m+1}$. We shall now prove that the word t is Lyndon.

Assume the contrary. Then, t is not Lyndon. Let q be the (lexicographically) smallest nonempty suffix of t. Then, Proposition 6.1.19(b) (applied to t and q instead of w and v) yields that there exists a nonempty $p \in \mathfrak{A}^*$ such that $t = pq$, $p \geq q$ and $pq \geq qp$ [982]. Consider this p. Since q is nonempty, we have $\ell(pq) > \ell(p)$, so that $\ell\left(\underbrace{u}_{=t=pq}\right) = \ell(pq) > \ell(p)$. From $pq \geq qp$, we obtain $qp \leq pq$. Thus, it is easy to see that $(qp)^{m+1} \leq (pq)^{m+1}$ [983].

We have $v = t^m$. Since $t = pq$, this rewrites as $v = (pq)^m$. Now,

$$q \underbrace{v}_{=(pq)^m} p = \underbrace{q(pq)^m}_{=(qp)^m q} p = (qp)^m qp = (qp)^{m+1} \leq (pq)^{m+1} = (pq)\left(\underbrace{pq}_{=t}\right)^m = (pq)\underbrace{t^m}_{=v} = (pq)v = pqv.$$

[980] *Proof.* Assume the contrary. Hence, any nonempty words u and v satisfying $w = uv$ satisfy $vu > uv$. In other words, w satisfies Assertion $\mathcal{D}$ of Theorem 6.1.20. Hence, w also satisfies Assertion $\mathcal{A}$ of Theorem 6.1.20 (since Theorem 6.1.20 yields the equivalence of these two assertions $\mathcal{D}$ and $\mathcal{A}$). In other words, w is a Lyndon word. This contradicts the knowledge that w is not a Lyndon word, qed.

[981] *Proof.* Assume the contrary. Hence, $n \neq 1$. Combined with $n \neq 0$, this leads to $n \geq 2$. As a consequence, $u = t^n$ can be rewritten as $u = tt^{n-1}$. The word t^{n-1+m} is nonempty (since t is nonempty and since $\underbrace{n}_{\geq 2} - 1 + \underbrace{m}_{\geq 0} \geq 2 - 1 + 0 = 1$). Now, $w = \underbrace{u}_{=tt^{n-1}}\underbrace{v}_{=t^m} = tt^{n-1}t^m = tt^{n-1+m}$. Also, $t^{n-1+m}t = t^{n-1+m+1} = tt^{n-1+m}$. Hence, (13.145.1) (applied to $u' = t$ and $v' = t^{n-1+m}$) yields $\ell(t) \geq \ell(u)$ (since t and t^{n-1+m} are nonempty). Since $u = t^n$, this rewrites as $\ell(t) \geq \ell(t^n) = \underbrace{n}_{\geq 2}\ell(t) \geq 2\ell(t)$, whence $\ell(t) = 0$, which contradicts the fact that t is nonempty. This contradiction shows that our assumption was wrong, qed.

[982] Notice that the variable p here is what has been called u in Proposition 6.1.19(b). (We had to rename it since the letter u is already in use.)

[983] *Proof.* We have $qp \leq pq$. Hence, Proposition 6.1.2(d) (applied to qp, $(qp)^m$, pq and $(pq)^m$ instead of a, b, c and d) yields that either we have $(qp)(qp)^m \leq (pq)(pq)^m$ or the word qp is a prefix of pq. In the first of these two cases, we are done (because in the first of these two cases, we have $(qp)(qp)^m \leq (pq)(pq)^m$ and thus $(qp)^{m+1} = (qp)(qp)^m \leq (pq)(pq)^m = (pq)^{m+1}$). Hence, we can WLOG assume that we are in the second of these two cases. Assume this. Then, the word qp is a prefix of pq. Since the word qp has the same length as pq (in fact, $\ell(qp) = \ell(q) + \ell(p) = \ell(p) + \ell(q) = \ell(pq)$), this yields that $qp = pq$, so that $\left(\underbrace{qp}_{=pq}\right)^{m+1} = (pq)^{m+1} \leq (pq)^{m+1}$, qed.

Since $w = \underbrace{u}_{=t=pq} v = pqv$, we can thus apply (13.145.1) to $u' = p$ and $v' = qv$. As a result, we obtain $\ell(p) \geq \ell(u)$. This contradicts $\ell(u) > \ell(p)$. This contradiction shows that our assumption (that t is not Lyndon) was wrong. Hence, t is Lyndon. Thus, w is a power of a Lyndon word (since $w = t^{m+1}$ is a power of t). Thus, Assertion $\mathcal{A}'$ is satisfied, so we have proven the implication $\mathcal{D}' \Longrightarrow \mathcal{A}'$.

Now we have proven both implications $\mathcal{A}' \Longrightarrow \mathcal{D}'$ and $\mathcal{D}' \Longrightarrow \mathcal{A}'$. Therefore, the equivalence $\mathcal{A}' \Longleftrightarrow \mathcal{D}'$ follows. Thus, Exercise 6.1.32(a) is solved.

(b) Consider the letter m and the alphabet $\mathfrak{A} \cup \{m\}$ defined in Assertion $\mathcal{F}''$. We notice that the lexicographic order on $\mathfrak{A}^*$ is the restriction of the lexicographic order on $(\mathfrak{A} \cup \{m\})^*$ to $\mathfrak{A}^*$. Therefore, when we have two words p and q in $\mathfrak{A}^*$, statements like "$p < q$" do not depend on whether we are regarding p and q as elements of $\mathfrak{A}^*$ or as elements of $(\mathfrak{A} \cup \{m\})^*$. It is easy to see that the one-letter word m satisfies

$$m > p \qquad \text{for every } p \in \mathfrak{A}^*. \tag{13.145.2}$$

[984]

We shall prove the implications $\mathcal{B}' \Longrightarrow \mathcal{E}'$, $\mathcal{C}' \Longrightarrow \mathcal{E}'$, $\mathcal{G}' \Longrightarrow \mathcal{H}'$, $\mathcal{E}' \Longrightarrow \mathcal{F}''$, $\mathcal{F}'' \Longrightarrow \mathcal{B}'$, $\mathcal{F}' \Longrightarrow \mathcal{C}'$, $\mathcal{B}' \Longrightarrow \mathcal{G}'$ and $\mathcal{H}' \Longrightarrow \mathcal{B}'$.

First of all, the implication $\mathcal{B}' \Longrightarrow \mathcal{E}'$ holds for obvious reasons (in fact, if two words u and v satisfying $w = uv$ satisfy $v \geq w$, then $v \geq u$ (because $v \geq w = uv \geq u$)). Also, the implication $\mathcal{C}' \Longrightarrow \mathcal{E}'$ holds for obvious reasons (in fact, if two words u and v satisfying $w = uv$ satisfy (v is a prefix of u), then (v is a prefix of w) (because v is a prefix of u, and u in turn is a prefix of $uv = w$)). The implication $\mathcal{G}' \Longrightarrow \mathcal{H}'$ is also trivially true.

Proof of the implication $\mathcal{E}' \Longrightarrow \mathcal{F}''$: Assume that Assertion $\mathcal{E}'$ holds.

Assume (for the sake of contradiction) that the word $wm \in (\mathfrak{A} \cup \{m\})^*$ is not a Lyndon word. Clearly, this word wm is nonempty. Let v' be the (lexicographically) smallest nonempty suffix of this word $wm \in (\mathfrak{A} \cup \{m\})^*$. Since wm is not a Lyndon word, we can apply Proposition 6.1.19(b) to $\mathfrak{A} \cup \{m\}$, wm and v' instead of $\mathfrak{A}$, w and v. As a result, we conclude that there exists a nonempty $u \in (\mathfrak{A} \cup \{m\})^*$ such that $wm = uv'$, $u \geq v'$ and $uv' \geq v'u$. Consider this u.

We know that v' is a suffix of wm. Thus, we must be in one of the following two cases (depending on whether this suffix begins before the suffix m of wm begins or afterwards):

Case 1: The word v' is a nonempty suffix of m. (Note that $v' = m$ is allowed.)

Case 2: The word v' has the form vm where v is a nonempty suffix of w.

Let us consider Case 1 first. In this case, the word v' is a nonempty suffix of m. Since the only nonempty suffix of m is m itself (because m is a one-letter word), this yields $v' = m$. Now, $wm = u \underbrace{v'}_{=m} = um$. Cancelling m from this equality, we obtain $w = u$, so that $u = w \in \mathfrak{A}^*$. Hence, $m > u$ (by (13.145.2), applied to $p = u$). This contradicts $u \geq v' = m$. Thus, we have obtained a contradiction in Case 1.

Let us now consider Case 2. In this case, the word v' has the form vm where v is a nonempty suffix of w. Consider this v. We have $wm = u \underbrace{v'}_{=vm} = uvm$. By cancelling m from this equality, we obtain $w = uv$. Thus, u and v are subwords of w, and therefore belong to $\mathfrak{A}^*$ (since $w \in \mathfrak{A}^*$). Moreover, u and v are nonempty. Hence, Assertion $\mathcal{E}'$ yields that either we have $v \geq u$ or the word v is a prefix of w. Since we cannot have $v \geq u$ (because if we had $v \geq u$, then we would have $v' = vm > v \geq u \geq v'$, which is absurd), we therefore must have that v is a prefix of w. In other words, there exists a $q \in \mathfrak{A}^*$ such that $w = vq$. Consider this q. We have $m > q$ (by (13.145.2), applied to $p = q$). Thus, $q \leq m$. Hence, Proposition 6.1.2(b) (applied to $\mathfrak{A} \cup \{m\}$, v, q and m instead of $\mathfrak{A}$, a, c and d) yields $vq \leq vm$. Therefore, $vm \geq vq = w$ (since $w = vq$), so that $v' = vm \geq w = uv > u$ (since v is nonempty). This contradicts $u \geq v'$. Thus, we have found a contradiction in Case 2.

[984] *Proof of (13.145.2):* Let $p \in \mathfrak{A}^*$. If p is empty, then (13.145.2) is obviously satisfied. Hence, for the rest of the proof of (13.145.2), we WLOG assume that p is nonempty. Then, p has a first letter, and we have (the first letter of the word p) $\in \mathfrak{A}$ (since $p \in \mathfrak{A}^*$). Thus,

$$\begin{aligned}(\text{the first letter of the word } p) &< m \qquad (\text{since } a < m \text{ for every } a \in \mathfrak{A}) \\ &= (\text{the first letter of the word } m)\,.\end{aligned}$$

Hence, $p < m$ (by the definition of the lexicographic order), that is, $m > p$. This proves (13.145.2).

We have therefore obtained a contradiction in each of the two Cases 1 and 2. Since these two Cases cover all possibilities, this shows that we always get a contradiction. Hence, our assumption (that the word $wm \in (\mathfrak{A} \cup \{m\})^*$ is not a Lyndon word) was false. Hence, the word $wm \in (\mathfrak{A} \cup \{m\})^*$ is a Lyndon word. That is, Assertion $\mathcal{F}''$ holds. Hence, we have proven the implication $\mathcal{E}' \Longrightarrow \mathcal{F}''$.

Proof of the implication $\mathcal{F}'' \Longrightarrow \mathcal{B}'$*:* Assume that Assertion $\mathcal{F}''$ holds. Thus, the word $wm \in (\mathfrak{A} \cup \{m\})^*$ is a Lyndon word.

Let u and v be nonempty words satisfying $w = uv$. We shall prove that either we have $v \geq w$ or the word v is a prefix of w.

Indeed, v is a suffix of w (since $w = uv$), so that vm is a suffix of wm. Clearly, vm is nonempty. Thus, Corollary 6.1.15 (applied to $\mathfrak{A} \cup \{m\}$, wm and vm instead of $\mathfrak{A}$, w and v) yields $vm \geq wm$. In other words, $wm \leq vm$. Hence, Proposition 6.1.2(e) (applied to $\mathfrak{A} \cup \{m\}$, w, m, v and m instead of $\mathfrak{A}$, a, b, c and d) yields that either we have $w \leq v$ or the word v is a prefix of w. In other words, either we have $v \geq w$ or the word v is a prefix of w.

Now, let us forget that we fixed u and v. We thus have proven that if u and v are nonempty words satisfying $w = uv$, then either we have $v \geq w$ or the word v is a prefix of w. In other words, Assertion $\mathcal{B}'$ holds. Hence, we have proven the implication $\mathcal{F}'' \Longrightarrow \mathcal{B}'$.

Proof of the implication $\mathcal{F}'' \Longrightarrow \mathcal{C}'$*:* Assume that Assertion $\mathcal{F}''$ holds. Thus, the word $wm \in (\mathfrak{A} \cup \{m\})^*$ is a Lyndon word.

Let u and v be nonempty words satisfying $w = uv$. We shall prove that either we have $v \geq u$ or the word v is a prefix of u.

Indeed, Proposition 6.1.14(b) (applied to $\mathfrak{A} \cup \{m\}$, wm and vm instead of $\mathfrak{A}$, w and v) yields $vm > u$ (since $\underbrace{w}_{=uv} m = uvm$, and since vm is nonempty). Hence, $vm \geq u = u\varnothing$, so that $u\varnothing \leq vm$. Thus, Proposition 6.1.2(e) (applied to $\mathfrak{A} \cup \{m\}$, u, $\varnothing$, v and m instead of $\mathfrak{A}$, a, b, c and d) yields that either we have $u \leq v$ or the word v is a prefix of u. In other words, either we have $v \geq u$ or the word v is a prefix of u.

Now, let us forget that we fixed u and v. We thus have proven that if u and v are nonempty words satisfying $w = uv$, then either we have $v \geq u$ or the word v is a prefix of u. In other words, Assertion $\mathcal{C}'$ holds. Hence, we have proven the implication $\mathcal{F}'' \Longrightarrow \mathcal{C}'$.

Proof of the implication $\mathcal{B}' \Longrightarrow \mathcal{G}'$*:* Assume that Assertion $\mathcal{B}'$ holds.

Let s be the longest suffix v of w satisfying $v < w$. (This is well-defined, because there exists a suffix v of w satisfying $v < w$ – namely, the empty word.) So we know that s is a suffix v of w satisfying $v < w$. In other words, s is a suffix of w and satisfies $s < w$. As a consequence, s is a proper suffix of w (because otherwise, s would be w, and this would contradict $s < w$). Hence, there exists a nonempty word $h \in \mathfrak{A}^*$ satisfying $w = hs$. Consider this h. Using Assertion $\mathcal{B}'$, it is easy to see that s is a prefix of w [985]. In other words, there exists a $g \in \mathfrak{A}^*$ such that $w = sg$. Consider this g.

We know that s is the longest suffix v of w satisfying $v < w$. Hence,

$$(\text{if } v \text{ is a suffix of } w \text{ satisfying } v < w, \text{ then } \ell(v) \leq \ell(s)). \tag{13.145.3}$$

There exists a nonnegative integer m such that h^m is a prefix of s (for example, the nonnegative integer $m = 0$). Consider the **maximal** such integer m [986]. Then, h^m is a prefix of s, but h^{m+1} is not a prefix of s. Since h^m is a prefix of s, there exists a word $q \in \mathfrak{A}^*$ such that $s = h^m q$. Consider this q. Clearly, $w = h \underbrace{s}_{=h^m q} = \underbrace{hh^m}_{=h^{m+1}} q = \underbrace{h^{m+1}}_{=h^m h} q = h^m hq$. Hence, $h^m hq = w = \underbrace{s}_{=h^m q} g = h^m qg$. Cancelling h^m from this

[985] *Proof.* Assume the contrary. Thus, s is not a prefix of w. Hence, s is nonempty. Therefore, Assertion $\mathcal{B}'$ (applied to $u = h$ and $v = s$) yields that either we have $s \geq w$ or the word s is a prefix of w. Since s is not a prefix of w, we must thus have $s \geq w$. But this contradicts $s < w$. This contradiction shows that our assumption was wrong, qed.

[986] This is well-defined, because of the following reason:

We have $\ell(h) \geq 1$ (since the word h is nonempty). Thus, for every $m \in \mathbb{N}$ satisfying $m > \ell(s)$, we have $\ell(h^m) = m \underbrace{\ell(h)}_{\geq 1} \geq m > \ell(s)$. In other words, for every $m \in \mathbb{N}$ satisfying $m > \ell(s)$, the word h^m is longer than s. Hence, for every $m \in \mathbb{N}$ satisfying $m > \ell(s)$, the word h^m cannot be a prefix of s. Thus, for every sufficiently high $m \in \mathbb{N}$, the word h^m cannot be a prefix of s.

We thus know the following:

- There exists a nonnegative integer m such that h^m is a prefix of s.
- For every sufficiently high $m \in \mathbb{N}$, the word h^m cannot be a prefix of s.

Consequently, there exists a **maximal** nonnegative integer m such that h^m is a prefix of s, qed.

equality, we obtain $hq = qg$. It is now easy to see that $h > q$ [987]. Hence, $q \leq h \leq hq = qg$. Thus, Proposition 6.1.2(g) (applied to q, h and g instead of a, b and c) yields that q is a prefix of h.

Next, we shall prove that the word h is Lyndon.

In fact, assume the contrary. Then, h is not Lyndon. Let v be the (lexicographically) smallest nonempty suffix of h. Then, Proposition 6.1.19(b) (applied to h instead of w) yields that there exists a nonempty $u \in \mathfrak{A}^*$ such that $h = uv$, $u \geq v$ and $uv \geq vu$. Consider this u. Since $w = \underbrace{h}_{=uv} s = uvs$, it is clear that the word vs is a suffix of w. If we had $vs < w$, then we could therefore obtain $\ell(vs) \leq \ell(s)$ (by (13.145.3), applied to vs instead of v), which would contradict $\ell(vs) = \underbrace{\ell(v)}_{\substack{>0 \\ (\text{since } v \text{ is nonempty})}} + \ell(s) > \ell(s)$. Thus, we cannot have $vs < w$. We thus have $vs \geq w = hs \geq h = uv$, so that $uv \leq vs$.

Recall that $s < w$. Hence, $vs \leq vw$ (by Proposition 6.1.2(b), applied to v, s and w instead of a, c and d). Now, $uv \geq vu$, so that $vu \leq uv \leq vs \leq v \underbrace{w}_{=hs} = v \underbrace{h}_{=uv} s = vuvs$. Hence, Proposition 6.1.2(g) (applied to vu, uv and vs instead of a, b and c) yields that vu is a prefix of uv. Since vu has the same length as uv (because $\ell(vu) = \ell(v) + \ell(u) = \ell(u) + \ell(v) = \ell(uv)$), this yields that $vu = uv$. Thus, the elements u and v of the monoid $\mathfrak{A}^*$ commute. Thus, the submonoid of $\mathfrak{A}^*$ generated by u and v is commutative. Since $h = uv$, the element h lies in this submonoid, and therefore the element h^m lies in it as well. Thus, h^m commutes with v (since this submonoid is commutative), i.e., we have $vh^m = h^m v$. Thus, $v \underbrace{s}_{=h^m q} = \underbrace{vh^m}_{=h^m v} q = h^m vq$. Thus, $h^m vq = vs \geq w = h \underbrace{s}_{=h^m q} = \underbrace{hh^m}_{=h^{m+1} = h^m h} q = h^m hq$, so that $h^m hq \leq h^m vq$. Hence, Proposition 6.1.2(c) (applied to h^m, hq and vq instead of a, c and d) yields $hq \leq vq$. But since q is a prefix of h, there exists a word $z \in \mathfrak{A}^*$ such that $h = qz$. Consider this z. We have

$$\begin{aligned} v \underbrace{qz}_{=h=uv} = \underbrace{vu}_{=uv=h} v = hv &\leq h \underbrace{vu}_{=uv=h=qz} \qquad (\text{since } hv \text{ is a prefix of } hvu) \\ &= hqz. \end{aligned}$$

Also, $\ell\left(\underbrace{h}_{=uv} q\right) = \ell(uvq) = \ell(u(vq)) = \underbrace{\ell(u)}_{\substack{>0 \\ (\text{since } u \text{ is nonempty})}} + \ell(vq) > \ell(vq)$, so that $\ell(vq) \leq \ell(hq)$. Hence, Proposition 6.1.2(f) (applied to vq, z, hq and z instead of a, b, c and d) yields $vq \leq hq$. Combined with $hq \leq vq$, this yields $vq = hq$. Hence, $\ell\left(\underbrace{vq}_{=hq}\right) = \ell(hq) > \ell(vq)$, which is absurd. This contradiction proves that our assumption is wrong. Thus, we have shown that the word h is Lyndon.

We now know that $h \in \mathfrak{A}^*$ is a Lyndon word, $m+1$ is a positive integer, and q is a prefix of h, and we have $w = h^{m+1} q$. Hence, there exists a Lyndon word $t \in \mathfrak{A}^*$, a positive integer ℓ and a prefix p of t (possibly empty) such that $w = t^\ell p$ (namely, $t = h$, $\ell = m+1$ and $p = q$). In other words, Assertion $\mathcal{G}'$ holds. This proves the implication $\mathcal{B}' \Longrightarrow \mathcal{G}'$.

We are now going to prove the implication $\mathcal{F}' \Longrightarrow \mathcal{B}'$; this implication will later be used in the proof of the implication $\mathcal{H}' \Longrightarrow \mathcal{B}'$.

Proof of the implication $\mathcal{F}' \Longrightarrow \mathcal{B}'$: Assume that Assertion $\mathcal{F}'$ holds. In other words, the word w is a prefix of a Lyndon word in $\mathfrak{A}^*$. Let z be this Lyndon word. Thus, w is a prefix of z. In other words, there exists a word $q \in \mathfrak{A}^*$ such that $z = wq$. Consider this q.

Let u and v be nonempty words satisfying $w = uv$. We are going to prove that either we have $v \geq w$ or the word v is a prefix of w.

[987] *Proof.* Assume the contrary. Then, $h \leq q$. Hence, $h \leq q \leq qg = hq$ (since $hq = qg$). Therefore, Proposition 6.1.2(g) (applied to h, q and q instead of a, b and c) yields that h is a prefix of q. In other words, there exists a word $r \in \mathfrak{A}^*$ such that $q = hr$. Consider this r. Now, $s = h^m \underbrace{q}_{=hr} = \underbrace{h^m h}_{=h^{m+1}} r = h^{m+1} r$, so that h^{m+1} is a prefix of s. This contradicts the fact that h^{m+1} is not a prefix of s. This contradiction proves that our assumption was wrong, qed.

We have $z = \underbrace{w}_{=uv} q = uvq$. Thus, Proposition 6.1.14(a) (applied to z and vq instead of w and v) yields $vq \geq z$ (since vq is nonempty (since v is nonempty)). Hence, $vq \geq z = wq$. In other words, $wq \leq vq$. Thus, Proposition 6.1.2(e) (applied to w, q, v and q instead of a, b, c and d) yields that either we have $w \leq v$ or the word v is a prefix of w. In other words, either we have $v \geq w$ or the word v is a prefix of w.

Now, forget that we fixed u and v. We thus have shown that if u and v are nonempty words satisfying $w = uv$, then either we have $v \geq w$ or the word v is a prefix of w. In other words, Assertion $\mathcal{B}'$ holds. Thus, the implication $\mathcal{F}' \Longrightarrow \mathcal{B}'$ is proven.

Proof of the implication $\mathcal{H}' \Longrightarrow \mathcal{B}'$: Assume that Assertion $\mathcal{H}'$ holds. In other words, there exists a Lyndon word $t \in \mathfrak{A}^*$, a nonnegative integer ℓ and a prefix p of t (possibly empty) such that $w = t^\ell p$. Consider this t, this ℓ and this p.

We are going to prove that for every $m \in \mathbb{N}$,

$$\text{(every suffix } s \text{ of } t^m p \text{ satisfies either } s \geq t^m p \text{ or } (\text{the word } s \text{ is a prefix of } t^m p)) . \tag{13.145.4}$$

Proof of (13.145.4): We will prove (13.145.4) by induction over m:

Induction base: Using the implication $\mathcal{F}' \Longrightarrow \mathcal{B}'$, it is easy to see that (13.145.4) holds for $m = 0$ [988]. This completes the induction base.

Induction step: Let M be a positive integer. Assume that (13.145.4) is proven for $m = M - 1$. We will now show that (13.145.4) holds for $m = M$.

Let r denote the word $t^{M-1}p$. It is easy to see that r is a prefix of $t^M p$ [989]. In other words, there exists a word $g \in \mathfrak{A}^*$ such that $t^M p = rg$. Consider this g.

Let s be a suffix of $t^M p$. We shall show that either $s \geq t^M p$ or $\left(\text{the word } s \text{ is a prefix of } t^M p\right)$.

In order to prove this, let us assume the contrary (for the sake of contradiction). Then, neither $s \geq t^M p$ nor $\left(\text{the word } s \text{ is a prefix of } t^M p\right)$. In other words, we have $s < t^M p$, and the word s is not a prefix of $t^M p$. If the word s was a prefix of r, then the word s would be a prefix of $t^M p$ (since r is a prefix of $t^M p$), which would contradict the fact that the word s is not a prefix of $t^M p$. Hence, the word s cannot be a prefix of r. In other words, the word s cannot be a prefix of $t^{M-1}p$ (since $r = t^{M-1}p$).

The word s is a suffix of $\underbrace{t^M}_{=tt^{M-1}} p = t \underbrace{t^{M-1}p}_{=r} = tr$. Therefore, we must be in one of the following two cases (depending on whether this suffix begins before the suffix r of tr begins or afterwards):

Case 1: The word s is a suffix of r. (Note that $s = r$ is allowed.)

Case 2: The word s has the form $s'r$ where s' is a nonempty suffix of t.

Let us consider Case 1 first. In this case, the word s is a suffix of r. In other words, the word s is a suffix of $t^{M-1}p$ (since $r = t^{M-1}p$). Hence, (13.145.4) (applied to $m = M - 1$) yields that either $s \geq t^{M-1}p$

[988] *Proof.* Assume that $m = 0$. Then, $t^m p = \underbrace{t^0}_{=\varnothing} p = \varnothing p = p$.

Let s be a suffix of $t^m p$. Then, s is a suffix of $t^m p = p$. In other words, there exists a word $g \in \mathfrak{A}^*$ satisfying $p = gs$. Consider this g.

We are going to prove that either $s \geq t^m p$ or (the word s is a prefix of $t^m p$). If the word s is empty, then this is obvious (because if the word s is empty, then the word s is a prefix of $t^m p$). Hence, we WLOG assume that the word s is nonempty. If the word g is empty, then it is also clear that either $s \geq t^m p$ or (the word s is a prefix of $t^m p$) (because if the word g is empty, then $g = \varnothing$ and thus $t^m p = p = \underbrace{g}_{=\varnothing} s = s$, so that the word s is a prefix of $t^m p$). Hence, we WLOG assume that the word g is nonempty.

But the word p is a prefix of a Lyndon word in $\mathfrak{A}^*$ (since p is a prefix of t, and since t is a Lyndon word in $\mathfrak{A}^*$). In other words, Assertion $\mathcal{F}'$ with w replaced by p is satisfied. Hence, Assertion $\mathcal{B}'$ with w replaced by p is satisfied as well (since we have already proven the implication $\mathcal{F}' \Longrightarrow \mathcal{B}'$). In other words,

$$\left(\begin{array}{c} \text{if } u \text{ and } v \text{ are nonempty words satisfying } p = uv, \text{ then} \\ \text{either we have } v \geq p \text{ or the word } v \text{ is a prefix of } p \end{array} \right) . \tag{13.145.5}$$

Since the words g and s are nonempty, we can apply (13.145.5) to $u = g$ and $v = s$. As a result, we obtain that either we have $s \geq p$ or the word s is a prefix of p. In other words, either $s \geq p$ or (the word s is a prefix of p). In other words, either $s \geq t^m p$ or (the word s is a prefix of $t^m p$) (since $t^m p = p$). This proves (13.145.4).

[989] *Proof.* There exists a word $q \in \mathfrak{A}^*$ such that $t = pq$ (since p is a prefix of t). Consider this q. We have $\underbrace{t^M}_{=t^{M-1}t} p = t^{M-1} \underbrace{t}_{=pq} p = \underbrace{t^{M-1}p}_{=r} qp = rqp = r(qp)$. Hence, r is a prefix of $t^M p$, qed.

or $\left(\text{the word } s \text{ is a prefix of } t^{M-1}p\right)$ (since (13.145.4) holds for $m = M-1$). Since the word s cannot be a prefix of $t^{M-1}p$, we thus must have $s \geq t^{M-1}p$. Thus, $t^{M-1}p \leq s < t^M p = rg$. Hence, $r = t^{M-1}p \leq s < rg$. Therefore, Proposition 6.1.2(g) (applied to r, s and g instead of a, b and c) yields that r is a prefix of s. Since $\ell(r) \geq \ell(s)$ (because s is a suffix of r), this can only hold if $r = s$. We thus have $r = s$. Thus, $s = r$, so that s is a prefix of $s = r$. This contradicts the fact that the word s cannot be a prefix of r. Thus, we have found a contradiction in Case 1.

Let us now consider Case 2. In this case, the word s has the form $s'r$ where s' is a nonempty suffix of t. Consider this s'. Corollary 6.1.15 (applied to t and s' instead of w and v) yields $s' \geq t$. But s' is a suffix of t, so that $\ell(s') \leq \ell(t)$. Also, $s = s'r$, so that $s'r = s < t^M p = tr$. Hence, Proposition 6.1.2(f) (applied to s', r, t and r instead of a, b, c and d) yields that $s' \leq t$. Combined with $s' \geq t$, this yields $s' = t$. Hence, $s = \underbrace{s'}_{=t} r = t^M p$ (since $t^M p = tr$), which contradicts the fact that the word s is not a prefix of $t^M p$. Thus, we have found a contradiction in Case 2.

We have thus obtained a contradiction in each of the two Cases 1 and 2. Since these two Cases cover all possibilities, this shows that we always get a contradiction. This completes the proof that either $s \geq t^M p$ or $\left(\text{the word } s \text{ is a prefix of } t^M p\right)$.

Now, forget that we fixed s. We thus have shown that every suffix s of $t^M p$ satisfies either $s \geq t^M p$ or $\left(\text{the word } s \text{ is a prefix of } t^M p\right)$. In other words, (13.145.4) holds for $m = M$. This completes the induction step, and thus (13.145.4) is proven by induction.

Now, let u and v be nonempty words satisfying $w = uv$. Then, v is a suffix of $w = t^\ell p$. Hence, (13.145.4) (applied to $m = \ell$ and $s = v$) yields that either $v \geq t^\ell p$ or $\left(\text{the word } v \text{ is a prefix of } t^\ell p\right)$. In other words, either $v \geq w$ or (the word v is a prefix of w) (since $w = t^\ell p$). In other words, either we have $v \geq w$ or the word v is a prefix of w.

Now, forget that we fixed u and v. We thus have shown that if u and v are nonempty words satisfying $w = uv$, then either we have $v \geq w$ or the word v is a prefix of w. In other words, Assertion $\mathcal{B}'$ holds. Thus, the implication $\mathcal{H}' \Longrightarrow \mathcal{B}'$ is proven.

We have thus proven the implications $\mathcal{B}' \Longrightarrow \mathcal{E}'$, $\mathcal{C}' \Longrightarrow \mathcal{E}'$, $\mathcal{G}' \Longrightarrow \mathcal{H}'$, $\mathcal{E}' \Longrightarrow \mathcal{F}''$, $\mathcal{F}'' \Longrightarrow \mathcal{B}'$, $\mathcal{F}' \Longrightarrow \mathcal{C}'$, $\mathcal{B}' \Longrightarrow \mathcal{G}'$ and $\mathcal{H}' \Longrightarrow \mathcal{B}'$. Combined, these yield the equivalence $\mathcal{B}' \Longleftrightarrow \mathcal{C}' \Longleftrightarrow \mathcal{E}' \Longleftrightarrow \mathcal{F}'' \Longleftrightarrow \mathcal{G}' \Longleftrightarrow \mathcal{H}'$. This solves Exercise 6.1.32(b).

(c) The implication $\mathcal{F}' \Longrightarrow \mathcal{B}'$ has already been proven in our solution of Exercise 6.1.32(b). Hence, Exercise 6.1.32(c) is solved.

(d) Assume that Assertion $\mathcal{D}'$ holds. Then, Assertion $\mathcal{A}'$ holds as well (because of the equivalence $\mathcal{A}' \Longleftrightarrow \mathcal{D}'$). In other words, the word w is a power of a Lyndon word. In other words, there exist a Lyndon word $z \in \mathfrak{A}^*$ and a nonnegative integer m such that $w = z^m$. Consider these z and m. There exists a Lyndon word $t \in \mathfrak{A}^*$, a nonnegative integer ℓ and a prefix p of t (possibly empty) such that $w = t^\ell p$ (namely, $t = z$, $\ell = m$ and $p = \varnothing$). In other words, Assertion $\mathcal{H}'$ holds. Since Assertion $\mathcal{H}'$ is equivalent to Assertion $\mathcal{B}'$ (because of the equivalence $\mathcal{B}' \Longleftrightarrow \mathcal{C}' \Longleftrightarrow \mathcal{E}' \Longleftrightarrow \mathcal{F}'' \Longleftrightarrow \mathcal{G}' \Longleftrightarrow \mathcal{H}'$), this yields that Assertion $\mathcal{B}'$ holds. Thus, the implication $\mathcal{D}' \Longrightarrow \mathcal{B}'$ is proven.

(e) Assume that there exists a letter $\mu \in \mathfrak{A}$ such that ($\mu > a$ for every letter a of w). Consider this μ. We need to prove that the equivalence $\mathcal{F}' \Longleftrightarrow \mathcal{F}''$ holds.

Combining the implication $\mathcal{F}' \Longrightarrow \mathcal{B}'$ (which has already been proven) and the implication $\mathcal{B}' \Longrightarrow \mathcal{F}''$ (which follows from the equivalence $\mathcal{B}' \Longleftrightarrow \mathcal{C}' \Longleftrightarrow \mathcal{E}' \Longleftrightarrow \mathcal{F}'' \Longleftrightarrow \mathcal{G}' \Longleftrightarrow \mathcal{H}'$ proven above), we obtain the implication $\mathcal{F}' \Longrightarrow \mathcal{F}''$. Thus, in order to prove the equivalence $\mathcal{F}' \Longleftrightarrow \mathcal{F}''$, it is enough to verify the implication $\mathcal{F}'' \Longrightarrow \mathcal{F}'$. Let us do this now.

Assume that Assertion $\mathcal{F}''$ holds. Let $\mathfrak{B}$ denote the alphabet consisting of all letters that appear in w. Clearly, $\mathfrak{B}$ is a subalphabet of $\mathfrak{A}$, and we have $w \in \mathfrak{B}^*$. Moreover, we have $\mu > a$ for every letter a of w. Therefore, $\mu > a$ for every $a \in \mathfrak{B}$ (because the elements of $\mathfrak{B}$ are precisely the letters of w). In other words, $a < \mu$ for every $a \in \mathfrak{B}$. As a consequence, $\mu \notin \mathfrak{B}$; that is, μ is an object not in the alphabet $\mathfrak{B}$.

We know that Assertion $\mathcal{F}''$ holds. Due to the implication $\mathcal{F}'' \Longrightarrow \mathcal{B}'$ (which follows from the equivalence $\mathcal{B}' \Longleftrightarrow \mathcal{C}' \Longleftrightarrow \mathcal{E}' \Longleftrightarrow \mathcal{F}'' \Longleftrightarrow \mathcal{G}' \Longleftrightarrow \mathcal{H}'$ proven above), this yields that Assertion $\mathcal{B}'$ holds. Thus, Assertion $\mathcal{B}'$ with $\mathfrak{A}$ replaced by $\mathfrak{B}$ holds as well (since the Assertion $\mathcal{B}'$ does not change if we extend our alphabet). Due to the implication $\mathcal{B}' \Longrightarrow \mathcal{F}''$ (which follows from the equivalence $\mathcal{B}' \Longleftrightarrow \mathcal{C}' \Longleftrightarrow \mathcal{E}' \Longleftrightarrow \mathcal{F}'' \Longleftrightarrow \mathcal{G}' \Longleftrightarrow \mathcal{H}'$ proven above), this yields that Assertion $\mathcal{F}''$ with $\mathfrak{A}$ replaced by $\mathfrak{B}$ holds as well. Thus, we can apply Assertion $\mathcal{F}''$ with $\mathfrak{A}$ replaced by $\mathfrak{B}$ to $m = \mu$ (because μ is an object not in the alphabet $\mathfrak{B}$, and the total order on $\mathfrak{A}$ satisfies ($a < \mu$ for every $a \in \mathfrak{B}$)). As a result, we conclude that the word $w\mu \in (\mathfrak{B} \cup \{\mu\})^*$ is a Lyndon word. Since $\mathfrak{B} \cup \{\mu\} \subset \mathfrak{A}$ (because $\mathfrak{B} \subset \mathfrak{A}$ and $\mu \in \mathfrak{A}$), we have $w\mu \in \left(\underbrace{\mathfrak{B} \cup \{\mu\}}_{\subset \mathfrak{A}}\right)^* \subset \mathfrak{A}^*$, and thus $w\mu \in \mathfrak{A}^*$ is a Lyndon word. Of course, the word w is a prefix of $w\mu$. As a consequence, the word w is a prefix of a Lyndon word in $\mathfrak{A}^*$ (namely, of the word $w\mu$). In other words, Assertion $\mathcal{F}'$ holds. This proves the implication $\mathcal{F}'' \Longrightarrow \mathcal{F}'$. Thus, the solution of Exercise 6.1.32(e) is complete.

(f) Assume that there exists a letter $\mu \in \mathfrak{A}$ such that ($\mu > a$ for some letter a of w). Consider this μ. We need to prove that the equivalence $\mathcal{F}' \Longleftrightarrow \mathcal{F}''$ holds.

Just as in our solution of Exercise 6.1.32(e) above, we can see that it is enough to verify the implication $\mathcal{F}'' \Longrightarrow \mathcal{F}'$. Let us do this now.

Assume that Assertion $\mathcal{F}''$ holds. Due to the implication $\mathcal{F}'' \Longrightarrow \mathcal{B}'$ (which follows from the equivalence $\mathcal{B}' \Longleftrightarrow \mathcal{C}' \Longleftrightarrow \mathcal{E}' \Longleftrightarrow \mathcal{F}'' \Longleftrightarrow \mathcal{G}' \Longleftrightarrow \mathcal{H}'$ proven above), this yields that Assertion $\mathcal{B}'$ holds.

Let β be the highest letter of the word $w\mu \in \mathfrak{A}^*$ (the concatenation of the word w with the one-letter word μ). The word $w\beta^{\ell(w)}$ is clearly nonempty (since w is nonempty). We shall now prove that the word $w\beta^{\ell(w)}$ is Lyndon.

Indeed, assume the contrary. Thus, $w\beta^{\ell(w)}$ is not Lyndon. Let v denote the (lexicographically) smallest nonempty suffix of $w\beta^{\ell(w)}$. Then, Proposition 6.1.19(b) (applied to $w\beta^{\ell(w)}$ instead of w) yields that there exists a nonempty $u \in \mathfrak{A}^*$ such that $w\beta^{\ell(w)} = uv$, $u \geq v$ and $uv \geq vu$.

Let f be the first letter of the word w (this is well-defined since w is nonempty). Then, we can write w in the form $w = fs$ for some word $s \in \mathfrak{A}^*$. Consider this s.

We know that $\mu > a$ for some letter a of w. Consider this a. We also know that β is the highest letter of the word $w\mu$. Thus, β is $\geq$ to every letter of the word $w\mu$. In particular, this yields that $\beta \geq \mu$ (since μ is a letter of the word $w\mu$), so that $\beta \geq \mu > a$. But it is fairly easy to see (using the fact that Assertion $\mathcal{B}'$ holds) that $a \geq f$ [990]. Hence, $\beta > a \geq f$.

The word v is a proper suffix of $w\beta^{\ell(w)}$ (since $w\beta^{\ell(w)} = uv$ and since u is nonempty) and is nonempty. Hence, v is a nonempty proper suffix of $w\beta^{\ell(w)}$. Therefore, we must be in one of the following two cases (depending on whether this suffix begins before the suffix $\beta^{\ell(w)}$ of $w\beta^{\ell(w)}$ begins or afterwards):

Case 1: The word v is a nonempty suffix of $\beta^{\ell(w)}$. (Note that $v = \beta^{\ell(w)}$ is allowed.)

Case 2: The word v has the form $q\beta^{\ell(w)}$ where q is a nonempty proper suffix of w.

Let us first consider Case 1. In this case, the word v is a nonempty suffix of $\beta^{\ell(w)}$. Thus, v has the form β^i for some $i \in \{0, 1, \ldots, \ell(w)\}$. Consider this i. We have $v = \beta^i$. Thus, $i \neq 0$ (since v is nonempty). Hence, (the first letter of v) $= \beta$. But since the word w is nonempty, we have

$$\left(\text{the first letter of } w\beta^{\ell(w)}\right) = (\text{the first letter of } w) = f$$

(since f was defined to be the first letter of w). Now,

$$(\text{the first letter of } v) = \beta > f = \left(\text{the first letter of } w\beta^{\ell(w)}\right).$$

By the definition of lexicographic order, this shows that $v > w\beta^{\ell(w)}$. But this contradicts $w\beta^{\ell(w)} = uv \geq vu \geq v$. Hence, we have found a contradiction in Case 1.

[990] *Proof.* Assume the contrary. Thus, $a < f$.

Since a is a letter of w, the word w must have a suffix which begins with the letter a. Let p be this suffix. Then, there exists a word $q \in \mathfrak{A}^*$ such that $w = qp$ (since p is a suffix of w). Consider this q. The word p is nonempty (since it begins with a). Since the word p begins with the letter a, we have

$$(\text{the first letter of } p) = a < f = (\text{the first letter of } w)\,. \tag{13.145.6}$$

By the definition of the lexicographic order, this shows that $p < w$. Hence, $p \neq w$. Now, the word q is nonempty (since otherwise, we would have $q = \varnothing$ and thus $w = \underbrace{q}_{=\varnothing} p = p$, contradicting $p \neq w$). Hence, applying Assertion $\mathcal{B}'$ to q and p instead of u and v, we conclude that either we have $p \geq w$ or the word p is a prefix of w. Since $p \geq w$ is impossible (because $p < w$), this yields that the word p is a prefix of w. Since p is nonempty, this shows that p is a nonempty prefix of w. But this yields that

$$(\text{the first letter of } p) = (\text{the first letter of } w)\,;$$

this contradicts (13.145.6). This contradiction proves that our assumption was wrong, qed.

Let us now consider Case 2. In this case, the word v has the form $q\beta^{\ell(w)}$ where q is a nonempty proper suffix of w. Consider this q. Notice that $\ell(q) < \ell(w)$ (since q is a proper suffix of w), so that $\ell(w) > \ell(q)$.

Since q is a proper suffix of w, there exists a nonempty word $u \in \mathfrak{A}^*$ such that $w = uq$. Consider this u. Recall that Assertion $\mathcal{B}'$ holds. Applying this Assertion $\mathcal{B}'$ to q instead of v, we conclude that either we have $q \geq w$ or the word q is a prefix of w. If the word q is not a prefix of w, then it is very easy to derive a contradiction[991]. Hence, the word q must be a prefix of w. In other words, there exists a word $g \in \mathfrak{A}^*$ such that $w = qg$. Consider this g. Notice that $\ell\left(\underbrace{w}_{=qg}\right) = \ell(qg) = \underbrace{\ell(q)}_{\substack{>0 \\ (\text{since } q \text{ is nonempty})}} + \ell(g) > \ell(g)$, so that $\ell\left(\beta^{\ell(w)}\right) = \ell(w) > \ell(g)$. In other words, the word $\beta^{\ell(w)}$ is longer than g. It is now easy to see that $\beta^{\ell(w)} \geq g$ [992]. In other words, $g \leq \beta^{\ell(w)}$. Hence, Proposition 6.1.2(b) (applied to q, g and $\beta^{\ell(w)}$ instead of a, c and d) yields $qg \leq q\beta^{\ell(w)}$, so that $w = qg \leq q\beta^{\ell(w)} = v$ (since $v = q\beta^{\ell(w)}$). Thus, $v \geq w$, so that $u \geq v \geq w$.

But $v = q\beta^{\ell(w)}$, so that $w\beta^{\ell(w)} = u \underbrace{v}_{=q\beta^{\ell(w)}} = uq\beta^{\ell(w)}$. Cancelling $\beta^{\ell(w)}$ from this equation, we obtain $w = uq$. Thus, $u \leq uq = w$. Since $u \neq w$ [993], this becomes $u < w$. This contradicts $u \geq w$. Hence, we have found a contradiction in Case 2.

We have now obtained a contradiction in each of our two Cases 1 and 2. Since these two Cases cover all possibilities, this yields that we always obtain a contradiction. Thus, our assumption was wrong, and we conclude that the word $w\beta^{\ell(w)}$ is Lyndon. Hence, w is a prefix of a Lyndon word in $\mathfrak{A}^*$ (because w is a prefix of the word $w\beta^{\ell(w)}$). In other words, Assertion $\mathcal{F}'$ holds. This proves the implication $\mathcal{F}'' \Longrightarrow \mathcal{F}'$. This solves Exercise 6.1.32(f).

Remark: Of course, for a letter $\mu \in \mathfrak{A}$, if we have ($\mu > a$ for every letter a of w), then we also have ($\mu > a$ for some letter a of w) (since w is nonempty and thus has at least one letter). Hence, Exercise 6.1.32(e) is a particular case of Exercise 6.1.32(f).

13.146. **Solution to Exercise 6.1.33.** *Solution to Exercise 6.1.33.*

(a) If we replace the word "total" by "partial" throughout the proof of Proposition 6.1.2, then the resulting argument can be used in the partial-order setting to prove Proposition 6.1.2 with "a total order" replaced by "a partial order". Hence, Proposition 6.1.2 holds in the partial-order setting, as long as one replaces "a total order" by "a partial order" in part (a) of this Proposition. This solves Exercise 6.1.33(a).

[991] *Proof.* Assume that the word q is not a prefix of w. Then, we have $q \geq w$ (since we know that either we have $q \geq w$ or the word q is a prefix of w). In other words, $w \leq q$. Hence, Proposition 6.1.2(d) (applied to w, $\beta^{\ell(w)}$, q and $\beta^{\ell(w)}$ instead of a, b, c and d) yields that either we have $w\beta^{\ell(w)} \leq q\beta^{\ell(w)}$ or the word w is a prefix of q. Since the word w is not a prefix of q (because $\ell(w) > \ell(q)$), this yields that we have $w\beta^{\ell(w)} \leq q\beta^{\ell(w)}$. Since $\ell\left(w\beta^{\ell(w)}\right) = \underbrace{\ell(w)}_{>\ell(q)} + \ell\left(\beta^{\ell(w)}\right) > \ell(q) + \ell\left(\beta^{\ell(w)}\right) = \ell\left(q\beta^{\ell(w)}\right)$, this shows that $w\beta^{\ell(w)} < q\beta^{\ell(w)} = v$ (because $v = q\beta^{\ell(w)}$). Hence, $v > w\beta^{\ell(w)} = uv \geq vu \geq v$, which is a contradiction, qed.

[992] *Proof.* Assume the contrary. Then, $\beta^{\ell(w)} < g$. By the definition of the lexicographic order, this yields that

$$\begin{aligned}&\textbf{either } \text{there exists an } i \in \left\{1, 2, \ldots, \min\left\{\ell\left(\beta^{\ell(w)}\right), \ell(g)\right\}\right\} \\ &\qquad \text{such that } \left(\left(\beta^{\ell(w)}\right)_i < g_i, \text{ and every } j \in \{1, 2, \ldots, i-1\} \text{ satisfies } \left(\beta^{\ell(w)}\right)_j = g_j\right), \\ &\textbf{or } \text{the word } \beta^{\ell(w)} \text{ is a prefix of } g.\end{aligned}$$

Since the word $\beta^{\ell(w)}$ cannot be a prefix of g (because the word $\beta^{\ell(w)}$ is longer than g), this yields that there exists an $i \in \left\{1, 2, \ldots, \min\left\{\ell\left(\beta^{\ell(w)}\right), \ell(g)\right\}\right\}$ such that $\left(\left(\beta^{\ell(w)}\right)_i < g_i, \text{ and every } j \in \{1, 2, \ldots, i-1\} \text{ satisfies } \left(\beta^{\ell(w)}\right)_j = g_j\right)$. Consider this i. We have $\left(\beta^{\ell(w)}\right)_i = \beta$, so that $\beta = \left(\beta^{\ell(w)}\right)_i < g_i$.

But g is a suffix of w (since $w = qg$), and thus every letter of g is a letter of w. In particular, g_i is a letter of w. Hence, g_i is a letter of $w\mu$ as well (since every letter of w is a letter of $w\mu$ (because w is a prefix of $w\mu$)). Since β is the highest letter of $w\mu$, this yields $\beta \geq g_i$. But this contradicts $\beta < g_i$. This contradiction proves that our assumption was false, qed.

[993] *Proof.* Assume the contrary. Then, $u = w$. Hence, $u\varnothing = u = w = uq$. Cancelling u from this equation, we obtain $\varnothing = q$. Hence, the word q is empty. This contradicts the fact that q is nonempty. This contradiction shows that our assumption was wrong, qed.

(b) Let us work in the partial-order setting. Let $a, b, c, d \in \mathfrak{A}^*$ be four words such that the words ab and cd are comparable. We need to prove that the words a and c are comparable.

The words ab and cd are comparable. In other words, we have either $ab \leq cd$ or $ab \geq cd$. In other words, we have either $ab \leq cd$ or $cd \leq ab$. We can WLOG assume that $ab \leq cd$ (since otherwise, we can achieve $ab \leq cd$ by switching a and b with c and d, respectively). Assume this. We know (from Exercise 6.1.33(a)) that Proposition 6.1.2 holds in the partial-order setting, as long as one replaces "a total order" by "a partial order" in part (a) of this Proposition. In particular, Proposition 6.1.2(e) holds in the partial-order setting. Applying Proposition 6.1.2(e), we thus conclude that either we have $a \leq c$ or the word c is a prefix of a. Thus, either we have $a \leq c$ or we have $c \leq a$ (because if c is a prefix of a, then $c \leq a$). In other words, either we have $a \leq c$ or we have $a \geq c$. In other words, the words a and c are comparable. This solves Exercise 6.1.33(b).

(c) Recall (from Exercise 6.1.33(a)) that Proposition 6.1.2 holds in the partial-order setting, as long as one replaces "a total order" by "a partial order" in part (a) of this Proposition.

Proposition 6.1.4 holds in the partial-order setting, because its proof applies verbatim in this setting.

Proposition 6.1.5 holds in the partial-order setting, because its proof applies verbatim in this setting.

The proof of Corollary 6.1.6 given above is no longer applicable in the partial-order setting. However, the alternative proof of Corollary 6.1.6 given in the solution to Exercise 6.1.7 does apply verbatim in the partial-order setting. Thus, Corollary 6.1.6 holds in the partial-order setting.

Corollary 6.1.8 holds in the partial-order setting, because its proof applies verbatim in this setting.

Exercise 6.1.9 and Exercise 6.1.10 hold in the partial-order setting, because their solutions apply verbatim in this setting.

Exercise 6.1.11 and Exercise 6.1.12 hold in the partial-order setting, because their solutions apply verbatim in this setting.

Proposition 6.1.14 holds in the partial-order setting, because its proof applies verbatim in this setting.

Corollary 6.1.15 holds in the partial-order setting, because its proof applies verbatim in this setting.

Proposition 6.1.16 holds in the partial-order setting, because its proof applies verbatim in this setting.

Corollary 6.1.17 holds in the partial-order setting, because its proof applies verbatim in this setting.

Our proof of Proposition 6.1.18 does not directly apply in the partial-order setting; however, it can be tweaked so that it does:

Proof of Proposition 6.1.18 in the partial-order setting. If the words u and v are incomparable, then Proposition 6.1.18 is easily seen to hold[994]. Hence, for the rest of this proof, we can WLOG assume that the words u and v are comparable. Assume this.

The words u and v are comparable. In other words, we have either $u < v$ or $u = v$ or $u > v$. From here, we can proceed as in our proof of Proposition 6.1.18 in the total-order setting. Proposition 6.1.18 is thus proven in the partial-order setting. □

Exercise 6.1.21(a) holds in the partial-order setting, because its solution applies verbatim in this setting.

The proof of Theorem 6.1.20 we gave (using Proposition 6.1.19) does not work in the partial-order setting[995]. However, the alternative proof of Theorem 6.1.20 given in Exercise 6.1.21(b) does apply verbatim in the partial-order setting. Thus, Theorem 6.1.20 holds in the partial-order setting.

Exercise 6.1.23 and Exercise 6.1.24 hold in the partial-order setting, because their solutions apply verbatim in this setting.

Exercise 6.1.31(a) and Exercise 6.1.31(b) hold in the partial-order setting, because their solutions apply verbatim in this setting.

Exercise 6.1.33(c) is thus solved.

[994] *Proof.* Assume that u and v are incomparable. If the words uv and vu were comparable, then the words u and v would also be comparable (by Exercise 6.1.33(b), applied to $a = u$, $b = v$, $c = v$ and $d = u$), which would contradict the fact that u and v are incomparable. Hence, the words uv and vu are incomparable. Thus, we cannot have $uv \geq vu$. But we cannot have $u \geq v$ either (since u and v are incomparable). Thus, neither $u \geq v$ nor $uv \geq vu$ holds. Hence, $u \geq v$ if and only if $uv \geq vu$; therefore, Proposition 6.1.18 holds, qed.

[995] One might try tweaking Proposition 6.1.19 for the partial-order setting by replacing "the (lexicographically) smallest nonempty suffix" by "a nonempty suffix which is (lexicographically) minimal among the nonempty suffices", and by replacing "$u \geq v$ and $uv \geq vu$" by "neither $u < v$ nor $uv < vu$". But this still would not hold. For example, if w is the word $XX12$ over the partially ordered alphabet $\{X, 1, 2\}$ with relation $1 < 2$, then the nonempty suffix $X12$ is lexicographically minimal among such suffixes, but we do have $X < X12$. (At least $uv < vu$ does not hold indeed.)

[*Remark:* In the above solution of Exercise 6.1.33(c), we have used the results of Exercise 6.1.21(b) and Exercise 6.1.7. This is the main reason why the latter two exercises have been written. However, there is a way to avoid them and still prove that Theorem 6.1.20 and Corollary 6.1.6 hold in the partial-order setting. This uses a trick, which we shall now explain.

First, a definition:

Definition 13.146.1. Let P be a poset.

(a) A poset Q is said to be an *extension* of P if and only if the following two statements hold:
- We have $Q = P$ as sets.
- Any two elements a and b of P satisfying $a \leq b$ in P satisfy $a \leq b$ in Q.

(b) An extension Q of P is called a *linear extension* of P if and only if the poset Q is totally ordered.[996]

The following fact about extensions of posets is well-known:

Proposition 13.146.2. *Let P be a finite poset.*

(a) *If a and b are two incomparable elements of P, then there exists an extension Q of P such that we have $a < b$ in Q.*

(b) *If a and b are two elements of P which don't satisfy $a \geq b$, then there exists an extension Q of P such that we have $a < b$ in Q.*

(c) *There exists a linear extension of P.*

(d) *If a and b are two elements of P which don't satisfy $a \geq b$, then there exists a linear extension Q of P such that we have $a < b$ in Q.*

(Actually, the requirement that P be finite in Proposition 13.146.2 can be dropped if you accept Zorn's lemma, but we do not need this generality.)

We can now apply this all to alphabets. In the partial-order setting, alphabets are posets, and so it makes sense to speak of an extension of an alphabet. We notice the following fact:

Proposition 13.146.3. *Let $\mathfrak{A}$ be a finite poset.*

(a) *If $\mathfrak{B}$ is an extension of the poset $\mathfrak{A}$, then $\mathfrak{B}^*$ is an extension of the poset $\mathfrak{A}^*$.*

(b) *Let $u \in \mathfrak{A}^*$ and $v \in \mathfrak{A}^*$ be two words. Then, $u < v$ holds in $\mathfrak{A}^*$ if and only if we have*

$$(u < v \text{ in } \mathfrak{B}^* \text{ for every linear extension } \mathfrak{B} \text{ of } \mathfrak{A}) .$$

Proof of Proposition 13.146.3. (a) Let $\mathfrak{B}$ be an extension of the poset $\mathfrak{A}$. Then, by the definition of an "extension", we see that:

- We have $\mathfrak{B} = \mathfrak{A}$ as sets.
- Any two elements a and b of $\mathfrak{A}$ satisfying $a \leq b$ in $\mathfrak{A}$ satisfy $a \leq b$ in $\mathfrak{B}$.

Now, it is clear that $\mathfrak{B}^* = \mathfrak{A}^*$ as sets (since $\mathfrak{B} = \mathfrak{A}$ as sets). Also, any two elements a and b of $\mathfrak{A}^*$ satisfying $a \leq b$ in $\mathfrak{A}^*$ satisfy $a \leq b$ in $\mathfrak{B}^*$ [997]. Consequently, $\mathfrak{B}^*$ is an extension of the poset $\mathfrak{A}^*$. This proves Proposition 13.146.3(a).

[996] This notion of a linear extension is identical to the one used in Theorem 5.2.11, except that we don't require P to be finite here.

[997] *Proof.* Let u and v be two elements of $\mathfrak{A}^*$ satisfying $u \leq v$ in $\mathfrak{A}^*$. We are going to prove that $u \leq v$ in $\mathfrak{B}^*$.

According to the definition of the relation $\leq$ on $\mathfrak{A}^*$, we have that

$$\begin{aligned} &\textbf{either } \text{there exists an } i \in \{1, 2, \ldots, \min\{\ell(u), \ell(v)\}\} \\ &\qquad \text{such that } (u_i < v_i \text{ in } \mathfrak{A}, \text{ and every } j \in \{1, 2, \ldots, i-1\} \text{ satisfies } u_j = v_j), \\ &\textbf{or } \text{the word } u \text{ is a prefix of } v \end{aligned}$$

(because $u \leq v$ in $\mathfrak{A}$). In other words, we must be in one of the following two cases:

Case 1: There exists an $i \in \{1, 2, \ldots, \min\{\ell(u), \ell(v)\}\}$ such that ($u_i < v_i$ in $\mathfrak{A}$, and every $j \in \{1, 2, \ldots, i-1\}$ satisfies $u_j = v_j$).

Case 2: The word u is a prefix of v.

Let us first consider Case 1. In this case, there exists an $i \in \{1, 2, \ldots, \min\{\ell(u), \ell(v)\}\}$ such that ($u_i < v_i$ in $\mathfrak{A}$, and every $j \in \{1, 2, \ldots, i-1\}$ satisfies $u_j = v_j$). Let i' be this i. Thus, $i' \in \{1, 2, \ldots, \min\{\ell(u), \ell(v)\}\}$ satisfies ($u_{i'} < v_{i'}$ in $\mathfrak{A}$, and every $j \in \{1, 2, \ldots, i'-1\}$ satisfies $u_j = v_j$).

We have $u_{i'} < v_{i'}$ in $\mathfrak{A}$; thus, $u_{i'} \leq v_{i'}$ in $\mathfrak{A}$ and $u_{i'} \neq v_{i'}$.

Recall that any two elements a and b of $\mathfrak{A}$ satisfying $a \leq b$ in $\mathfrak{A}$ satisfy $a \leq b$ in $\mathfrak{B}$. Applied to $a = u_{i'}$ and $b = v_{i'}$, this yields that $u_{i'} \leq v_{i'}$ in $\mathfrak{B}$ (since $u_{i'} \leq v_{i'}$ in $\mathfrak{A}$). Combined with $u_{i'} \neq v_{i'}$, this yields $u_{i'} < v_{i'}$ in $\mathfrak{B}$. Thus, we have ($u_{i'} < v_{i'}$ in $\mathfrak{B}$, and every $j \in \{1, 2, \ldots, i'-1\}$ satisfies $u_j = v_j$). Consequently, there exists an $i \in \{1, 2, \ldots, \min\{\ell(u), \ell(v)\}\}$

(b) It is easy to prove the logical implication

$$\left(u < v \text{ in } \mathfrak{A}^*\right) \Longrightarrow \left(u < v \text{ in } \mathfrak{B}^* \text{ for every linear extension } \mathfrak{B} \text{ of } \mathfrak{A}\right) \tag{13.146.1}$$

[998]. We shall now focus on proving the implication

$$\left(u < v \text{ in } \mathfrak{B}^* \text{ for every linear extension } \mathfrak{B} \text{ of } \mathfrak{A}\right) \Longrightarrow \left(u < v \text{ in } \mathfrak{A}^*\right). \tag{13.146.2}$$

Proof of (13.146.1): Assume that

$$u < v \text{ in } \mathfrak{B}^* \text{ for every linear extension } \mathfrak{B} \text{ of } \mathfrak{A}. \tag{13.146.3}$$

We need to show that $u < v$ in $\mathfrak{A}^*$.

First of all, it is impossible that $u \geq v$ in $\mathfrak{A}^*$ [999]. Consequently, we have $u \neq v$ (because otherwise, we would have $u = v$ and thus $u \geq v$ in $\mathfrak{A}^*$, which would contradict the fact that it is impossible that $u \geq v$ in $\mathfrak{A}^*$).

Let $m = \min\{\ell(u), \ell(v)\}$. Then, $m \leq \ell(u)$ and $m \leq \ell(v)$. Recall that $u = \left(u_1, u_2, \ldots, u_{\ell(u)}\right)$ and $v = \left(v_1, v_2, \ldots, v_{\ell(v)}\right)$. If every $i \in \{1, 2, \ldots, m\}$ satisfies $u_i = v_i$, then it is easy to see that $u < v$ in $\mathfrak{A}^*$ [1000]. Hence, for the rest of the proof of (13.146.1), we can WLOG assume that not every $i \in \{1, 2, \ldots, m\}$ satisfies $u_i = v_i$. Assume this.

Not every $i \in \{1, 2, \ldots, m\}$ satisfies $u_i = v_i$. In other words, there exists an $i \in \{1, 2, \ldots, m\}$ which does not satisfy $u_i = v_i$. Let k be the smallest such i. Thus, $k \in \{1, 2, \ldots, m\}$ does not satisfy $u_k = v_k$, whereas

$$\text{every } i \in \{1, 2, \ldots, m\} \text{ satisfying } i < k \text{ satisfies } u_i = v_i. \tag{13.146.4}$$

such that ($u_i < v_i$ in $\mathfrak{B}$, and every $j \in \{1, 2, \ldots, i-1\}$ satisfies $u_j = v_j$) (namely, $i = i'$). Thus,

$$\begin{aligned} &\textbf{either} \text{ there exists an } i \in \{1, 2, \ldots, \min\{\ell(u), \ell(v)\}\} \\ &\qquad \text{such that } \left(u_i < v_i \text{ in } \mathfrak{B}, \text{ and every } j \in \{1, 2, \ldots, i-1\} \text{ satisfies } u_j = v_j\right), \\ &\textbf{or} \text{ the word } u \text{ is a prefix of } v. \end{aligned}$$

In other words, $u \leq v$ in $\mathfrak{B}^*$. Thus, $u \leq v$ in $\mathfrak{B}^*$ is proven in Case 1.

Let us now consider Case 2. In this case, the word u is a prefix of v. Hence, $u \leq v$ in $\mathfrak{B}^*$. Thus, $u \leq v$ in $\mathfrak{B}^*$ is proven in Case 2.

Now, $u \leq v$ in $\mathfrak{B}^*$ is proven in each of the two Cases 1 and 2. Since these two Cases cover all possibilities, this yields that $u \leq v$ in $\mathfrak{B}^*$ always holds.

Now, let us forget that we fixed u and v. We have thus shown that any two elements u and v of $\mathfrak{A}^*$ satisfying $u \leq v$ in $\mathfrak{A}^*$ satisfy $u \leq v$ in $\mathfrak{B}^*$. Renaming the variables u and v as a and b in this statement, we conclude the following: Any two elements a and b of $\mathfrak{A}^*$ satisfying $a \leq b$ in $\mathfrak{A}^*$ satisfy $a \leq b$ in $\mathfrak{B}^*$. Qed.

[998] *Proof of (13.146.1):* Assume that $u < v$ in $\mathfrak{A}^*$. Thus, $u \neq v$ and $u \leq v$ in $\mathfrak{A}^*$. Let $\mathfrak{B}$ be a linear extension of $\mathfrak{A}$. Then, $\mathfrak{B}^*$ is an extension of the poset $\mathfrak{A}^*$ (by Proposition 13.146.3(a)). Hence, $u \leq v$ in $\mathfrak{B}^*$ (since $u \leq v$ in $\mathfrak{A}^*$). Combined with $u \neq v$, this yields $u < v$ in $\mathfrak{B}^*$.

Now, let us forget that we fixed $\mathfrak{B}$. We thus have proven that $u < v$ in $\mathfrak{B}^*$ for every linear extension $\mathfrak{B}$ of $\mathfrak{A}$. This proves (13.146.1).

[999] *Proof.* Assume the contrary. Thus, $u \geq v$ in $\mathfrak{A}^*$. Hence, $v \leq u$ in $\mathfrak{A}^*$.

Proposition 13.146.2(c) (applied to $P = \mathfrak{A}^*$) yields that there exists a linear extension of $\mathfrak{A}$. Let $\mathfrak{B}$ be such a linear extension. Then, $u < v$ in $\mathfrak{B}^*$ (by (13.146.3)). But $\mathfrak{B}^*$ is an extension of the poset $\mathfrak{A}^*$ (by Proposition 13.146.2(a)). Thus, $v \leq u$ in $\mathfrak{B}^*$ (since $v \leq u$ in $\mathfrak{A}^*$). This contradicts $u < v$ in $\mathfrak{B}^*$. This contradiction proves that our assumption was wrong, qed.

[1000] *Proof.* Assume that every $i \in \{1, 2, \ldots, m\}$ satisfies $u_i = v_i$. Thus, $(u_1, u_2, \ldots, u_m) = (v_1, v_2, \ldots, v_m)$.

Let us first assume (for the sake of contradiction) that $\ell(u) \geq \ell(v)$. Then, $m = \min\{\ell(u), \ell(v)\} = \ell(v)$ (since $\ell(u) \geq \ell(v)$). Now,

$$\begin{aligned} v &= \left(v_1, v_2, \ldots, v_{\ell(v)}\right) = (v_1, v_2, \ldots, v_m) \qquad (\text{since } \ell(v) = m) \\ &= (u_1, u_2, \ldots, u_m). \end{aligned}$$

But the word $(u_1, u_2, \ldots, u_m)$ is clearly a prefix of $\left(u_1, u_2, \ldots, u_{\ell(u)}\right)$ (since $m \leq \ell(u)$). In other words, the word v is a prefix of u (since $v = (u_1, u_2, \ldots, u_m)$ and $u = \left(u_1, u_2, \ldots, u_{\ell(u)}\right)$). Hence, $v \leq u$ in $\mathfrak{A}^*$. In other words, $u \geq v$ in $\mathfrak{A}^*$. This contradicts the fact that it is impossible that $u \geq v$ in $\mathfrak{A}^*$.

This contradiction proves that our assumption (that $\ell(u) \geq \ell(v)$) was wrong. Hence, we cannot have $\ell(u) \geq \ell(v)$. We thus have $\ell(u) < \ell(v)$. Thus, $m = \min\{\ell(u), \ell(v)\} = \ell(u)$ (since $\ell(u) < \ell(v)$). Now,

$$\begin{aligned} u &= \left(u_1, u_2, \ldots, u_{\ell(v)}\right) = (u_1, u_2, \ldots, u_m) \qquad (\text{since } \ell(u) = m) \\ &= (v_1, v_2, \ldots, v_m). \end{aligned}$$

But the word $(v_1, v_2, \ldots, v_m)$ is clearly a prefix of $\left(v_1, v_2, \ldots, v_{\ell(v)}\right)$ (since $m \leq \ell(v)$). In other words, the word u is a prefix of v (since $u = (v_1, v_2, \ldots, v_m)$ and $v = \left(v_1, v_2, \ldots, v_{\ell(v)}\right)$). Hence, $u \leq v$ in $\mathfrak{A}^*$. Combined with $u \neq v$, this yields that $u < v$ in $\mathfrak{A}^*$, qed.

We have $u_k \neq v_k$ (since k does not satisfy $u_k = v_k$). We must have $v_k \geq u_k$ in $\mathfrak{A}$ [1001]. That is, $u_k \leq v_k$ in $\mathfrak{A}$. Combined with $u_k \neq v_k$, this yields that $u_k < v_k$ in $\mathfrak{A}$.

Every $j \in \{1, 2, \ldots, k-1\}$ satisfies $u_j = v_j$ (by (13.146.4), applied to $i = j$). Also, we have $k \in \{1, 2, \ldots, m\} = \{1, 2, \ldots, \min\{\ell(u), \ell(v)\}\}$ (since $m = \min\{\ell(u), \ell(v)\}$). Altogether, we thus have shown that $k \in \{1, 2, \ldots, \min\{\ell(u), \ell(v)\}\}$ satisfies ($u_k < v_k$ in $\mathfrak{A}$, and every $j \in \{1, 2, \ldots, k-1\}$ satisfies $u_j = v_j$). Thus, there exists an $i \in \{1, 2, \ldots, \min\{\ell(u), \ell(v)\}\}$ such that
($u_i < v_i$ in $\mathfrak{A}$, and every $j \in \{1, 2, \ldots, i-1\}$ satisfies $u_j = v_j$) (namely, $i = k$). Hence,

either there exists an $i \in \{1, 2, \ldots, \min\{\ell(u), \ell(v)\}\}$
such that ($u_i < v_i$ in $\mathfrak{A}$, and every $j \in \{1, 2, \ldots, i-1\}$ satisfies $u_j = v_j$),
or the word u is a prefix of v.

In other words, $u < v$ in $\mathfrak{A}^*$ (by the definition of the relation $<$ on $\mathfrak{A}^*$). This proves (13.146.2).

Combining the implications (13.146.1) and (13.146.2), we obtain the following equivalence of statements:

$$(u < v \text{ in } \mathfrak{A}^*) \iff (u < v \text{ in } \mathfrak{B}^* \text{ for every linear extension } \mathfrak{B} \text{ of } \mathfrak{A}).$$

Thus, Proposition 13.146.3(b) is proven. □

Let us now give a new proof of the fact that Corollary 6.1.6 holds in the partial-order setting:

Alternative proof of Corollary 6.1.6 in the partial-order setting. We need to prove that $uw \geq wu$. If $uw = wu$, then this is obvious. Hence, for the rest of this proof, we can WLOG assume that $uw \neq wu$. Assume this.

Let $\mathfrak{B}$ be any linear extension of the alphabet $\mathfrak{A}$. Then, $\mathfrak{B}^*$ is an extension of $\mathfrak{A}^*$ (according to Proposition 13.146.3(a)). Hence, $uv \geq vu$ in $\mathfrak{B}^*$ (since $uv \geq vu$ in $\mathfrak{A}^*$) and $vw \geq wv$ in $\mathfrak{B}^*$ (since $vw \geq wv$ in $\mathfrak{A}^*$). But the alphabet $\mathfrak{B}$ is totally ordered, and thus Corollary 6.1.6 (applied to $\mathfrak{B}$ instead of $\mathfrak{A}$) yields that $uw \geq wu$ in $\mathfrak{B}^*$ (since we know that Corollary 6.1.6 holds in the total-order setting). Thus, $uw > wu$ in $\mathfrak{B}^*$ (since $uw \neq wu$). In other words, $wu < uw$ in $\mathfrak{B}^*$.

Now, let us forget that we fixed $\mathfrak{B}$. We thus have shown that $wu < uw$ in $\mathfrak{B}^*$ for every linear extension $\mathfrak{B}$ of $\mathfrak{A}$. But Proposition 13.146.3(b) (applied to wu and uw instead of u and v) yields that $wu < uw$ holds in $\mathfrak{A}^*$ if and only if

$$(wu < uw \text{ in } \mathfrak{B}^* \text{ for every linear extension } \mathfrak{B} \text{ of } \mathfrak{A}).$$

Thus, we conclude that $wu < uw$ holds in $\mathfrak{A}^*$ (since we already know that $wu < uw$ in $\mathfrak{B}^*$ for every linear extension $\mathfrak{B}$ of $\mathfrak{A}$). Thus, $uw > wu$ in $\mathfrak{A}^*$; hence, $uw \geq wu$ in $\mathfrak{A}^*$. This proves Corollary 6.1.6 in the partial-order setting. □

Next, let us give a new proof of the fact that Theorem 6.1.20 holds in the partial-order setting:

Alternative proof of Theorem 6.1.20 in the partial-order setting. We need to prove the equivalence $\mathcal{A} \iff \mathcal{B} \iff \mathcal{C} \iff \mathcal{D}$. We can prove the implications $\mathcal{A} \Longrightarrow \mathcal{B}$, $\mathcal{A} \Longrightarrow \mathcal{C}$, $\mathcal{A} \Longrightarrow \mathcal{D}$ and $\mathcal{B} \Longrightarrow \mathcal{A}$ in the same way as we did in the total-order setting. Hence, in order to prove Theorem 6.1.20, it will be enough to prove the implications $\mathcal{C} \Longrightarrow \mathcal{B}$ and $\mathcal{D} \Longrightarrow \mathcal{B}$.

Proof of the implication $\mathcal{C} \Longrightarrow \mathcal{B}$: Assume that Assertion $\mathcal{C}$ holds.

[1001] *Proof.* Assume the contrary. Thus, we don't have $v_k \geq u_k$ in $\mathfrak{A}$. Hence, Proposition 13.146.2(d) (applied to $P = \mathfrak{A}^*$, $a = v_k$ and $b = u_k$) yields that there exists a linear extension Q of $\mathfrak{A}$ such that we have $v_k < u_k$ in Q. Let $\mathfrak{B}$ be such a linear extension. Thus, $\mathfrak{B}$ is a linear extension of $\mathfrak{A}$ such that we have $v_k < u_k$ in $\mathfrak{B}$. We have $u < v$ in $\mathfrak{B}^*$ (by (13.146.3)).

Every $j \in \{1, 2, \ldots, k-1\}$ satisfies $u_j = v_j$ (by (13.146.4), applied to $i = j$). In other words, every $j \in \{1, 2, \ldots, k-1\}$ satisfies $v_j = u_j$. Also, we have $k \in \{1, 2, \ldots, m\} = \{1, 2, \ldots, \min\{\ell(v), \ell(u)\}\}$ (since $m = \min\{\ell(u), \ell(v)\} = \min\{\ell(v), \ell(u)\}$). Altogether, we thus have shown that $k \in \{1, 2, \ldots, \min\{\ell(v), \ell(u)\}\}$ satisfies ($v_k < u_k$ in $\mathfrak{B}$, and every $j \in \{1, 2, \ldots, k-1\}$ satisfies $v_j = u_j$). Thus, there exists an $i \in \{1, 2, \ldots, \min\{\ell(v), \ell(u)\}\}$ such that
($v_i < u_i$ in $\mathfrak{B}$, and every $j \in \{1, 2, \ldots, i-1\}$ satisfies $v_j = u_j$) (namely, $i = k$). Hence,

either there exists an $i \in \{1, 2, \ldots, \min\{\ell(v), \ell(u)\}\}$
such that ($v_i < u_i$ in $\mathfrak{B}$, and every $j \in \{1, 2, \ldots, i-1\}$ satisfies $v_j = u_j$),
or the word v is a prefix of u.

In other words, $v < u$ in $\mathfrak{B}^*$ (by the definition of the relation $<$ on $\mathfrak{B}^*$). This contradicts the fact that $u < v$ in $\mathfrak{B}^*$. This contradiction proves that our assumption was wrong, qed.

Let $\mathfrak{B}$ be any linear extension of the alphabet $\mathfrak{A}$. Then, $\mathfrak{B}^*$ is an extension of $\mathfrak{A}^*$ (according to Proposition 13.146.3(a)). It is easy to see that Assertion $\mathcal{C}$ (with $\mathfrak{A}$ replaced by $\mathfrak{B}$) holds[1002]. But we can apply Theorem 6.1.20 to $\mathfrak{B}$ instead of $\mathfrak{A}$ (since $\mathfrak{B}$ is totally ordered). As a consequence, we see that the four assertions $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ and $\mathcal{D}$ (all with $\mathfrak{A}$ replaced by $\mathfrak{B}$) are equivalent. In particular, Assertion $\mathcal{C}$ (with $\mathfrak{A}$ replaced by $\mathfrak{B}$) is equivalent to Assertion $\mathcal{B}$ (with $\mathfrak{A}$ replaced by $\mathfrak{B}$). Since Assertion $\mathcal{C}$ (with $\mathfrak{A}$ replaced by $\mathfrak{B}$) holds, we thus conclude that Assertion $\mathcal{B}$ (with $\mathfrak{A}$ replaced by $\mathfrak{B}$) holds. In other words,

$$\text{any two nonempty words } u \in \mathfrak{B}^* \text{ and } v \in \mathfrak{B}^* \text{ satisfying } w = uv \text{ satisfy } v > w \text{ in } \mathfrak{B}^*. \tag{13.146.5}$$

Now, let $u \in \mathfrak{A}^*$ and $v \in \mathfrak{A}^*$ be two nonempty words satisfying $w = uv$. It is easy to see that

$$(w < v \text{ in } \mathfrak{B}^* \text{ for every linear extension } \mathfrak{B} \text{ of } \mathfrak{A}) \tag{13.146.6}$$

[1003].

But Proposition 13.146.3(b) (applied to w instead of u) yields that $w < v$ holds in $\mathfrak{A}^*$ if and only if we have

$$(w < v \text{ in } \mathfrak{B}^* \text{ for every linear extension } \mathfrak{B} \text{ of } \mathfrak{A}).$$

Consequently, we conclude that $w < v$ holds in $\mathfrak{A}^*$ (since we know that we have $(w < v \text{ in } \mathfrak{B}^* \text{ for every linear extension } \mathfrak{B} \text{ of } \mathfrak{A})$). In other words, $v > w$ in $\mathfrak{A}^*$.

Now, let us forget that we fixed u and v. We have thus shown that any nonempty words $u \in \mathfrak{A}^*$ and $v \in \mathfrak{A}^*$ satisfying $w = uv$ satisfy $v > w$ in $\mathfrak{A}^*$. In other words, Assertion $\mathcal{B}$ holds. Thus, we have proven the implication $\mathcal{C} \Longrightarrow \mathcal{B}$.

Proof of the implication $\mathcal{D} \Longrightarrow \mathcal{B}$: The proof of the implication $\mathcal{D} \Longrightarrow \mathcal{B}$ is analogous to our above proof of the implication $\mathcal{C} \Longrightarrow \mathcal{B}$, and thus left to the reader.

The proof of Theorem 6.1.20 in the partial-order setting is complete. □

The examples of Theorem 6.1.20 and Corollary 6.1.6 should have illustrated how Proposition 13.146.3 allows deriving certain facts about the partial-order setting from the corresponding facts about the total-order setting. This trick, however, has its limits. For example, in the total-order setting, the fact that any Lyndon word w of length > 1 can be written in the form $w = uv$ for two Lyndon words u and v satisfying $u < w < v$ is a consequence of Theorem 6.1.30. But in the partial-order setting, it is not clear how to derive it from the total-order setting, although it **is** true in the partial-order setting (and follows from Exercise 6.1.31).]

(d) If $\mathfrak{A}$ is the partially ordered alphabet $\{1, 2\}$ with no relations whatsoever (i.e., a 2-element antichain), and w is the word 12, then w does satisfy (if $w \leq t^n$, then $w \leq t$) for every nonempty word t and every positive integer n, but w is not Lyndon.

[*Remark:* One direction of Exercise 6.1.22 does hold in the partial-order setting: Namely, if w is Lyndon, then every nonempty word t and every positive integer n satisfy (if $w \leq t^n$, then $w \leq t$). The proof of this is the same as in the total-order setting.]

(e) The following statement is clearly equivalent to Exercise 6.1.22 in the total-order setting, while still being valid in the partial-order setting:

Proposition 13.146.4. *Let w be a nonempty word. Then, w is Lyndon if and only if every nonempty word t and every positive integer n satisfy (if $w > t$, then $w > t^n$).*

Proof of Proposition 13.146.4 in the partial-order setting. Let us first assume that w is Lyndon. We shall prove that

$$(\text{every nonempty word } t \text{ and every positive integer } n \text{ satisfy (if } w > t, \text{ then } w > t^n)). \tag{13.146.7}$$

[1002] *Proof.* Let $u \in \mathfrak{B}^*$ and $v \in \mathfrak{B}^*$ be two nonempty words satisfying $w = uv$. We have $u \in \mathfrak{B}^* = \mathfrak{A}^*$ and $v \in \mathfrak{B}^* = \mathfrak{A}^*$, and thus $u \in \mathfrak{A}^*$ and $v \in \mathfrak{A}^*$ are two nonempty words satisfying $w = uv$. Hence, we have $v > u$ in $\mathfrak{A}^*$ (since we assumed that Assertion $\mathcal{C}$ holds). Thus, $u < v$ in $\mathfrak{A}^*$, so that $u \neq v$ and $u \leq v$ in $\mathfrak{A}^*$.

So we have $u \leq v$ in $\mathfrak{A}^*$. Hence, $u \leq v$ in $\mathfrak{B}^*$ (since $\mathfrak{B}^*$ is an extension of $\mathfrak{A}^*$). Since $u \neq v$, this becomes $u < v$ in $\mathfrak{B}^*$. In other words, $v > u$ in $\mathfrak{B}^*$.

Now, let us forget that we fixed u and v. We thus have shown that any nonempty words $u \in \mathfrak{B}^*$ and $v \in \mathfrak{B}^*$ satisfying $w = uv$ satisfy $v > u$ in $\mathfrak{B}^*$. In other words, Assertion $\mathcal{C}$ (with $\mathfrak{A}$ replaced by $\mathfrak{B}$) holds

[1003] *Proof.* Let $\mathfrak{B}$ be a linear extension of $\mathfrak{A}$. Then, $\mathfrak{B}^*$ is an extension of $\mathfrak{A}^*$ (according to Proposition 13.146.3(a)). Thus, $\mathfrak{B}^* = \mathfrak{A}^*$ as sets, so that $u \in \mathfrak{A}^* = \mathfrak{B}^*$ and $v \in \mathfrak{A}^* = \mathfrak{B}^*$. Now, (13.146.5) yields $v > w$ in $\mathfrak{B}^*$. In other words, $w < v$ in $\mathfrak{B}^*$. This proves (13.146.6).

Let t be a nonempty word, and let n be a positive integer. Assume that $w > t$. We need to prove that $w > t^n$.

Assume the contrary. Thus, we don't have $w > t^n$. Thus, there exists an $i \in \{1, 2, \ldots, n\}$ such that we don't have $w > t^i$ (namely, $i = n$). Let m be the **minimal** such i. Thus, $m \in \{1, 2, \ldots, n\}$, and we don't have $w > t^m$. Hence, $m \neq 1$ (since we don't have $w > t^m$, but we do have $w > t = t^1$). Thus, $m \geq 2$, so that $m - 1$ is also an element of $\{1, 2, \ldots, n\}$. If we did not have $w > t^{m-1}$, then $m - 1$ would therefore be an $i \in \{1, 2, \ldots, n\}$ such that we don't have $w > t^i$. But this would contradict the fact that m is the **minimal** such i. Thus, we must have $w > t^{m-1}$. In other words, $t^{m-1} < w$.

Notice that $m - 1 \geq 1$ (since $m \geq 2$). Hence, the word t^{m-1} is nonempty (since t is nonempty). In other words, $t^{m-1} \neq \varnothing$.

Recall that we don't have $w > t^m$. It is now easy to see that we don't have $t^m \leq w$ [1004]. In other words, we don't have $t^{m-1}t \leq w\varnothing$ (since $t^m = t^{m-1}t$ and $w = w\varnothing$).

But recall that $t^{m-1} < w$. Hence, Proposition 6.1.2(d) (applied to $a = t^{m-1}$, $b = t$, $c = w$ and $d = \varnothing$) yields that either we have $t^{m-1}t \leq w\varnothing$ or the word t^{m-1} is a prefix of w. Thus, the word t^{m-1} is a prefix of w (since we don't have $t^{m-1}t \leq w\varnothing$). In other words, there exists a $v \in \mathfrak{A}^*$ such that $w = t^{m-1}v$. Consider this v. We have $v \neq \varnothing$ (because otherwise, we would have $v = \varnothing$ and thus $w = t^{m-1}\underbrace{v}_{=\varnothing} = t^{m-1}$, contradicting the fact that $w > t^{m-1}$), so that v is nonempty. Hence, Proposition 6.1.14(b) (applied to $u = t^{m-1}$) now yields $v > t^{m-1}$. Hence, $t^{m-1} < v$. Thus, Proposition 6.1.2(b) (applied to $a = t^{m-1}$, $c = t^{m-1}$ and $d = v$) yields $t^{m-1}t^{m-1} \leq t^{m-1}v = w$. Hence,

$$\begin{aligned} w &\geq t^{m-1}t^{m-1} = t^{2(m-1)} = t^{m+(m-2)} = t^m t^{m-2} \qquad \text{(this makes sense since } m \geq 2\text{)} \\ &\geq t^m \qquad \left(\text{since } t^m \text{ is a prefix of } t^m t^{m-2}\right). \end{aligned}$$

Hence, $t^m \leq w$. This contradicts the fact that we don't have $t^m \leq w$. This contradiction proves that our assumption (that we don't have $w > t^n$) was false. Hence, $w > t^n$.

Forget now that we assumed that $w > t$. We thus have proven that if $w > t$, then $w > t^n$.

Now, forget that we fixed t and assumed that w is Lyndon. We thus have shown that

$$\text{(if } w \text{ is Lyndon, then (13.146.7) holds)}. \tag{13.146.8}$$

Now, conversely, assume that (13.146.7) holds. We will prove that w is Lyndon.

Let u and v be any nonempty words satisfying $w = uv$. We have $w \neq u$ [1005]. Combined with $w = uv \geq u$ (since u is a prefix of uv), this yields $w > u$. But (13.146.7) (applied to $t = u$ and $n = 2$) yields that if $w > u$, then $w > u^2$. Hence, $w > u^2$ (since we know that $w > u$). Thus, $uv = w > u^2 = uu$, so that $uu < uv$. Thus, Proposition 6.1.2(c) (applied to $a = u$, $c = u$ and $d = v$) yields $u \leq v$. Since $u \neq v$ [1006], this becomes $u < v$, so that $v > u$.

Now, let us forget that we fixed u and v. We thus have proven that any nonempty words u and v satisfying $w = uv$ satisfy $v > u$. In other words, Assertion $\mathcal{C}$ of Theorem 6.1.20 holds. Hence, Assertion $\mathcal{A}$ of Theorem 6.1.20 holds as well (since Theorem 6.1.20 yields that these Assertions $\mathcal{C}$ and $\mathcal{A}$ are equivalent). In other words, the word w is Lyndon.

Now, forget that we fixed w. We thus have proven that

$$\text{(if (13.146.7) holds, then } w \text{ is Lyndon)}.$$

Combined with (13.146.8), this yields that w is Lyndon if and only if (13.146.7) holds. This proves Proposition 13.146.4. □

[1004] *Proof.* Assume the contrary. Thus, we have $t^m \leq w$. Hence, $w \geq t^m$, so that $w = t^m$ (since we don't have $w > t^m$). Thus, $w = t^m = t^{m-1}t$. Thus, t is a suffix of w. Thus, Corollary 6.1.15 (applied to t instead of v) yields $t \geq w$. Thus, $t \geq w = t^m = tt^{m-1}$. Combined with $tt^{m-1} \geq t$ (since t is a prefix of tt^{m-1}), this yields $t = tt^{m-1}$. Thus, $tt^{m-1} = t = t\varnothing$. Cancelling t from this equality, we obtain $t^{m-1} = \varnothing$. This contradicts $t^{m-1} \neq \varnothing$. This contradiction proves that our assumption was wrong, qed.

[1005] *Proof.* Assume the contrary. Then, $w = u$. Hence, $uv = w = u = u\varnothing$. Cancelling u from this equality, we obtain $v = \varnothing$, so that v is empty. This contradicts the fact that v is nonempty. This contradiction proves that our assumption was wrong, qed.

[1006] *Proof.* Assume the contrary. Then, $u = v$. Hence, $w = u\underbrace{v}_{=u} = uu = u^2$. This contradicts $w > u^2$. This contradiction proves that our assumption was wrong, qed.

Thus, we have salvaged Exercise 6.1.22 in the partial-order setting. (And, as a side effect, we have obtained an alternative solution for Exercise 6.1.22 in the total-order setting.)

[*Remark:* Speaking of salvaging, there is a little piece of Proposition 6.1.19 which can be salvaged for the partial-order case:

Proposition 13.146.5. *Let w be a nonempty word. Let v be a (lexicographically) minimal nonempty suffix of w* [1007]*. Then, there exists a $u \in \mathfrak{A}^*$ such that $w = uv$ but we don't have $uv < vu$.*

Proof of Proposition 13.146.5 in the partial-order case. We know that v is a suffix of w. Hence, there exists a $u \in \mathfrak{A}^*$ such that $w = uv$. Consider this u. It will clearly be enough to show that we don't have $uv < vu$.

So let us prove that we don't have $uv < vu$. Indeed, assume the contrary. Then, $uv < vu$. Thus, there exists at least one suffix t of u such that $tv < vt$ (namely, $t = u$). Let p be the **minimum-length** such suffix. Then, $pv < vp$. Thus, p is nonempty. In other words, $p \neq \varnothing$.

Since p is a suffix of u, it is clear that pv is a suffix of $uv = w$. So we know that pv is a nonempty suffix of w. Since v is a minimal such suffix, this yields that we don't have $pv < v$. Hence, we don't have $pv \leq v$ [1008].

But $pv\varnothing = pv < vp$. Hence, Proposition 6.1.2(e) (applied to $a = pv$, $b = \varnothing$, $c = v$ and $d = p$) yields that either we have $pv \leq v$ or the word v is a prefix of pv. Thus, the word v is a prefix of pv (since we don't have $pv \leq v$). In other words, there exists a $q \in \mathfrak{A}^*$ such that $pv = vq$. Consider this q. This q is nonempty (because otherwise we would have $pv = v \underbrace{q}_{=\varnothing} = v$, contradicting the fact that p is nonempty). From $vq = pv < vp$, we obtain $q \leq p$ (by Proposition 6.1.2(c), applied to $a = v$, $c = q$ and $d = p$).

We know that q is a suffix of pv (since $vq = pv$), whereas pv is a suffix of w. Thus, q is a suffix of w. So q is a nonempty suffix of w. Since v is a minimal such suffix, this yields that we don't have $q < v$. But we have $p\varnothing = p \leq pv < vp$. Hence, Proposition 6.1.2(e) (applied to $a = p$, $b = \varnothing$, $c = v$ and $d = p$) yields that either we have $p \leq v$ or the word v is a prefix of p. From this, it is easy to obtain that v is a prefix of p [1009]. In other words, there exists an $r \in \mathfrak{A}^*$ such that $p = vr$. Consider this r. Clearly, r is a suffix of p, while p is a suffix of u; therefore, r is a suffix of u. Also, $pv < vp$ rewrites as $vrv < vvr$ (because $p = vr$). Thus, Proposition 6.1.2(c) (applied to $a = v$, $c = rv$ and $d = vr$) yields $rv \leq vr$. Since $rv \neq vr$ (because otherwise, we would have $rv = vr$, thus $v \underbrace{rv}_{=vr} = vvr$, contradicting $vrv < vvr$), this becomes $rv < vr$.

The word v is nonempty; thus, $\ell(v) > 0$.

Now, r is a suffix of u such that $rv < vr$. Since p is the minimum-length such suffix, this yields $\ell(r) \geq \ell(p)$. But this contradicts the fact that $\ell\left(\underbrace{p}_{=vr}\right) = \ell(vr) = \underbrace{\ell(v)}_{>0} + \ell(r) > \ell(r)$. This contradiction proves our assumption wrong. Thus, we have proven that we don't have $uv < vu$. This completes the proof of Proposition 13.146.5 in the partial-order case. □

]

(f) Let us first establish a lemma which plays a role similar to that of Lemma 6.1.28 in the total-order setting:

Lemma 13.146.6. *Let $(a_1, a_2, \ldots, a_k)$ be a Hazewinkel-CFL factorization of a nonempty word w (in the partial-order setting). Let p be a suffix of w such that p is Lyndon. Then, $p \geq a_k$.*

[1007] By this, we mean that:

- v is a nonempty suffix of w;
- no nonempty suffix s of w satisfies $s < v$.

Such a v always exists (since w is nonempty), but is not always unique.

[1008] *Proof.* Assume the contrary. Thus, $pv \leq v$. This yields that $pv = v$ (since we don't have $pv < v$). Thus, $pv = v = \varnothing v$. Cancelling v from this equation, we obtain $p = \varnothing$. This contradicts $p \neq \varnothing$. This contradiction proves that our assumption was wrong, qed.

[1009] *Proof.* Assume the contrary. Thus, v is not a prefix of p. Hence, $p \leq v$ (since either we have $p \leq v$ or the word v is a prefix of p). Now, $q \leq p \leq v$. Now, $q = v$ (since $q \leq v$ but not $q < v$). Hence, $v = q \leq p$. Combined with $p \leq v$, this yields $v = p$. Hence, $p \underbrace{v}_{=p} = \underbrace{p}_{=v} p = vp$, which contradicts $pv < vp$. This contradiction proves that our assumption was wrong, qed.

Proof of Lemma 13.146.6. We will prove Lemma 13.146.6 by induction over the (obviously) positive integer k.

Induction base: Assume that $k = 1$. By the definition of a Hazewinkel-CFL factorization, $(a_1, a_2, \ldots, a_k)$ is a tuple of Lyndon words satisfying $w = a_1 a_2 \cdots a_k$ (since $(a_1, a_2, \ldots, a_k)$ is a Hazewinkel-CFL factorization of w). Thus, $w = a_1 a_2 \cdots a_k = a_1$ (since $k = 1$), so that w is a Lyndon word (since a_1 is a Lyndon word). But p is nonempty (since p is Lyndon). Thus, Corollary 6.1.15 (applied to $v = p$) yields $p \geq w = a_1 = a_k$ (since $1 = k$). Thus, Lemma 13.146.6 is proven in the case $k = 1$. The induction base is complete.

Induction step: Let K be a positive integer. Assume (as the induction hypothesis) that Lemma 13.146.6 is proven for $k = K$. We now need to show that Lemma 13.146.6 holds for $k = K + 1$.

So let $(a_1, a_2, \ldots, a_{K+1})$ be a Hazewinkel-CFL factorization of a nonempty word w. Let p be a nonempty suffix of w such that p is Lyndon. We need to prove that $p \geq a_{K+1}$.

The tuple $(a_1, a_2, \ldots, a_{K+1})$ is a Hazewinkel-CFL factorization of w. By the definition of a Hazewinkel-CFL factorization, this yields that $(a_1, a_2, \ldots, a_{K+1})$ is a tuple of Lyndon words such that $w = a_1 a_2 \cdots a_{K+1}$ and such that no $i \in \{1, 2, \ldots, K\}$ satisfies $a_i < a_{i+1}$.

Let $w' = a_2 a_3 \cdots a_{K+1}$; then, $w = a_1 a_2 \cdots a_{K+1} = a_1 \underbrace{(a_2 a_3 \cdots a_{K+1})}_{=w'} = a_1 w'$. Hence, every nonempty suffix of w is either a nonempty suffix of w', or has the form qw' for a nonempty suffix q of a_1. Since p is a nonempty suffix of w, we thus must be in one of the following two cases:

Case 1: The word p is a nonempty suffix of w'.

Case 2: The word p has the form qw' for a nonempty suffix q of a_1.

Let us first consider Case 1. In this case, p is a nonempty suffix of w'. The K-tuple $(a_2, a_3, \ldots, a_{K+1})$ of Lyndon words satisfies $w' = a_2 a_3 \cdots a_{K+1}$ and has the property that no $i \in \{1, 2, \ldots, K-1\}$ satisfies $a_{i+1} < a_{(i+1)+1}$ [1010]. Therefore, $(a_2, a_3, \ldots, a_{K+1})$ is a Hazewinkel-CFL factorization of w'. We can thus apply Lemma 13.146.6 to K, w' and $(a_2, a_3, \ldots, a_{K+1})$ instead of k, w and $(a_1, a_2, \ldots, a_k)$ (because we assumed that Lemma 13.146.6 is proven for $k = K$). As a result, we obtain that $p \geq a_{K+1}$. Thus, $p \geq a_{K+1}$ is proven in Case 1.

Let us now consider Case 2. In this case, p has the form qw' for a nonempty suffix q of a_1. Consider this q. Since a_1 is a Lyndon word, we have $q \geq a_1$ (by Corollary 6.1.15, applied to a_1 and q instead of w and v). Our goal, however, is to prove that $q \geq a_{K+1}$.

We will show that

$$q \geq a_i \qquad \text{for every } i \in \{1, 2, \ldots, K+1\}. \tag{13.146.9}$$

Proof of (13.146.9): We will prove (13.146.9) by induction over i:

Induction base: We know that $q \geq a_1$. In other words, (13.146.9) holds for $i = 1$. This completes the induction base.

Induction step: Let $j \in \{1, 2, \ldots, K\}$. Assume that (13.146.9) holds for $i = j$. We must prove that (13.146.9) holds for $i = j + 1$.

We have $j \leq K$. Hence, the product $a_{j+1} a_{j+2} \cdots a_{K+1}$ contains at least one factor. Also, the factors of this product are nonempty words (because the words a_{j+1}, a_{j+2}, $\ldots$, a_{K+1} are nonempty (since these words are Lyndon)). Hence, $a_{j+1} a_{j+2} \cdots a_{K+1}$ is a nonempty product of nonempty words. Consequently, $a_{j+1} a_{j+2} \cdots a_{K+1}$ is a nonempty word.

The product $a_{j+2} a_{j+3} \cdots a_{K+1}$ is well-defined (because $j \leq K$). Denote this product by g. Thus, $g = a_{j+2} a_{j+3} \cdots a_{K+1}$.

We have $q \geq a_j$ (since (13.146.9) holds for $i = j$). But

$$p = q \underbrace{\underbrace{w'}_{= a_2 a_3 \cdots a_{K+1}}}_{= (a_2 a_3 \cdots a_j)(a_{j+1} a_{j+2} \cdots a_{K+1})} = q (a_2 a_3 \cdots a_j)(a_{j+1} a_{j+2} \cdots a_{K+1}) = (q (a_2 a_3 \cdots a_j))(a_{j+1} a_{j+2} \cdots a_{K+1}).$$

[1010] *Proof.* We already know that $w' = a_2 a_3 \cdots a_{K+1}$. It remains to show that no $i \in \{1, 2, \ldots, K-1\}$ satisfies $a_{i+1} < a_{(i+1)+1}$.

To prove this, let us assume this contrary. Thus, there exists an $i \in \{1, 2, \ldots, K-1\}$ satisfying $a_{i+1} < a_{(i+1)+1}$. Let j be such an i. Then, $j \in \{1, 2, \ldots, K-1\}$ satisfies $a_{j+1} < a_{(j+1)+1}$. Hence, we have $a_i < a_{i+1}$ for $i = j + 1$. This contradicts the fact that no $i \in \{1, 2, \ldots, K\}$ satisfies $a_i < a_{i+1}$. This contradiction shows that our assumption was wrong, qed.

Consequently, $a_{j+1}a_{j+2}\cdots a_{K+1}$ is a suffix of p. Thus, Corollary 6.1.15 (applied to p and $a_{j+1}a_{j+2}\cdots a_{K+1}$ instead of w and v) yields $a_{j+1}a_{j+2}\cdots a_{K+1} \geq p$ (since p is Lyndon). Thus,

$$a_{j+1}a_{j+2}\cdots a_{K+1} \geq p = qw' \geq q \geq a_j.$$

In other words,

$$a_j \leq a_{j+1}a_{j+2}\cdots a_{K+1} = a_{j+1}\underbrace{(a_{j+2}a_{j+3}\cdots a_{K+1})}_{=g} = a_{j+1}g.$$

Hence, $a_j\varnothing = a_j \leq a_{j+1}g$. Proposition 6.1.2(e) (applied to a_j, $\varnothing$, a_{j+1} and g instead of a, b, c and d) thus yields that either we have $a_j \leq a_{j+1}$ or the word a_{j+1} is a prefix of a_j. From this, it is easy to conclude that the word a_{j+1} is a prefix of a_j [1011]. Consequently, $a_{j+1} \leq a_j$, so that $a_j \geq a_{j+1}$ and $q \geq a_j \geq a_{j+1}$. In other words, (13.146.9) holds for $i = j+1$. This completes the induction step. Thus, (13.146.9) is proven by induction.

Now, (13.146.9) (applied to $i = K+1$) yields $q \geq a_{K+1}$. Hence, $p = qw' \geq q \geq a_{K+1}$. Thus, $p \geq a_{K+1}$ is proven in Case 2.

We have now proven $p \geq a_{K+1}$ in all cases. This proves that Lemma 13.146.6 holds for $k = K+1$. The induction step is thus finished, and with it the proof of Lemma 13.146.6. □

We can now conclude the solution of Exercise 6.1.33(f) by proving the following proposition:

Proposition 13.146.7. *Let w be a word (in the partial-order setting). Then, there exists a unique Hazewinkel-CFL factorization of w.*

The proof is an almost literal adaptation of the proof of Theorem 6.1.27:

Proof of Proposition 13.146.7. Let us first prove that there exists a Hazewinkel-CFL factorization of w.

Indeed, there clearly exists a tuple $(a_1, a_2, \ldots, a_k)$ of Lyndon words satisfying $w = a_1a_2\cdots a_k$ [1012]. Fix such a tuple with **minimum** k. We claim that no $i \in \{1, 2, \ldots, k-1\}$ satisfies $a_i < a_{i+1}$.

Indeed, if some $i \in \{1, 2, \ldots, k-1\}$ would satisfy $a_i < a_{i+1}$, then the word a_ia_{i+1} would be Lyndon (by Proposition 6.1.16(a), applied to $u = a_i$ and $v = a_{i+1}$), whence $(a_1, a_2, \ldots, a_{i-1}, a_ia_{i+1}, a_{i+2}, a_{i+3}, \ldots, a_k)$ would also be a tuple of Lyndon words satisfying $w = a_1a_2\cdots a_{i-1}(a_ia_{i+1})a_{i+2}a_{i+3}\cdots a_k$ but having length $k-1 < k$, contradicting the fact that k is the minimum length of such a tuple. Hence, no $i \in \{1, 2, \ldots, k-1\}$ can satisfy $a_i < a_{i+1}$. Thus, $(a_1, a_2, \ldots, a_k)$ is a Hazewinkel-CFL factorization of w, so we have shown that such a Hazewinkel-CFL factorization exists.

It remains to show that there exists at most one Hazewinkel-CFL factorization of w. We shall prove this by induction over $\ell(w)$. Thus, we fix a word w and assume that

(13.146.10)

$$\text{for every word } v \text{ with } \ell(v) < \ell(w), \text{ there exists at most one Hazewinkel-CFL factorization of } v.$$

We now have to prove that there exists at most one Hazewinkel-CFL factorization of w.

Indeed, let $(a_1, a_2, \ldots, a_k)$ and $(b_1, b_2, \ldots, b_m)$ be two Hazewinkel-CFL factorizations of w. We need to prove that $(a_1, a_2, \ldots, a_k) = (b_1, b_2, \ldots, b_m)$. If w is empty, then this is obvious, so we WLOG assume that it is not; thus, $k > 0$ and $m > 0$.

The tuple $(a_1, a_2, \ldots, a_k)$ is a Hazewinkel-CFL factorization of w. Thus, $(a_1, a_2, \ldots, a_k)$ is a tuple of Lyndon words satisfying $w = a_1a_2\cdots a_k$ and such that no $i \in \{1, 2, \ldots, k-1\}$ satisfies $a_i < a_{i+1}$.

The tuple $(b_1, b_2, \ldots, b_m)$ is a Hazewinkel-CFL factorization of w. Thus, $(b_1, b_2, \ldots, b_m)$ is a tuple of Lyndon words satisfying $w = b_1b_2\cdots b_m$ and such that no $i \in \{1, 2, \ldots, m-1\}$ satisfies $b_i < b_{i+1}$. Now, b_m is a suffix of w (since $w = b_1b_2\cdots b_m$). Also, b_m is Lyndon. Thus, Lemma 13.146.6 (applied to $p = b_m$) yields $b_m \geq a_k$. The same argument (but with the roles of $(a_1, a_2, \ldots, a_k)$ and $(b_1, b_2, \ldots, b_m)$ switched) shows that $a_k \geq b_m$. Combined with $b_m \geq a_k$, this yields $a_k = b_m$. Now let $v = a_1a_2\cdots a_{k-1}$. Then, $(a_1, a_2, \ldots, a_{k-1})$ is a tuple of Lyndon words satisfying $v = a_1a_2\cdots a_{k-1}$ and such that no $i \in \{1, 2, \ldots, (k-1)-1\}$ satisfies

[1011] *Proof.* Assume the contrary. Thus, the word a_{j+1} is not a prefix of a_j. Hence, we have $a_j \leq a_{j+1}$ (because we know that either we have $a_j \leq a_{j+1}$ or the word a_{j+1} is a prefix of a_j). Since we cannot have $a_j < a_{j+1}$ (because no $i \in \{1, 2, \ldots, K\}$ satisfies $a_i < a_{i+1}$), this yields that we have $a_j = a_{j+1}$. Therefore, a_{j+1} is a prefix of a_j; this contradicts our assumption that the word a_{j+1} is not a prefix of a_j. This contradiction proves that our assumption was wrong, qed.

[1012] For instance, the tuple $(w_1, w_2, \ldots, w_{\ell(w)})$ of one-letter words is a valid example (recall that one-letter words are always Lyndon).

$a_i < a_{i+1}$ (because no $i \in \{1, 2, \ldots, k-1\}$ satisfies $a_i < a_{i+1}$). In other words, $(a_1, a_2, \ldots, a_{k-1})$ is a Hazewinkel-CFL factorization of v.

But $w = a_1 a_2 \cdots a_k = \underbrace{a_1 a_2 \cdots a_{k-1}}_{=v} \underbrace{a_k}_{=b_m} = v b_m$, so that

$$v b_m = w = b_1 b_2 \cdots b_m = b_1 b_2 \cdots b_{m-1} b_m.$$

Cancelling b_m from this equality yields $v = b_1 b_2 \cdots b_{m-1}$. Thus, $(b_1, b_2, \ldots, b_{m-1})$ is a tuple of Lyndon words satisfying $v = b_1 b_2 \cdots b_{m-1}$ and such that no $i \in \{1, 2, \ldots, (m-1)-1\}$ satisfies $b_i < b_{i+1}$ (since no $i \in \{1, 2, \ldots, m-1\}$ satisfies $b_i < b_{i+1}$). In other words, $(b_1, b_2, \ldots, b_{m-1})$ is a Hazewinkel-CFL factorization of v. Since $\ell(v) < \ell(w)$ (because $v = a_1 a_2 \cdots a_{k-1}$ is shorter than $w = a_1 a_2 \cdots a_k$), we can apply (13.146.10) to obtain that there exists at most one Hazewinkel-CFL factorization of v. But we already know two such Hazewinkel-CFL factorizations: $(a_1, a_2, \ldots, a_{k-1})$ and $(b_1, b_2, \ldots, b_{m-1})$. Thus, $(a_1, a_2, \ldots, a_{k-1}) = (b_1, b_2, \ldots, b_{m-1})$. Combining this with $a_k = b_m$, we obtain $(a_1, a_2, \ldots, a_k) = (b_1, b_2, \ldots, b_m)$. This is exactly what we needed to prove. So we have shown (by induction) that there exists at most one Hazewinkel-CFL factorization of w. This completes the proof of Proposition 13.146.7. □

(g) *Solution to Exercise 6.1.32 in the partial-order setting.* Let us first observe the following fact (in the partial-order setting): If h is a nonempty word which is not a Lyndon word, then

(13.146.11) there exist nonempty words u and v such that $h = uv$ and not $vu > uv$.

[1013] Renaming u and v as p and q in this result, we obtain the following: If h is a nonempty word which is not a Lyndon word, then

(13.146.12) there exist nonempty words p and q such that $h = pq$ and not $qp > pq$.

Furthermore, if h is a nonempty word which is not a Lyndon word, then

(13.146.13) there exist nonempty words u and v such that $h = uv$ and not $v > u$.

[1014]

Now, let us come to the solution of Exercise 6.1.32 in the partial-order setting.

Solution to Exercise 6.1.32(a) in the partial-order setting. The implication $\mathcal{A}' \Longrightarrow \mathcal{D}'$ can be proven in the same way as it was proven in the total-order setting.

Let us now prove the implication $\mathcal{D}' \Longrightarrow \mathcal{A}'$:

Proof of the implication $\mathcal{D}' \Longrightarrow \mathcal{A}'$: Assume that Assertion $\mathcal{D}'$ holds. Thus, if u and v are nonempty words satisfying $w = uv$, then we have $vu \geq uv$.

We need to prove that Assertion $\mathcal{A}'$ holds, i.e., that w is a power of a Lyndon word. Assume the contrary. Thus, w is not a power of a Lyndon word; hence, w is not a Lyndon word itself. Consequently, (13.146.11) (applied to $h = w$) shows that there exist nonempty words u and v such that $w = uv$ and not $vu > uv$.

[1013] *Proof of (13.146.11):* Let h be a nonempty word which is not a Lyndon word. We need to prove that (13.146.11) holds.

In fact, assume the contrary. Then, there exist no nonempty words u and v such that $h = uv$ and not $vu > uv$. In other words, any nonempty words u and v satisfying $h = uv$ must satisfy $vu > uv$. In other words, Assertion $\mathcal{D}$ of Theorem 6.1.20 (with h instead of w) holds.

But we know (from Exercise 6.1.33(c)) that Theorem 6.1.20 holds in the partial-order setting. Hence, we can apply Theorem 6.1.20 to h instead of w. We thus conclude that Assertions $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ and $\mathcal{D}$ of Theorem 6.1.20 (with h instead of w) are equivalent. Hence, Assertion $\mathcal{A}$ of Theorem 6.1.20 (with h instead of w) holds (since Assertion $\mathcal{D}$ of Theorem 6.1.20 (with h instead of w) holds). In other words, the word h is Lyndon. This contradicts the fact that the word h is not Lyndon. This contradiction shows that our assumption was wrong. Hence, (13.146.11) is proven, qed.

[1014] *Proof of (13.146.13):* Let h be a nonempty word which is not a Lyndon word. We need to prove that (13.146.13) holds.

In fact, assume the contrary. Then, there exist no nonempty words u and v such that $h = uv$ and not $v > u$. In other words, any nonempty words u and v satisfying $h = uv$ must satisfy $v > u$. In other words, Assertion $\mathcal{C}$ of Theorem 6.1.20 (with h instead of w) holds.

But we know (from Exercise 6.1.33(c)) that Theorem 6.1.20 holds in the partial-order setting. Hence, we can apply Theorem 6.1.20 to h instead of w. We thus conclude that Assertions $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ and $\mathcal{D}$ of Theorem 6.1.20 (with h instead of w) are equivalent. Hence, Assertion $\mathcal{A}$ of Theorem 6.1.20 (with h instead of w) holds (since Assertion $\mathcal{C}$ of Theorem 6.1.20 (with h instead of w) holds). In other words, the word h is Lyndon. This contradicts the fact that the word h is not Lyndon. This contradiction shows that our assumption was wrong. Hence, (13.146.13) is proven, qed.

Consider such a pair of nonempty words u and v with **minimum** $\ell(u)$. The minimality of $\ell(u)$ shows that

$$\left(\begin{array}{c} \text{if } u' \text{ and } v' \text{ are nonempty words such that } w = u'v' \text{ and not } v'u' > u'v', \\ \text{then } \ell(u') \geq \ell(u) \end{array} \right). \tag{13.146.14}$$

We have $vu \geq uv$ (according to Assertion $\mathcal{D}'$) but not $vu > uv$. Thus, $vu = uv$. Therefore, Proposition 6.1.4 yields that there exist a $t \in \mathfrak{A}^*$ and two nonnegative integers n and m such that $u = t^n$ and $v = t^m$. Consider this t and these n and m. We have $n \neq 0$ (since $t^n = u$ is nonempty) and $m \neq 0$ (since $t^m = v$ is nonempty), and the word t is nonempty (since $t^n = u$ is nonempty). Moreover, we have $n = 1$ [1015]. Hence, $t^n = t^1 = t$, so that $u = t^n = t$ and $w = \underbrace{u}_{=t}\underbrace{v}_{=t^m} = tt^m = t^{m+1}$. We shall now prove that the word t is Lyndon.

Assume the contrary. Then, t is not Lyndon. Hence, (13.146.12) (applied to $h = t$) shows that there exist nonempty words p and q such that $t = pq$ and not $qp > pq$. Consider these p and q. Since q is nonempty, we have $\ell(q) > 0$, so that $\ell\left(\underbrace{u}_{=t=pq}\right) = \ell(pq) = \ell(p) + \underbrace{\ell(q)}_{>0} > \ell(p)$.

We have $w = \underbrace{u}_{=t=pq} v = pqv$, and the words p and qv are nonempty[1016]. Now, using (13.146.14), it is easy to see that $qvp > pqv$ [1017].

Notice that $q(pq)^i = (qp)^i q$ for every $i \in \mathbb{N}$ [1018]. Applied to $i = m$, this yields $q(pq)^m = (qp)^m q$.

But $v = t^m$. Since $t = pq$, this rewrites as $v = (pq)^m$. Hence,

$$q \underbrace{v}_{=(pq)^m} p = \underbrace{q(pq)^m}_{=(qp)^m q} p = (qp)^m qp = (qp)^m (qp) = (qp)^{m+1},$$

so that $(qp)^{m+1} = qvp > pqv$. But

$$(pq)^{m+1} = (pq)\left(\underbrace{pq}_{=t}\right)^m = (pq)\underbrace{t^m}_{=v} = (pq)v = pqv < (qp)^{m+1}$$

[1015] *Proof.* Assume the contrary. Hence, $n \neq 1$. Combined with $n \neq 0$, this leads to $n \geq 2$. As a consequence, $u = t^n$ can be rewritten as $u = tt^{n-1}$. The word t^{n-1+m} is nonempty (since t is nonempty and since $\underbrace{n}_{\geq 2} - 1 + \underbrace{m}_{\geq 0} \geq 2 - 1 + 0 = 1$). Now, $w = \underbrace{u}_{=tt^{n-1}}\underbrace{v}_{=t^m} = tt^{n-1}t^m = tt^{n-1+m}$. Also, $t^{n-1+m}t = t^{n-1+m+1} = tt^{n-1+m}$. Thus, we do not have $t^{n-1+m}t > tt^{n-1+m}$. Hence, (13.146.14) (applied to $u' = t$ and $v' = t^{n-1+m}$) yields $\ell(t) \geq \ell(u)$ (since t and t^{n-1+m} are nonempty). Since $u = t^n$, this rewrites as $\ell(t) \geq \ell(t^n) = \underbrace{n}_{\geq 2}\ell(t) \geq 2\ell(t)$, whence $\ell(t) = 0$, which contradicts the fact that t is nonempty. This contradiction shows that our assumption was wrong, qed.

[1016] For qv, this follows from the nonemptiness of q.

[1017] *Proof.* Assume the contrary. Then, we do not have $qvp > pqv$. We can thus apply (13.146.14) to $u' = p$ and $v' = qv$. As a result, we obtain $\ell(p) \geq \ell(u)$. This contradicts $\ell(u) > \ell(p)$. This contradiction shows that our assumption was wrong, qed.

[1018] *Proof.* We shall prove the equality $q(pq)^i = (qp)^i q$ by induction over i:

Induction base: We have $q\underbrace{(pq)^0}_{=\varnothing} = q\varnothing = q = \underbrace{\varnothing}_{=(qp)^0} q = (qp)^0 q$. In other words, the equality $q(pq)^i = (qp)^i q$ holds for $i = 0$. This completes the induction base.

Induction step: Let $I \in \mathbb{N}$ be such that the equality $q(pq)^i = (qp)^i q$ holds for $i = I$. We need to show that the equality $q(pq)^i = (qp)^i q$ also holds for $i = I + 1$.

We have $q(pq)^I = (qp)^I q$ (since the equality $q(pq)^i = (qp)^i q$ holds for $i = I$). Thus,

$$q \underbrace{(pq)^{I+1}}_{=(pq)^I(pq)} = \underbrace{q(pq)^I}_{=(qp)^I q}(pq) = (qp)^I \underbrace{q(pq)}_{=(qp)q} = \underbrace{(qp)^I(qp)}_{=(qp)^{I+1}} q = (qp)^{I+1} q.$$

In other words, the equality $q(pq)^i = (qp)^i q$ holds for $i = I + 1$. This completes the induction step. The induction proof of the equality $q(pq)^i = (qp)^i q$ is thus complete.

(since $(qp)^{m+1} > pqv$). From this, it is easy to see that $pq < qp$ [1019], so that $qp > pq$. This contradicts the fact that we do not have $qp > pq$. This contradiction shows that our assumption (that t is not Lyndon) was wrong. Hence, t is Lyndon. Thus, w is a power of a Lyndon word (since $w = t^{m+1}$ is a power of t). Thus, Assertion $\mathcal{A}'$ is satisfied, so we have proven the implication $\mathcal{D}' \Longrightarrow \mathcal{A}'$.

Now we have proven both implications $\mathcal{A}' \Longrightarrow \mathcal{D}'$ and $\mathcal{D}' \Longrightarrow \mathcal{A}'$. Therefore, the equivalence $\mathcal{A}' \Longleftrightarrow \mathcal{D}'$ follows. Thus, Exercise 6.1.32(a) is solved in the partial-order case.

Solution to Exercise 6.1.32(b) in the partial-order setting. Consider the letter m and the alphabet $\mathfrak{A} \cup \{m\}$ defined in Assertion $\mathcal{F}''$. We notice that the lexicographic order on $\mathfrak{A}^*$ is the restriction of the lexicographic order on $(\mathfrak{A} \cup \{m\})^*$ to $\mathfrak{A}^*$. Therefore, when we have two words p and q in $\mathfrak{A}^*$, statements like "$p < q$" do not depend on whether we are regarding p and q as elements of $\mathfrak{A}^*$ or as elements of $(\mathfrak{A} \cup \{m\})^*$. It is easy to see that the one-letter word m satisfies

$$m > p \qquad \text{for every } p \in \mathfrak{A}^*. \tag{13.146.15}$$

[1020]

The implications $\mathcal{B}' \Longrightarrow \mathcal{E}'$, $\mathcal{C}' \Longrightarrow \mathcal{E}'$, $\mathcal{G}' \Longrightarrow \mathcal{H}'$, $\mathcal{F}'' \Longrightarrow \mathcal{B}'$, $\mathcal{F}'' \Longrightarrow \mathcal{C}'$ and $\mathcal{F}' \Longrightarrow \mathcal{B}'$ can be proven in the same way as they were proven in the total-order setting. We shall now prove some further implications.

Proof of the implication $\mathcal{E}' \Longrightarrow \mathcal{F}''$: Assume that Assertion $\mathcal{E}'$ holds.

Assume (for the sake of contradiction) that the word $wm \in (\mathfrak{A} \cup \{m\})^*$ is not a Lyndon word. Clearly, this word wm is nonempty. Thus, (13.146.13) (applied to $h = wm$) shows that there exist nonempty words u and v in $(\mathfrak{A} \cup \{m\})^*$ such that $wm = uv$ and not $v > u$. Denote these two nonempty words u and v by u and v'. Then, u and v' are nonempty words in $(\mathfrak{A} \cup \{m\})^*$ such that $wm = uv'$ and not $v' > u$.

The word v' is a proper suffix of wm (since $wm = uv'$ and since w is nonempty). Hence, v' is a nonempty suffix of wm. Thus, we must be in one of the following two cases (depending on whether this suffix begins before the suffix m of wm begins or afterwards):

Case 1: The word v' is a nonempty suffix of m. (Note that $v' = m$ is allowed.)

Case 2: The word v' has the form vm where v is a nonempty proper suffix of w.

Let us consider Case 1 first. In this case, the word v' is a nonempty suffix of m. Since the only nonempty suffix of m is m itself (because m is a one-letter word), this yields $v' = m$. Now, $wm = u \underbrace{v'}_{=m} = um$. Cancelling m from this equality, we obtain $w = u$. But $v' = m > w$ (by (13.146.15), applied to $p = w$), thus $v' > w = u$. This contradicts the fact that we don't have $v' > u$. Thus, we have obtained a contradiction in Case 1.

Let us now consider Case 2. In this case, the word v' has the form vm where v is a nonempty proper suffix of w. Consider this v. We have $wm = u \underbrace{v'}_{=vm} = uvm$. By cancelling m from this equality, we obtain $w = uv$. Thus, u and v are subwords of w, and therefore belong to $\mathfrak{A}^*$ (since $w \in \mathfrak{A}^*$). Moreover, u and v are nonempty. Hence, Assertion $\mathcal{E}'$ yields that either we have $v \geq u$ or the word v is a prefix of w. Since we cannot have $v \geq u$ (because if we had $v \geq u$, then we would have $v' = vm > v \geq u$, which would contradict the fact that we don't have $v' > u$), we therefore must have that v is a prefix of w. In other words, there exists a $q \in \mathfrak{A}^*$ such that $w = vq$. Consider this q. We have $m > q$ (by (13.146.15), applied to $p = q$). Thus, $q \leq m$. Hence, Proposition 6.1.2(b) (applied to $\mathfrak{A} \cup \{m\}$, v, q and m instead of $\mathfrak{A}$, a, c and d) yields $vq \leq vm$. Therefore, $vm \geq vq = w$ (since $w = vq$), so that $v' = vm \geq w = uv > u$ (since v is nonempty). This contradicts the fact that we don't have $v' > u$. Thus, we have found a contradiction in Case 2.

We have therefore obtained a contradiction in each of the two Cases 1 and 2. Since these two Cases cover all possibilities, this shows that we always get a contradiction. Hence, our assumption (that the word $wm \in (\mathfrak{A} \cup \{m\})^*$ is not a Lyndon word) was false. Hence, the word $wm \in (\mathfrak{A} \cup \{m\})^*$ is a Lyndon word. That is, Assertion $\mathcal{F}''$ holds. Hence, we have proven the implication $\mathcal{E}' \Longrightarrow \mathcal{F}''$.

Proof of the implication $\mathcal{B}' \Longrightarrow \mathcal{G}'$: Assume that Assertion $\mathcal{B}'$ holds.

[1019] *Proof.* Assume the contrary. Thus, we don't have $pq < qp$. But $(pq)(pq)^m = (pq)^{m+1} < (qp)^{m+1} = (qp)(qp)^{m+1}$ and $\ell(pq) = \ell(p) + \ell(q) = \ell(q) + \ell(p) = \ell(qp)$. Hence, Proposition 6.1.2(f) (applied to $a = pq$, $b = (pq)^m$, $c = qp$ and $d = (qp)^m$) yields $pq \leq qp$. Thus, $pq = qp$ (since we don't have $pq < qp$). Taking both sides of this equality to the $(m+1)$-th power, we obtain $(pq)^{m+1} = (qp)^{m+1}$, which contradicts $(pq)^{m+1} < (qp)^{m+1}$. This contradiction proves that our assumption was wrong, qed.

[1020] This can be proven in the same way as it was proven in the total-order setting.

Let s be the longest suffix v of w which does not satisfy $v \geq w$. (This is well-defined, because there exists a suffix v of w which does not satisfy $v \geq w$ – namely, the empty word.) So we know that s is a suffix v of w which does not satisfy $v \geq w$. In other words, s is a suffix of w and does not satisfy $s \geq w$. Hence, $s \neq w$ (because otherwise, we would have $s = w$ and thus $s \geq w$; but this would contradict the fact that s does not satisfy $s \geq w$). Thus, s is a proper suffix of w (because s is a suffix of w and satisfies $s \neq w$). Hence, there exists a nonempty word $h \in \mathfrak{A}^*$ satisfying $w = hs$. Consider this h. Using Assertion $\mathcal{B}'$, it is easy to see that s is a prefix of w [1021]. In other words, there exists a $g \in \mathfrak{A}^*$ such that $w = sg$. Consider this g. We have $g \neq \varnothing$ (because otherwise, we would have $g = \varnothing$ and thus $w = s \underbrace{g}_{=\varnothing} = s \neq w$, which would be absurd). In other words, the word g is nonempty. Thus, $s < sg = w$.

We know that s is the longest suffix v of w which does not satisfy $v \geq w$. Hence,

$$(\text{if } v \text{ is a suffix of } w \text{ which does not satisfy } v \geq w, \text{ then } \ell(v) \leq \ell(s)). \tag{13.146.16}$$

There exists a nonnegative integer m such that h^m is a prefix of s (for example, the nonnegative integer $m = 0$). Consider the **maximal** such integer m [1022]. Then, h^m is a prefix of s, but h^{m+1} is not a prefix of s. Since h^m is a prefix of s, there exists a word $q \in \mathfrak{A}^*$ such that $s = h^m q$. Consider this q. Clearly, $w = h \underbrace{s}_{=h^m q} = \underbrace{hh^m}_{=h^{m+1}} q = \underbrace{h^{m+1}}_{=h^m h} q = h^m hq$. Hence, $h^m hq = w = \underbrace{s}_{=h^m q} g = h^m qg$. Cancelling h^m from this equality, we obtain $hq = qg$. It is now easy to see that we don't have $h \leq q$ [1023]. But $h\varnothing = h \leq hq = qg$. Thus, Proposition 6.1.2(e) (applied to h, $\varnothing$, q and g instead of a, b, c and d) yields that either we have $h \leq q$ or the word q is a prefix of h. Thus, the word q is a prefix of h (since we don't have $h \leq q$).

Next, we shall prove that the word h is Lyndon.

In fact, assume the contrary. Then, h is not Lyndon. Hence, (13.146.11) shows that there exist nonempty words u and v such that $h = uv$ and not $vu > uv$. Consider these u and v. Since $w = \underbrace{h}_{=uv} s = uvs = u(vs)$, it is clear that the word vs is a suffix of w. If this suffix vs would not satisfy $vs \geq w$, then we could therefore obtain $\ell(vs) \leq \ell(s)$ (by (13.146.16), applied to vs instead of v), which would contradict $\ell(vs) = \underbrace{\ell(v)}_{\substack{>0 \\ (\text{since } v \text{ is nonempty})}} + \ell(s) > \ell(s)$. Thus, the suffix vs must satisfy $vs \geq w$. We thus have $vs \geq w = hs \geq h = uv$, so that $uv \leq vs$.

Recall that $s < w$. Hence, $vs \leq vw$ (by Proposition 6.1.2(b), applied to v, s and w instead of a, c and d). Now, $uv\varnothing = uv \leq vs \leq v \underbrace{w}_{=hs} = v \underbrace{h}_{=uv} s = vuvs$ and $\ell(uv) = \ell(u) + \ell(v) = \ell(v) + \ell(u) = \ell(vu) \leq \ell(vu)$. Hence, Proposition 6.1.2(f) (applied to uv, $\varnothing$, vu and vs instead of a, b, c and d) yields that $uv \leq vu$. In other words, $vu \geq uv$. Since we don't have $vu > uv$, we therefore must have $vu = uv$. Thus, the elements u and v of the monoid $\mathfrak{A}^*$ commute. Thus, the submonoid of $\mathfrak{A}^*$ generated by u and v is commutative. Since $h = uv$, the element h lies in this submonoid, and therefore the element h^m lies in it as well. Thus, h^m commutes

[1021] *Proof.* Assume the contrary. Thus, s is not a prefix of w. Hence, s is nonempty. Therefore, Assertion $\mathcal{B}'$ (applied to $u = h$ and $v = s$) yields that either we have $s \geq w$ or the word s is a prefix of w. Since s is not a prefix of w, we must thus have $s \geq w$. But this contradicts the fact that s does not satisfy $s \geq w$. This contradiction shows that our assumption was wrong, qed.

[1022] This is well-defined, because of the following reason:

We have $\ell(h) \geq 1$ (since the word h is nonempty). Thus, for every $m \in \mathbb{N}$ satisfying $m > \ell(s)$, we have $\ell(h^m) = m \underbrace{\ell(h)}_{\geq 1} \geq m > \ell(s)$. In other words, for every $m \in \mathbb{N}$ satisfying $m > \ell(s)$, the word h^m is longer than s. Hence, for every $m \in \mathbb{N}$ satisfying $m > \ell(s)$, the word h^m cannot be a prefix of s. Thus, for every sufficiently high $m \in \mathbb{N}$, the word h^m cannot be a prefix of s.

We thus know the following:

- There exists a nonnegative integer m such that h^m is a prefix of s.
- For every sufficiently high $m \in \mathbb{N}$, the word h^m cannot be a prefix of s.

Consequently, there exists a **maximal** nonnegative integer m such that h^m is a prefix of s, qed.

[1023] *Proof.* Assume the contrary. Then, $h \leq q$. Hence, $h \leq q \leq qg = hq$ (since $hq = qg$). Therefore, Proposition 6.1.2(g) (applied to h, q and q instead of a, b and c) yields that h is a prefix of q. In other words, there exists a word $r \in \mathfrak{A}^*$ such that $q = hr$. Consider this r. Now, $s = h^m \underbrace{q}_{=hr} = \underbrace{h^m h}_{=h^{m+1}} r = h^{m+1} r$, so that h^{m+1} is a prefix of s. This contradicts the fact that h^{m+1} is not a prefix of s. This contradiction proves that our assumption was wrong, qed.

with v (since this submonoid is commutative), i.e., we have $vh^m = h^m v$. Thus, $v \underbrace{s}_{=h^m q} = \underbrace{vh^m}_{=h^m v} q = h^m vq$. Thus, $h^m vq = vs \geq w = h \underbrace{s}_{=h^m q} = \underbrace{hh^m}_{=h^{m+1}=h^m h} q = h^m hq$, so that $h^m hq \leq h^m vq$. Hence, Proposition 6.1.2(c) (applied to h^m, hq and vq instead of a, c and d) yields $hq \leq vq$. But since q is a prefix of h, there exists a word $z \in \mathfrak{A}^*$ such that $h = qz$. Consider this z. We have

$$v \underbrace{qz}_{=h=uv} = \underbrace{vu}_{=uv=h} v = hv \leq h \underbrace{vu}_{=uv=h=qz} \qquad \text{(since } hv \text{ is a prefix of } hvu\text{)}$$
$$= hqz.$$

Also, $\ell\left(\underbrace{h}_{=uv} q\right) = \ell(uvq) = \ell(u(vq)) = \underbrace{\ell(u)}_{\substack{>0 \\ \text{(since } u \text{ is nonempty)}}} + \ell(vq) > \ell(vq)$, so that $\ell(vq) \leq \ell(hq)$. Hence, Proposition 6.1.2(f) (applied to vq, z, hq and z instead of a, b, c and d) yields $vq \leq hq$. Combined with $hq \leq vq$, this yields $vq = hq$. Hence, $\ell\left(\underbrace{vq}_{=hq}\right) = \ell(hq) > \ell(vq)$, which is absurd. This contradiction proves that our assumption is wrong. Thus, we have shown that the word h is Lyndon.

We now know that $h \in \mathfrak{A}^*$ is a Lyndon word, $m+1$ is a positive integer, and q is a prefix of h, and we have $w = h^{m+1}q$. Hence, there exists a Lyndon word $t \in \mathfrak{A}^*$, a positive integer ℓ and a prefix p of t (possibly empty) such that $w = t^\ell p$ (namely, $t = h$, $\ell = m+1$ and $p = q$). In other words, Assertion $\mathcal{G}'$ holds. This proves the implication $\mathcal{B}' \Longrightarrow \mathcal{G}'$.

Furthermore, the implication $\mathcal{F}' \Longrightarrow \mathcal{B}'$ holds. (In fact, it can be proven in the same way as it was proven in the total-order setting.)

Proof of the implication $\mathcal{H}' \Longrightarrow \mathcal{B}'$: Assume that Assertion $\mathcal{H}'$ holds. In other words, there exists a Lyndon word $t \in \mathfrak{A}^*$, a nonnegative integer ℓ and a prefix p of t (possibly empty) such that $w = t^\ell p$. Consider this t, this ℓ and this p.

We are going to prove that for every $m \in \mathbb{N}$,

$$\text{(13.146.17)} \qquad \left(\text{every suffix } s \text{ of } t^m p \text{ satisfies either } s \geq t^m p \text{ or } \left(\text{the word } s \text{ is a prefix of } t^m p\right)\right).$$

Proof of (13.146.17): We will prove (13.146.17) by induction over m:

Induction base: Using the implication $\mathcal{F}' \Longrightarrow \mathcal{B}'$, it is easy to see that (13.146.17) holds for $m = 0$ [1024]. This completes the induction base.

Induction step: Let M be a positive integer. Assume that (13.146.17) is proven for $m = M-1$. We will now show that (13.146.17) holds for $m = M$.

[1024] *Proof.* Assume that $m = 0$. Then, $t^m p = \underbrace{t^0}_{=\varnothing} p = \varnothing p = p$.

Let s be a suffix of $t^m p$. Then, s is a suffix of $t^m p = p$. In other words, there exists a word $g \in \mathfrak{A}^*$ satisfying $p = gs$. Consider this g.

We are going to prove that either $s \geq t^m p$ or (the word s is a prefix of $t^m p$). If the word s is empty, then this is obvious (because if the word s is empty, then the word s is a prefix of $t^m p$). Hence, we WLOG assume that the word s is nonempty. If the word g is empty, then it is also clear that either $s \geq t^m p$ or (the word s is a prefix of $t^m p$) (because if the word g is empty, then $g = \varnothing$ and thus $t^m p = p = \underbrace{g}_{=\varnothing} s = s$, so that the word s is a prefix of $t^m p$). Hence, we WLOG assume that the word g is nonempty.

But the word p is a prefix of a Lyndon word in $\mathfrak{A}^*$ (since p is a prefix of t, and since t is a Lyndon word in $\mathfrak{A}^*$). In other words, Assertion $\mathcal{F}'$ with w replaced by p is satisfied. Hence, Assertion $\mathcal{B}'$ with w replaced by p is satisfied as well (since we have already proven the implication $\mathcal{F}' \Longrightarrow \mathcal{B}'$). In other words,

$$\text{(13.146.18)} \qquad \left(\begin{array}{c}\text{if } u \text{ and } v \text{ are nonempty words satisfying } p = uv \text{, then} \\ \text{either we have } v \geq p \text{ or the word } v \text{ is a prefix of } p\end{array}\right).$$

Since the words g and s are nonempty, we can apply (13.146.18) to $u = g$ and $v = s$. As a result, we obtain that either we have $s \geq p$ or the word s is a prefix of p. In other words, either $s \geq p$ or (the word s is a prefix of p). In other words, either $s \geq t^m p$ or (the word s is a prefix of $t^m p$) (since $t^m p = p$). This proves (13.146.17).

Let r denote the word $t^{M-1}p$. It is easy to see that r is a prefix of $t^M p$ [1025]. In other words, there exists a word $g \in \mathfrak{A}^*$ such that $t^M p = rg$. Consider this g.

Let s be a suffix of $t^M p$. We shall show that either $s \geq t^M p$ or $\left(\text{the word } s \text{ is a prefix of } t^M p\right)$.

In order to prove this, let us assume the contrary (for the sake of contradiction). Then, neither $s \geq t^M p$ nor $\left(\text{the word } s \text{ is a prefix of } t^M p\right)$. In other words, we don't have $s \geq t^M p$, and the word s is not a prefix of $t^M p$. If the word s was a prefix of r, then the word s would be a prefix of $t^M p$ (since r is a prefix of $t^M p$), which would contradict the fact that the word s is not a prefix of $t^M p$. Hence, the word s cannot be a prefix of r. In other words, the word s cannot be a prefix of $t^{M-1}p$ (since $r = t^{M-1}p$).

We don't have $s \geq t^M p$. Since $s = s\varnothing$ and $t^M p = rg$, this rewrites as follows: We don't have $s\varnothing \geq rg$. In other words, we don't have $rg \leq s\varnothing$.

The word s is a suffix of $\underbrace{t^M}_{=tt^{M-1}} p = t \underbrace{t^{M-1}p}_{=r} = tr$. Therefore, we must be in one of the following two cases (depending on whether this suffix begins before the suffix r of tr begins or afterwards):

Case 1: The word s is a suffix of r. (Note that $s = r$ is allowed.)

Case 2: The word s has the form $s'r$ where s' is a nonempty suffix of t.

Let us consider Case 1 first. In this case, the word s is a suffix of r. In other words, the word s is a suffix of $t^{M-1}p$ (since $r = t^{M-1}p$). Hence, (13.146.17) (applied to $m = M-1$) yields that either $s \geq t^{M-1}p$ or $\left(\text{the word } s \text{ is a prefix of } t^{M-1}p\right)$ (since (13.146.17) holds for $m = M-1$). Since the word s cannot be a prefix of $t^{M-1}p$, we thus must have $s \geq t^{M-1}p$. Thus, $t^{M-1}p \leq s$, so that $r = t^{M-1}p \leq s$. Therefore, Proposition 6.1.2(d) (applied to r, g, s and $\varnothing$ instead of a, b, c and d) yields that either we have $rg \leq s\varnothing$ or the word r is a prefix of s. Thus, the word r is a prefix of s (since we don't have $rg \leq s\varnothing$). But $\ell(r) \geq \ell(s)$ (since s is a suffix of r). Hence, r is a prefix of s which is at least as long as s itself. Consequently, $r = s$. Hence, $s = r$, so that s is a prefix of $s = r$. This contradicts the fact that the word s cannot be a prefix of r. Thus, we have found a contradiction in Case 1.

Let us now consider Case 2. In this case, the word s has the form $s'r$ where s' is a nonempty suffix of t. Consider this s'. Corollary 6.1.15 (applied to t and s' instead of w and v) yields $s' \geq t$. That is, $t \leq s'$. But s' is a suffix of t, so that $\ell(s') \leq \ell(t)$. Hence, $\ell(t) \geq \ell(s')$. Thus, Proposition 6.1.2(j) (applied to t, s' and r instead of a, b and c) yields $tr \leq s'r$. Hence, $s'r \geq tr$, so that $s = s'r \geq tr = t^M p$. This contradicts the fact that we don't have $s \geq t^M p$. Thus, we have found a contradiction in Case 2.

We have thus obtained a contradiction in each of the two Cases 1 and 2. Since these two Cases cover all possibilities, this shows that we always get a contradiction. This completes the proof that either $s \geq t^M p$ or $\left(\text{the word } s \text{ is a prefix of } t^M p\right)$.

Now, forget that we fixed s. We thus have shown that every suffix s of $t^M p$ satisfies either $s \geq t^M p$ or $\left(\text{the word } s \text{ is a prefix of } t^M p\right)$. In other words, (13.146.17) holds for $m = M$. This completes the induction step, and thus (13.146.17) is proven by induction.

Now, let u and v be nonempty words satisfying $w = uv$. Then, v is a suffix of $w = t^\ell p$. Hence, (13.146.17) (applied to $m = \ell$ and $s = v$) yields that either $v \geq t^\ell p$ or $\left(\text{the word } v \text{ is a prefix of } t^\ell p\right)$. In other words, either $v \geq w$ or (the word v is a prefix of w) (since $w = t^\ell p$). In other words, either we have $v \geq w$ or the word v is a prefix of w.

Now, forget that we fixed u and v. We thus have shown that if u and v are nonempty words satisfying $w = uv$, then either we have $v \geq w$ or the word v is a prefix of w. In other words, Assertion $\mathcal{B}'$ holds. Thus, the implication $\mathcal{H}' \Longrightarrow \mathcal{B}'$ is proven.

We have thus proven the implications $\mathcal{B}' \Longrightarrow \mathcal{E}'$, $\mathcal{C}' \Longrightarrow \mathcal{E}'$, $\mathcal{G}' \Longrightarrow \mathcal{H}'$, $\mathcal{E}' \Longrightarrow \mathcal{F}''$, $\mathcal{F}'' \Longrightarrow \mathcal{B}'$, $\mathcal{F}' \Longrightarrow \mathcal{C}'$, $\mathcal{B}' \Longrightarrow \mathcal{G}'$ and $\mathcal{H}' \Longrightarrow \mathcal{B}'$. Combined, these yield the equivalence $\mathcal{B}' \Longleftrightarrow \mathcal{C}' \Longleftrightarrow \mathcal{E}' \Longleftrightarrow \mathcal{F}'' \Longleftrightarrow \mathcal{G}' \Longleftrightarrow \mathcal{H}'$. This solves Exercise 6.1.32(b) in the partial-order setting.

Solution to Exercise 6.1.32(c) in the partial-order setting. The implication $\mathcal{F}' \Longrightarrow \mathcal{B}'$ has already been proven in our solution of Exercise 6.1.32(b). Hence, Exercise 6.1.32(c) is solved in the partial-order setting.

Solution to Exercise 6.1.32(d) in the partial-order setting. The solution of Exercise 6.1.32(d) in the partial-order setting proceeds precisely as in the total-order setting.

[1025] *Proof.* There exists a word $q \in \mathfrak{A}^*$ such that $t = pq$ (since p is a prefix of t). Consider this q. We have $\underbrace{t^M}_{=t^{M-1}t} p = t^{M-1} \underbrace{t}_{=pq} p = \underbrace{t^{M-1}p}_{=r} qp = rqp = r(qp)$. Hence, r is a prefix of $t^M p$, qed.

Solution to Exercise 6.1.32(e) in the partial-order setting. The solution of Exercise 6.1.32(e) in the partial-order setting proceeds precisely as in the total-order setting.

Solution to Exercise 6.1.32(f) in the partial-order setting. Assume that there exists a letter $\mu \in \mathfrak{A}$ such that ($\mu > a$ for some letter a of w). Consider this μ. We need to prove that the equivalence $\mathcal{F}' \Longleftrightarrow \mathcal{F}''$ holds.

Combining the implication $\mathcal{F}' \Longrightarrow \mathcal{B}'$ (which has already been proven) and the implication $\mathcal{B}' \Longrightarrow \mathcal{F}''$ (which follows from the equivalence $\mathcal{B}' \Longleftrightarrow \mathcal{C}' \Longleftrightarrow \mathcal{E}' \Longleftrightarrow \mathcal{F}'' \Longleftrightarrow \mathcal{G}' \Longleftrightarrow \mathcal{H}'$ proven above), we obtain the implication $\mathcal{F}' \Longrightarrow \mathcal{F}''$. Thus, in order to prove the equivalence $\mathcal{F}' \Longleftrightarrow \mathcal{F}''$, it is enough to verify the implication $\mathcal{F}'' \Longrightarrow \mathcal{F}'$. Let us do this now.

Assume that Assertion $\mathcal{F}''$ holds. Due to the implication $\mathcal{F}'' \Longrightarrow \mathcal{G}'$ (which follows from the equivalence $\mathcal{B}' \Longleftrightarrow \mathcal{C}' \Longleftrightarrow \mathcal{E}' \Longleftrightarrow \mathcal{F}'' \Longleftrightarrow \mathcal{G}' \Longleftrightarrow \mathcal{H}'$ proven above), this yields that Assertion $\mathcal{G}'$ holds. In other words, there exists a Lyndon word $t \in \mathfrak{A}^*$, a positive integer ℓ and a prefix p of t (possibly empty) such that $w = t^\ell p$. Consider these t, ℓ and p.

We shall first prove that

$$\text{there exists a letter } z \in \mathfrak{A} \text{ such that } z > t. \tag{13.146.19}$$

Proof of (13.146.19): The word t is Lyndon and thus nonempty. Hence, $\ell(t) \geq 1$. We thus are in one of the following two cases:

Case 1: We have $\ell(t) = 1$.

Case 2: We have $\ell(t) > 1$.

Let us consider Case 1 first. In this case, we have $\ell(t) = 1$. Thus, t is a one-letter word. In other words, $t = b$ for some letter b. Let us consider this b.

Recall that $\mu > a$ for some letter a of w. Consider this letter a.

The word $t^\ell p$ is a prefix of $t^\ell t$ (since p is a prefix of t). In other words, the word w is a prefix of $t^{\ell+1}$ (since $w = t^\ell p$ and $t^{\ell+1} = t^\ell t$). Hence, each letter of w is a letter of $t^{\ell+1}$. Since each letter of $t^{\ell+1}$ is a letter of t, this shows that each letter of w is a letter of t. Applying this to the letter a, we conclude that a is a letter of t (since a is a letter of w).

Now, both a and b are letters of t. Since the word t has only one letter (because $\ell(t) = 1$), this yields that $a = b$. Hence, $\mu > a = b = t$ (since $t = b$). Hence, there exists a letter $z \in \mathfrak{A}$ such that $z > t$ (namely, $z = \mu$). Thus, (13.146.19) is proven in Case 1.

Let us now consider Case 2. In this case, we have $\ell(t) > 1$. Let g be the last letter of the word t (this is well-defined since t is nonempty). Then, the one-letter word g is a suffix of the word t, and therefore there exists a word $t' \in \mathfrak{A}^*$ such that $t = t'g$. Consider this t'. Since g is a one-letter word, we have $\ell(g) = 1$. Thus, $\ell\left(\underbrace{t}_{=t'g}\right) = \ell(t') + \underbrace{\ell(g)}_{=1} = \ell(t') + 1$, so that $\ell(t') = \underbrace{\ell(t)}_{>1} - 1 > 1 - 1 = 0$. Hence, the word t' is nonempty. Thus, g is a proper suffix of t (since $t = t'g$). Also, g is nonempty (since $\ell(g) = 1$). But recall that the word t is Lyndon. By the definition of a Lyndon word, this yields that every nonempty proper suffix v of t satisfies $v > t$. Applying this to $v = g$, we obtain $g > t$. Thus, there exists a letter $z \in \mathfrak{A}$ such that $z > t$ (namely, $z = g$). Thus, (13.146.19) is proven in Case 2.

We have thus proven (13.146.19) in each of the two Cases 1 and 2. Since these two Cases cover all possibilities, this yields that (13.146.19) always holds.

Now, (13.146.19) shows that there exists a letter $z \in \mathfrak{A}$ such that $z > t$. Consider this letter z. We have $z > t$, thus $t < z$, hence $t \leq z$. The one-letter word z is Lyndon (since every one-letter word is Lyndon). Thus, $\underbrace{t, t, \ldots, t}_{\ell+1 \text{ times } t}, z$ are $\ell + 2$ Lyndon words (since both t and z are Lyndon) satisfying $t \leq t \leq \cdots \leq t \leq z$ (since $t \leq z$) and $t < z$. Hence, Exercise 6.1.23 (applied to $\ell + 2$ and $\left(\underbrace{t, t, \ldots, t}_{\ell+1 \text{ times } t}, z\right)$ instead of n and $(w_1, w_2, \ldots, w_n)$) yields that $\underbrace{tt\cdots t}_{\ell+1 \text{ times } t} z$ is a Lyndon word. In other words, $t^{\ell+1}z$ is a Lyndon word (since $\underbrace{tt\cdots t}_{\ell+1 \text{ times } t} = t^{\ell+1}$).

But p is a prefix of t. In other words, there exists a $q \in \mathfrak{A}^*$ such that $t = pq$. Consider this q. We then have $\underbrace{t^{\ell+1}}_{=t^\ell t} z = t^\ell \underbrace{t}_{=pq} z = \underbrace{t^\ell p}_{=w} qz = wqz = w(qz)$. Hence, w is a prefix of the word $t^{\ell+1}z$. Thus, w is a prefix of a Lyndon word in $\mathfrak{A}^*$ (because $t^{\ell+1}z$ is a Lyndon word in $\mathfrak{A}^*$). In other words, Assertion $\mathcal{F}'$ holds. This proves the implication $\mathcal{F}'' \Longrightarrow \mathcal{F}'$. This solves Exercise 6.1.32(f) in the partial-order case.

Remark: Of course, for a letter $\mu \in \mathfrak{A}$, if we have ($\mu > a$ for every letter a of w), then we also have ($\mu > a$ for some letter a of w) (since w is nonempty and thus has at least one letter). Hence, Exercise 6.1.32(e) is a particular case of Exercise 6.1.32(f).

Altogether, we have now solved all parts of Exercise 6.1.32 in the partial-order case. Exercise 6.1.33(g) is thus solved.

[*Remark:* The validity of some parts of Exercise 6.1.32 in the partial-order case can also be deduced from their validity in the total-order case using Proposition 13.146.3. For example, the equivalence $\mathcal{B}' \Longleftrightarrow \mathcal{C}' \Longleftrightarrow \mathcal{E}'$ can be treated this way. We shall not give any details of this alternative approach, however.]

13.147. **Solution to Exercise 6.1.34.** *Solution to Exercise 6.1.34.* In the following, whenever G is a group and P is a (left) G-set, we denote by G_u the stabilizer of u in G, that is, the subgroup $\{g \in G \mid gu = u\}$ of G.

Let $C^{(n)}$ denote the subgroup $\langle c^n \rangle$ of the infinite cyclic group C. Then, $C/C^{(n)}$ is a cyclic group with n elements. Hence, $\left|C/C^{(n)}\right| = n$.

Also, recall that c acts on $\mathfrak{A}^n$ by cyclically rotating n-tuples one step to the left ($c \cdot (a_1, a_2, \ldots, a_n) = (a_2, a_3, \ldots, a_n, a_1)$). Thus, c^n acts trivially on $\mathfrak{A}^n$ (since cyclically rotating an n-tuple n steps to the left does nothing). Therefore, the whole subgroup $C^{(n)}$ acts trivially on $\mathfrak{A}^n$ (since $C^{(n)} = \langle c^n \rangle$).

The word "divisor" shall mean "positive divisor" throughout this solution.

(a) Let N be any n-necklace. Then, N is an orbit of the C-action, thus a nonempty set. Fix an element w of N (such a w exists since N is nonempty).

Recall that the subgroup $C^{(n)}$ acts trivially on $\mathfrak{A}^n$. In particular, $C^{(n)}$ stabilizes w. Hence, $C^{(n)} \subset C_w$. Thus, $\left[C : C^{(n)}\right] = [C : C_w] \cdot \left[C_w : C^{(n)}\right]$. As a consequence, $[C : C_w] \mid [C : C_w] \cdot \left[C_w : C^{(n)}\right] = \left[C : C^{(n)}\right] = \left|C/C^{(n)}\right| = n$, so that $[C : C_w] < \infty$.

But N is an orbit of the C-action containing w. Hence, N is the C-orbit of w. Thus, by the orbit-stabilizer theorem, we have $|N| = [C : C_w]$. This yields $|N| = [C : C_w] < \infty$, so that N is a finite set.

Also, $|N| = [C : C_w] \mid n$. This solves Exercise 6.1.34(a).

(b) Let $w = (w_1, w_2, \ldots, w_n) \in \mathfrak{A}^n$ be an n-tuple. We need to prove the equivalence between the following two assertions:

> *Assertion $\mathcal{A}$:* The n-necklace $[w]$ is aperiodic.
>
> *Assertion $\mathcal{B}$:* Every $k \in \{1, 2, \ldots, n-1\}$ satisfies $(w_{k+1}, w_{k+2}, \ldots, w_n, w_1, w_2, \ldots, w_k) \neq w$.

We will achieve this by proving the implications $\mathcal{A} \Longrightarrow \mathcal{B}$ and $\mathcal{B} \Longrightarrow \mathcal{A}$. But let us first do some preparatory work.

Let N denote the n-necklace $[w]$. Thus, $w \in N$ and therefore (as we have shown in the solution of Exercise 6.1.34(a)) we have $C^{(n)} \subset C_w$ and $|N| = [C : C_w]$.

We are now ready to prove implications $\mathcal{A} \Longrightarrow \mathcal{B}$ and $\mathcal{B} \Longrightarrow \mathcal{A}$.

Proof of the implication $\mathcal{A} \Longrightarrow \mathcal{B}$: Assume that Assertion $\mathcal{A}$ holds. In other words, the n-necklace $[w]$ is aperiodic.

In other words, N is aperiodic (since $N = [w]$). In other words, the period of N is n (by the definition of "aperiodic"). In other words, $|N|$ is n (since the period of N is defined as $|N|$). In other words, $|N| = n$. Thus, $n = |N| = [C : C_w]$. But

$$n = \left[C : C^{(n)}\right] = \underbrace{[C : C_w]}_{=n} \cdot \left[C_w : C^{(n)}\right] \qquad \left(\text{since } C^{(n)} \subset C_w\right)$$
$$= n \cdot \left[C_w : C^{(n)}\right].$$

Solving this for $\left[C_w : C^{(n)}\right]$, we obtain $\left[C_w : C^{(n)}\right] = 1$, so that $C^{(n)} = C_w$.

Let now $k \in \{1,2,\ldots,n-1\}$. We are going to prove that $(w_{k+1}, w_{k+2}, \ldots, w_n, w_1, w_2, \ldots, w_k) \neq w$. Indeed, assume the contrary. Thus, $(w_{k+1}, w_{k+2}, \ldots, w_n, w_1, w_2, \ldots, w_k) = w$.

Recall that c acts on $\mathfrak{A}^n$ by cyclically rotating n-tuples one step to the left. Hence, c^k acts on $\mathfrak{A}^n$ by cyclically rotating n-tuples k steps to the left. Hence,

$$c^k w = (w_{k+1}, w_{k+2}, \ldots, w_n, w_1, w_2, \ldots, w_k) = w.$$

Thus, $c^k \in C_w = C^{(n)} = \langle c^n \rangle$. But if an integer $a \in \mathbb{Z}$ satisfies $c^a \in \langle c^n \rangle$, then we must have $n \mid a$ (this follows from the structure of the infinite cyclic group C). Applied to $a = k$, this yields $n \mid k$. But this is absurd, since $k \in \{1,2,\ldots,n-1\}$. This contradiction proves that our assumption was wrong. Hence, we have proven $(w_{k+1}, w_{k+2}, \ldots, w_n, w_1, w_2, \ldots, w_k) \neq w$.

Now, let us forget that we fixed k. We thus have shown that every $k \in \{1,2,\ldots,n-1\}$ satisfies $(w_{k+1}, w_{k+2}, \ldots, w_n, w_1, w_2, \ldots, w_k) \neq w$. In other words, Assertion $\mathcal{B}$ holds. This proves the implication $\mathcal{A} \Longrightarrow \mathcal{B}$.

Proof of the implication $\mathcal{B} \Longrightarrow \mathcal{A}$: Assume that Assertion $\mathcal{B}$ holds. In other words, every $k \in \{1,2,\ldots,n-1\}$ satisfies $(w_{k+1}, w_{k+2}, \ldots, w_n, w_1, w_2, \ldots, w_k) \neq w$.

We will prove that $C^{(n)} = C_w$. Indeed, assume the contrary. Thus, $C^{(n)} \neq C_w$. Since $C^{(n)} \subset C_w$, this yields that $C^{(n)}$ is a proper subset of C_w. Hence, there exists some $d \in C_w$ such that $d \notin C^{(n)}$. Consider this d. Write d in the form $d = c^h$ for some $h \in \mathbb{Z}$ (this is possible since c generates C). Let k denote the remainder of h modulo n. Then, $h - k$ is divisible by n, and thus $c^{h-k} \in \langle c^n \rangle = C^{(n)} \subset C_w$, so that $c^{h-k} w = w$. But also, $c^h w = w$ (since $c^h = d \in C_w$). Hence, $w = \underbrace{c^h}_{=c^k c^{h-k}} w = c^k \underbrace{c^{h-k} w}_{=w} = c^k w$.

But $k \in \{0,1,\ldots,n-1\}$ (since k is a remainder modulo n). We have $c^{h-k} \neq c^h$ (since $c^{h-k} \in C^{(n)}$ whereas $c^h = d \notin C^{(n)}$). Hence, $h-k \neq h$, so that $k \neq 0$. Combined with $k \in \{0,1,\ldots,n-1\}$, this yields $k \in \{0,1,\ldots,n-1\} \backslash \{0\} = \{1,2,\ldots,n-1\}$. Thus, Assertion $\mathcal{B}$ yields that $(w_{k+1}, w_{k+2}, \ldots, w_n, w_1, w_2, \ldots, w_k) \neq w$.

Recall that c acts on $\mathfrak{A}^n$ by cyclically rotating n-tuples one step to the left. Hence, c^k acts on $\mathfrak{A}^n$ by cyclically rotating n-tuples k steps to the left. Hence,

$$c^k w = (w_{k+1}, w_{k+2}, \ldots, w_n, w_1, w_2, \ldots, w_k) \neq w = c^k w.$$

This is absurd. This contradiction shows that our assumption was wrong. Hence, we have shown that $C^{(n)} = C_w$. Now, $|N| = \left[C : \underbrace{C_w}_{=C^{(n)}}\right] = \left[C : C^{(n)}\right] = n$. In other words, the period of N is n (since the period of N is defined as $|N|$). In other words, N is aperiodic (by the definition of "aperiodic"). In other words, $[w]$ is aperiodic (since $N = [w]$). In other words, Assertion $\mathcal{A}$ holds. This proves the implication $\mathcal{B} \Longrightarrow \mathcal{A}$.

Now we have proven both implications $\mathcal{A} \Longrightarrow \mathcal{B}$ and $\mathcal{B} \Longrightarrow \mathcal{A}$. Hence, the Assertions $\mathcal{A}$ and $\mathcal{B}$ are equivalent. Exercise 6.1.34(b) is solved.

Before we come to the solution of Exercise 6.1.34(c), let us state a simple lemma:

Lemma 13.147.1. *Let n be a positive integer. Assume that the set $\mathfrak{A}$ is totally ordered. Let $w \in \mathfrak{A}^n$ be a Lyndon word. Let $k \in \{1,2,\ldots,n-1\}$. Then, $c^k w > w$ in the lexicographic order.*

Proof of Lemma 13.147.1. We have $w = (w_1, w_2, \ldots, w_n)$ (since $w \in \mathfrak{A}^n$). Let $u = (w_1, w_2, \ldots, w_k)$ and $v = (w_{k+1}, w_{k+2}, \ldots, w_n)$. These words u and v are well-defined and nonempty (since $k \in \{1,2,\ldots,n-1\}$) and satisfy

$$\begin{aligned} uv &= (w_1, w_2, \ldots, w_k)(w_{k+1}, w_{k+2}, \ldots, w_n) = (w_1, w_2, \ldots, w_k, w_{k+1}, w_{k+2}, \ldots, w_n) \\ &= (w_1, w_2, \ldots, w_n) = w. \end{aligned}$$

But the n-tuple $c^k w$ is obtained from w by k-fold cyclic rotation to the left (since c acts on $\mathfrak{A}^n$ by cyclically rotating n-tuples one step to the left). In other words,

$$\begin{aligned} c^k w &= (w_{k+1}, w_{k+2}, \ldots, w_n, w_1, w_2, \ldots, w_k) && \left(\text{since } w = (w_1, w_2, \ldots, w_n) \text{ and } k \in \{1, 2, \ldots, n-1\}\right) \\ &= \underbrace{(w_{k+1}, w_{k+2}, \ldots, w_n)}_{=v} \underbrace{(w_1, w_2, \ldots, w_k)}_{=u} = vu > uv && (\text{by Proposition 6.1.14(c)}) \\ &= w. \end{aligned}$$

This proves Lemma 13.147.1. □

(c) Let N be any aperiodic n-necklace. We need to show that N contains exactly one Lyndon word.

Since the necklace N is aperiodic, we know that the period of N is n. In other words, $|N| = n$.

Let w be the lexicographically smallest word contained in N. The word w has n letters (since $w \in N \subset \mathfrak{A}^n$), and thus is nonempty. We will prove that the word w is Lyndon.

Clearly, N is a C-orbit (since N is an n-necklace), and thus N is the orbit of the word w (since $w \in N$). In other words, $N = Cw$.

We can see (as in the solution of Exercise 6.1.34(a)) that $C^{(n)} \subset C_w$ and $|N| = [C : C_w]$. Now, $n = \left[C : C^{(n)}\right] = \underbrace{[C : C_w]}_{=|N|=n} \cdot \left[C_w : C^{(n)}\right] = n \cdot \left[C_w : C^{(n)}\right]$. Thus, $\left[C_w : C^{(n)}\right] = 1$, so that $C_w = C^{(n)}$.

Now, let u and v be two nonempty words satisfying $w = uv$. We will prove that $vu > uv$.

Assume the contrary. Thus, $vu \leq uv$.

We have $w = (w_1, w_2, \ldots, w_n)$ (since $w \in N \subset \mathfrak{A}^n$). Thus, there exists some $k \in \{0, 1, \ldots, n\}$ such that $u = (w_1, w_2, \ldots, w_k)$ and $v = (w_{k+1}, w_{k+2}, \ldots, w_n)$ (since $w = uv$). Consider this k. We have $0 < k < n$ (since u and v are nonempty). Since $v = (w_{k+1}, w_{k+2}, \ldots, w_n)$ and $u = (w_1, w_2, \ldots, w_k)$, we have $vu = (w_{k+1}, w_{k+2}, \ldots, w_n, w_1, w_2, \ldots, w_k)$. In other words, the n-tuple vu is obtained from w by k-fold cyclic rotation to the left. In yet other words, $vu = c^k w$ (since c acts on $\mathfrak{A}^n$ by cyclically rotating n-tuples one step to the left). Hence, $vu = \underbrace{c^k}_{\in C} w \in Cw = N$. Thus, $vu \geq w$ (since w is the lexicographically smallest word contained in N). Combined with $vu \leq uv = w$, this yields $vu = w$. Hence, $c^k w = vu = w$, so that c^k stabilizes w. In other words, $c^k \in C_w = C^{(n)}$. But this is impossible, since $0 < k < n$ (and since $C^{(n)}$ is the subgroup $\langle c^n \rangle$ of C). This contradiction proves that our assumption was wrong. Thus, we have proven that $vu > uv$.

Let us now forget that we fixed u and v. We thus have proven that any nonempty words u and v satisfying $w = uv$ satisfy $vu > uv$. In other words, the word w satisfies Assertion $\mathcal{D}$ of Theorem 6.1.20. Consequently, the word w satisfies Assertion $\mathcal{A}$ of Theorem 6.1.20 as well (since Theorem 6.1.20 yields that these two assertions are equivalent); in other words, w is Lyndon. The orbit N thus contains at least one Lyndon word (namely, w).

We shall next prove that w is the only Lyndon word in N. Indeed, let p be any Lyndon word in N distinct from w. We will derive a contradiction.

It is easy to see that there exists a $k \in \{1, 2, \ldots, n-1\}$ satisfying $p = c^k w$ [1026]. Consider this k. Then, $w = c^{n-k} p$ [1027]. Notice that $n - k \in \{1, 2, \ldots, n-1\}$ (since $k \in \{1, 2, \ldots, n-1\}$) and $p \in N \subset \mathfrak{A}^n$. Hence, Lemma 13.147.1 (applied to p and $n-k$ instead of w and k) yields that $c^{n-k} p > p$. Hence, $p < c^{n-k} p = w$.

[1026] *Proof.* We have $p \in N = Cw$. Hence, there exists some $e \in C$ satisfying $p = ew$. Consider this e.

We know that $C^{(n)}$ acts trivially on $\mathfrak{A}^n$, and thus stabilizes w. Hence, $C^{(n)} w = \{w\}$.

By the well-known properties of the infinite cyclic group, we know that the elements c^0, c^1, ..., c^{n-1} of C form a system of unique representatives for the cosets of $C^{(n)}$ in C. Hence, every $d \in C$ belongs to the coset $c^\ell C^{(n)}$ for some $\ell \in \{0, 1, \ldots, n-1\}$. Applied to $d = e$, this yields that e belongs to the coset $c^\ell C^{(n)}$ for some $\ell \in \{0, 1, \ldots, n-1\}$. Consider this ℓ. We have $e \in c^\ell C^{(n)}$, thus $p = \underbrace{e}_{\in c^\ell C^{(n)}} w \in c^\ell \underbrace{C^{(n)} w}_{=\{w\}} = c^\ell \{w\} = \{c^\ell w\}$, so that $p = c^\ell w$.

If $\ell = 0$, then $p = \underbrace{c^\ell}_{=c^0 = 1_C} w = 1_C w = w$, which contradicts the fact that p is distinct from w. Hence, we cannot have $\ell = 0$. Thus, $\ell \in \{0, 1, \ldots, n-1\} \setminus \{0\} = \{1, 2, \ldots, n-1\}$. Hence, there exists a $k \in \{1, 2, \ldots, n-1\}$ satisfying $p = c^k w$ (namely, $k = \ell$), qed.

[1027] *Proof.* We know that c^n acts trivially on $\mathfrak{A}^n$. Thus, $c^n w = w$. Hence, $w = \underbrace{c^n}_{=c^{n-k} c^k} w = c^{n-k} \underbrace{c^k w}_{=p} = c^{n-k} p$, qed.

Since $p \in N$, this shows that there exists an element of N which is lexicographically smaller than w (namely, p). This contradicts the fact that w is the lexicographically smallest word contained in N.

Now, let us forget that we fixed w. We thus have obtained a contradiction for every Lyndon word p in N distinct from w. Thus, there exists no Lyndon word p in N distinct from w. Hence, w is the only Lyndon word in N. Since we already know that w is a Lyndon word in N, this yields that N contains exactly one Lyndon word (namely, w). This solves Exercise 6.1.34(c).

(d) Let N be an n-necklace which is not aperiodic. We shall prove that N contains no Lyndon word.

Indeed, assume the contrary. Then, N contains a Lyndon word. Let w be this Lyndon word.

But the n-necklace N is not aperiodic. Thus, $|N| \neq n$ (since N is aperiodic if and only if $|N| = n$). We can see (as in the solution of Exercise 6.1.34(a)) that $|N| = [C : C_w]$. Thus, $[C : C_w] = |N| \neq n = \left[C : C^{(n)}\right]$, so that $C_w \neq C^{(n)}$. But $C^{(n)} \subset C_w$ (this can be proven just as in the solution of Exercise 6.1.34(a)). Hence, $C^{(n)}$ is a proper subset of C_w. Hence, there exists some $e \in C_w$ such that $e \notin C^{(n)}$. Consider this e.

We have $ew = w$ (since $e \in C_w$). It is now easy to see that there exists a $k \in \{1, 2, \ldots, n-1\}$ such that $c^k w = w$ [1028]. Consider this k. Lemma 13.147.1 yields $c^k w > w$ (since $w \in N \subset \mathfrak{A}^n$), which contradicts $c^k w = w$. This contradiction proves that our assumption was false. Thus, N contains no Lyndon word. This solves Exercise 6.1.34(d).

(e) If w is a Lyndon word of length n, then $[w]$ is an aperiodic n-necklace[1029]. Hence, the map

$$\text{(the set of all Lyndon words of length } n) \to \text{(the set of all aperiodic } n\text{-necklaces)}\,,$$
$$w \mapsto [w]$$

is well-defined. Denote this map by Φ. This map Φ is injective[1030] and surjective[1031]. Hence, Φ is bijective. In other words, Φ is a bijection between the set of all Lyndon words of length n and the set of all aperiodic n-necklaces. Thus, the aperiodic n-necklaces are in bijection with Lyndon words of length n. This solves Exercise 6.1.34(e).

Before we start solving Exercise 6.1.34(f), we state a lemma about words:

Lemma 13.147.2. *Let N be a positive integer, and let $w \in \mathfrak{A}^N$. Let p be a positive divisor of N. Assume that $c^p w = w$. Then, there exists a word $q \in \mathfrak{A}^p$ such that $w = q^{N/p}$.*

Proof of Lemma 13.147.2. (The following proof is overkill, but it is the simplest proof to formalize.)

Notice that N/p is a positive integer (since p is a positive divisor of N), so that $N/p - 1 \in \mathbb{N}$.

[1028] *Proof.* We know that $C^{(n)}$ acts trivially on $\mathfrak{A}^n$, and thus stabilizes w. Hence, $C^{(n)} w = \{w\}$.

By the well-known properties of the infinite cyclic group, we know that the elements c^0, c^1, ..., c^{n-1} of C form a system of unique representatives for the cosets of $C^{(n)}$ in C. Hence, every $d \in C$ belongs to the coset $c^\ell C^{(n)}$ for some $\ell \in \{0, 1, \ldots, n-1\}$. Applied to $d = e$, this yields that e belongs to the coset $c^\ell C^{(n)}$ for some $\ell \in \{0, 1, \ldots, n-1\}$. Consider this ℓ. We have $e \in c^\ell C^{(n)}$, thus $w = \underbrace{e}_{\in c^\ell C^{(n)}} w \in c^\ell \underbrace{C^{(n)} w}_{=\{w\}} = c^\ell \{w\} = \left\{c^\ell w\right\}$, so that $w = c^\ell w$. Thus, $c^\ell w = w$.

If $\ell = 0$, then $e \in \underbrace{c^\ell}_{=c^0 = 1_C} C^{(n)} = 1^C C^{(n)} = C^{(n)}$, which contradicts $e \notin C^{(n)}$. Hence, we cannot have $\ell = 0$. Thus, $\ell \in \{0, 1, \ldots, n-1\} \setminus \{0\} = \{1, 2, \ldots, n-1\}$. Hence, there exists a $k \in \{1, 2, \ldots, n-1\}$ satisfying $c^k w = w$ (namely, $k = \ell$), qed.

[1029] *Proof.* Let w be a Lyndon word of length n. Then, $[w]$ is an n-necklace. If $[w]$ was not aperiodic, then the necklace $[w]$ would contain no Lyndon word (by Exercise 6.1.34(d), applied to $N = [w]$), which would contradict the fact that this necklace $[w]$ contains the Lyndon word w. Hence, $[w]$ must be aperiodic, qed.

[1030] *Proof.* Let w and w' be two Lyndon words of length n such that $\Phi(w) = \Phi(w')$. We shall prove that $w = w'$.

We have $\Phi(w) = [w]$ (by the definition of Φ) and $\Phi(w') = [w']$ (similarly). Thus, $[w'] = \Phi(w') = \Phi(w) = [w]$.

The word w is contained in $[w]$ (obviously), and the word w' is contained in $[w'] = [w]$. Thus, both w and w' are contained in $[w]$.

We know that $[w]$ is an aperiodic n-necklace. Hence, $[w]$ contains exactly one Lyndon word (by Exercise 6.1.34(c)). Hence, any two Lyndon words contained in $[w]$ must be identical. Applying this to the two Lyndon words w and w' both contained in $[w]$, we obtain that w and w' are identical, i.e., we have $w = w'$.

Let us now forget that we fixed w and w'. We have thus proven that any two Lyndon words w and w' of length n satisfying $\Phi(w) = \Phi(w')$ must satisfy $w = w'$. In other words, the map Φ is injective, qed.

[1031] *Proof.* Let N be an aperiodic n-necklace. Hence, N contains exactly one Lyndon word (by Exercise 6.1.34(c)). Let w be this Lyndon word. The definition of Φ yields $\Phi(w) = [w] = N$ (since N is the aperiodic n-necklace containing w). Thus, $N = \Phi(w) \in \operatorname{Im} \Phi$.

Now, let us forget that we fixed N. We have thus shown that $N \in \operatorname{Im} \Phi$ for every aperiodic n-necklace N. In other words, the map Φ is surjective, qed.

We have $w = (w_1, w_2, \ldots, w_N)$ (since $w \in \mathfrak{A}^N$). Let $u = (w_1, w_2, \ldots, w_p)$ and $v = (w_{p+1}, w_{p+2}, \ldots, w_N)$. (These u and v are well-defined since $p \in \{0, 1, \ldots, N\}$.) Then,

$$\begin{aligned}\underbrace{u}_{=(w_1,w_2,\ldots,w_p)} \underbrace{v}_{=(w_{p+1},w_{p+2},\ldots,w_N)} &= (w_1, w_2, \ldots, w_p)(w_{p+1}, w_{p+2}, \ldots, w_N) \\ &= (w_1, w_2, \ldots, w_p, w_{p+1}, w_{p+2}, \ldots, w_N) \\ &= (w_1, w_2, \ldots, w_N) = w.\end{aligned}$$

On the other hand, c acts on $\mathfrak{A}^N$ by cyclically rotating N-tuples one step to the left. Hence, $c^p w$ is the result of cyclically rotating the N-tuple w to the left p times. In other words,

$$c^p w = (w_{p+1}, w_{p+2}, \ldots, w_N, w_1, w_2, \ldots, w_p) = \underbrace{(w_{p+1}, w_{p+2}, \ldots, w_N)}_{=v} \underbrace{(w_1, w_2, \ldots, w_p)}_{=u} = vu.$$

Compared with $c^p w = w = uv$, this yields $uv = vu$. Hence, Proposition 6.1.4 yields that there exist a $t \in \mathfrak{A}^*$ and two nonnegative integers n and m such that $u = t^n$ and $v = t^m$. Consider this t and these n and m. We have $u = (w_1, w_2, \ldots, w_p)$, so that $u \in \mathfrak{A}^p$ and thus $\ell(u) = p$. Hence, $p = \ell\left(\underbrace{u}_{=t^n}\right) = \ell(t^n) = n\ell(t)$, so that $n\ell(t) = p$. On the other hand, $v = (w_{p+1}, w_{p+2}, \ldots, w_N)$, so that $\ell(v) = N - p$ and thus $N - p = \ell\left(\underbrace{v}_{=t^m}\right) = \ell(t^m) = m\ell(t)$ and thus $m\ell(t) = N - p$. Now, $n\ell(t) = p \neq 0$. Hence, $n \neq 0$ and $\ell(t) \neq 0$. Now,

$$\begin{aligned}\frac{m}{n} &= \frac{m\ell(t)}{n\ell(t)} = \frac{N-p}{p} \qquad \text{(since } m\ell(t) = N - p \text{ and } n\ell(t) = p\text{)} \\ &= N/p - 1,\end{aligned}$$

so that $m = n \cdot (N/p - 1)$. Hence, $t^m = t^{n \cdot (N/p-1)} = \left(\underbrace{t^n}_{=u}\right)^{N/p-1} = u^{N/p-1}$. Now, $w = u \underbrace{v}_{=t^m = u^{N/p-1}} = uu^{N/p-1} = u^{N/p}$. Hence, there exists a word $q \in \mathfrak{A}^p$ such that $w = q^{N/p}$ (namely, $q = u$). Lemma 13.147.2 is proven. □

We also recall a lemma about the functions μ and ϕ:

Lemma 13.147.3. *Every positive integer n satisfies*

$$\sum_{d|n} \mu(d) = \delta_{n,1}; \tag{13.147.1}$$

$$\sum_{d|n} \mu(d) \frac{n}{d} = \phi(n). \tag{13.147.2}$$

Proof of Lemma 13.147.3. Both equalities (13.147.1) and (13.147.2) have been proven in the solution of Exercise 2.9.6. (Indeed, (13.147.1) is (13.84.3), and (13.147.2) is (13.84.5).) □

In Exercise 6.1.34, an action of the cyclic group C on $\mathfrak{A}^n$ was defined. In the same way, we define an action of the cyclic group C on $\mathfrak{A}^m$ for any positive integer m. The orbits of this latter C-action will be called the *m-necklaces*. (The notion of "n-necklaces" defined in Exercise 6.1.34 is clearly a particular case of this.) We define the notions of "period" and "aperiodic" for m-necklaces in the same way as they were defined for n-necklaces in Exercise 6.1.34. We will use the same notations for m-necklaces as we do for n-necklaces (i.e., the m-necklace containing a given $w \in \mathfrak{A}^m$ will be denoted by $[w]$).

Let us now state a few lemmas about the actions of C on words:

Lemma 13.147.4. *Let m be a positive integer. Let $u \in \mathfrak{A}$ be a letter. Let $v \in \mathfrak{A}^{m-1}$ be a word. Then,*

$$c(uv) = vu$$

(where we identify the letter u with the one-letter word (u)).

Proof of Lemma 13.147.4. From $v \in \mathfrak{A}^{m-1}$, we obtain $v = (v_1, v_2, \ldots, v_{m-1})$, so that

$$\underbrace{u}_{=(u)} \underbrace{v}_{=(v_1,v_2,\ldots,v_{m-1})} = (u)(v_1, v_2, \ldots, v_{m-1}) = (u, v_1, v_2, \ldots, v_{m-1}).$$

But recall that c acts on $\mathfrak{A}^m$ by cyclically rotating m-tuples one step to the left. Thus,

$$c(u, v_1, v_2, \ldots, v_{m-1}) = (v_1, v_2, \ldots, v_{m-1}, u) = \underbrace{(v_1, v_2, \ldots, v_{m-1})}_{=v} \underbrace{(u)}_{=u} = vu.$$

Hence,

$$c \underbrace{(uv)}_{=(u,v_1,v_2,\ldots,v_{m-1})} = c(u, v_1, v_2, \ldots, v_{m-1}) = vu,$$

and thus Lemma 13.147.4 is proven. □

Lemma 13.147.5. *Let n be a positive integer. Let d be a positive divisor of n. Thus, n/d is a positive integer.*

Define a map

$$\Delta : \mathfrak{A}^{n/d} \to \mathfrak{A}^n,$$
$$w \mapsto w^d$$

(where, as we recall, w^d means the d-fold concatenation $\underbrace{ww\cdots w}_{d \text{ times}}$ of w with itself).

Then:

(a) *This map Δ is well-defined.*
(b) *The map Δ is C-equivariant (meaning that $\Delta(gw) = g \cdot \Delta(w)$ for every $g \in C$ and $w \in \mathfrak{A}^{n/d}$).*
(c) *The map Δ is injective.*

Proof of Lemma 13.147.5. (a) If $w \in \mathfrak{A}^{n/d}$, then the word w^d has length

$$\ell\left(w^d\right) = d \underbrace{\ell(w)}_{\substack{=n/d \\ (\text{since } w \in \mathfrak{A}^{n/d})}} = d(n/d) = n$$

and thus satisfies $w^d \in \mathfrak{A}^n$. Hence, the map Δ is well-defined. This proves Lemma 13.147.5(a).

(b) We need to show that Δ is C-equivariant. Clearly, it is enough to prove that $\Delta(cw) = c \cdot \Delta(w)$ for every $w \in \mathfrak{A}^{n/d}$ (since c generates the group C). So let us prove this.

Let $w \in \mathfrak{A}^{n/d}$. Thus, $w = \left(w_1, w_2, \ldots, w_{n/d}\right)$. Let $\overline{w}$ denote the word $\left(w_2, w_3, \ldots, w_{n/d}\right)$. Then, (identifying the letter w_1 with the one-letter word (w_1)) we have

$$\underbrace{w_1}_{=(w_1)} \underbrace{\overline{w}}_{=\left(w_2,w_3,\ldots,w_{n/d}\right)} = (w_1)\left(w_2, w_3, \ldots, w_{n/d}\right) = \left(w_1, w_2, w_3, \ldots, w_{n/d}\right) = \left(w_1, w_2, \ldots, w_{n/d}\right) = w.$$

Thus, $w = w_1\overline{w}$, so that

$$cw = c(w_1\overline{w}) = \overline{w}w_1 \qquad (\text{by Lemma 13.147.4, applied to } m = n/d,\ u = w_1 \text{ and } v = \overline{w}).$$

Also, it is easy to see that

$$(pq)^d = p(qp)^{d-1} q \qquad \text{for any } p \in \mathfrak{A}^* \text{ and } q \in \mathfrak{A}^*. \tag{13.147.3}$$

[1032].

Taking both sides of the equality $w = w_1\overline{w}$ to the d-th power, we obtain

$$w^d = (w_1\overline{w})^d = w_1(\overline{w}w_1)^{d-1}\overline{w} \qquad (\text{by (13.147.3), applied to } p = w_1 \text{ and } q = \overline{w}).$$

[1032] *Proof.* Let $p \in \mathfrak{A}^*$ and $q \in \mathfrak{A}^*$. Then,

$$(pq)^d = \underbrace{(pq)(pq)\cdots(pq)}_{d \text{ times}} = p \underbrace{\underbrace{(qp)(qp)\cdots(qp)}_{d-1 \text{ times}}}_{=(qp)^{d-1}} q = p(qp)^{d-1} q,$$

qed.

The definition of Δ yields $\Delta(w) = w^d = w_1 (\overline{w} w_1)^{d-1} \overline{w}$. Applying c to both sides of this identity, we obtain

$$\begin{aligned} c \cdot \Delta(w) &= c \cdot \left(w_1 (\overline{w} w_1)^{d-1} \overline{w} \right) = (\overline{w} w_1)^{d-1} \overline{w} w_1 \\ &\qquad \left(\text{by Lemma 13.147.4, applied to } m = n,\ u = w_1 \text{ and } v = (\overline{w} w_1)^{d-1} \overline{w} \right) \\ &= (\overline{w} w_1)^{d-1} (\overline{w} w_1) = \left(\underbrace{\overline{w} w_1}_{=cw} \right)^d = (cw)^d . \end{aligned}$$

Compared with $\Delta(cw) = (cw)^d$ (by the definition of Δ), this yields $\Delta(cw) = c \cdot \Delta(w)$. We thus have shown that Δ is C-equivariant. This proves Lemma 13.147.5(b).

(c) We need to prove that Δ is injective. In other words, we need to prove that every $w \in \mathfrak{A}^{n/d}$ can be reconstructed from $\Delta(w)$. But this is easy: Since $\Delta(w) = w^d$ (by the definition of Δ), the word w can be obtained from $\Delta(w)$ by taking the first n/d letters of $\Delta(w)$. Hence, w can be reconstructed from $\Delta(w)$. This proves Lemma 13.147.5(c).

□

We now step to the solution of Exercise 6.1.34(f):

(f) Exercise 6.1.34(a) yields that for any n-necklace N, we have $|N| \mid n$. Hence, for any n-necklace N, the cardinality $|N|$ is a divisor of n (since $|N| \in \mathbb{N}$). Now, recall that the n-necklaces are the orbits of the C-action on $\mathfrak{A}^n$, and therefore form a set partition of the set $\mathfrak{A}^n$. Hence,

$$|\mathfrak{A}^n| = \sum_{N \text{ is an } n\text{-necklace}} |N| = \sum_{d \mid n} \sum_{\substack{N \text{ is an } n\text{-necklace;} \\ |N| = d}} d$$

(because for any n-necklace N, the cardinality $|N|$ is a divisor of n). Hence,

$$|\mathfrak{A}|^n = |\mathfrak{A}^n| = \sum_{d \mid n} \underbrace{\sum_{\substack{N \text{ is an } n\text{-necklace;} \\ |N| = d}} d}_{= d |\{N \text{ is an } n\text{-necklace} \mid |N| = d\}|} = \sum_{d \mid n} d \left| \{ N \text{ is an } n\text{-necklace} \mid |N| = d \} \right| . \qquad (13.147.4)$$

Now, for every positive integer e, let $\operatorname{Aper}(e)$ denote the set of all aperiodic e-necklaces. We shall now prove that

$$\left| \{ N \text{ is an } n\text{-necklace} \mid |N| = n/d \} \right| = \left| \operatorname{Aper}(n/d) \right| \qquad (13.147.5)$$

for every divisor d of n.

Proof of (13.147.5): Let d be a divisor of n. Then, n/d is a positive integer. We can thus define a map

$$\Delta : \mathfrak{A}^{n/d} \to \mathfrak{A}^n, \qquad w \mapsto w^d$$

(where, as we recall, w^d means the d-fold concatenation $\underbrace{ww \cdots w}_{d \text{ times}}$ of w with itself). Lemma 13.147.5(a) shows that this map Δ is well-defined. Lemma 13.147.5(b) shows that this map Δ is C-equivariant (meaning that $\Delta(gw) = g \cdot \Delta(w)$ for every $g \in C$ and $w \in \mathfrak{A}^{n/d}$). Lemma 13.147.5(c) shows that this map Δ is injective.

Since the map Δ is C-equivariant, it gives rise to a map

$$\overline{\Delta} : \left(\text{the set of all } C\text{-orbits on } \mathfrak{A}^{n/d} \right) \to \left(\text{the set of all } C\text{-orbits on } \mathfrak{A}^n \right), \qquad N \mapsto \Delta(N) .$$

Consider this map $\overline{\Delta}$. The map $\overline{\Delta}$ is injective[1033]. Furthermore, $\overline{\Delta}$ is a map from the set of all C-orbits on $\mathfrak{A}^{n/d}$ to the set of all C-orbits on $\mathfrak{A}^n$. In other words, $\overline{\Delta}$ is a map from the set of all (n/d)-necklaces to the

[1033] *Proof.* Let P and Q be two C-orbits on $\mathfrak{A}^{n/d}$ such that $\overline{\Delta}(P) = \overline{\Delta}(Q)$. We want to show that $P = Q$.

By the definition of $\overline{\Delta}$, we have $\overline{\Delta}(P) = \Delta(P)$ and $\overline{\Delta}(Q) = \Delta(Q)$. Thus, $\Delta(P) = \overline{\Delta}(P) = \overline{\Delta}(Q) = \Delta(Q)$.

The orbit P is nonempty, and thus contains an element. Let p be such an element.

set of all n-necklaces (since the C-orbits on $\mathfrak{A}^{n/d}$ are the (n/d)-necklaces, and the C-orbits on $\mathfrak{A}^n$ are the n-necklaces). It is now easy to see that

$$\overline{\Delta}\left(\operatorname{Aper}\left(n/d\right)\right) \subset \left\{N \text{ is an } n\text{-necklace} \ \mid \ \left|N\right| = n/d\right\}. \tag{13.147.6}$$

[1034] But we also have

$$\left\{N \text{ is an } n\text{-necklace} \ \mid \ \left|N\right| = n/d\right\} \subset \overline{\Delta}\left(\operatorname{Aper}\left(n/d\right)\right). \tag{13.147.7}$$

Proof of (13.147.7): Let $P \in \left\{N \text{ is an } n\text{-necklace} \ \mid \ \left|N\right| = n/d\right\}$. Thus, P is an n-necklace such that $\left|P\right| = n/d$. We shall now prove that $P \in \overline{\Delta}\left(\operatorname{Aper}\left(n/d\right)\right)$.

The set P is an n-necklace, thus a C-orbit on $\mathfrak{A}^n$, thus nonempty. Pick some $w \in P$ (this clearly exists since P is nonempty). Then P is the C-orbit of w. Hence, by the orbit-stabilizer theorem, we have $\left|P\right| = \left[C : C_w\right]$. Hence, $\left[C : C_w\right] = \left|P\right| = n/d$. Hence, C_w is a subgroup of C having index n/d. Since the only subgroup of C having index n/d is $\left\langle c^{n/d}\right\rangle$ [1035], this yields that C_w is $\left\langle c^{n/d}\right\rangle$. Hence, $C_w = \left\langle c^{n/d}\right\rangle$, so that $c^{n/d} \in \left\langle c^{n/d}\right\rangle = C_w$, and thus $c^{n/d} w = w$. Now, Lemma 13.147.2 (applied to $N = n$ and $p = n/d$) yields that there exists a word $q \in \mathfrak{A}^{n/d}$ such that $w = q^{n/(n/d)}$. Consider this q. We have $w = q^{n/(n/d)} = q^d$, whereas the definition of Δ yields $\Delta\left(q\right) = q^d$. Thus, $w = q^d = \Delta\left(\underbrace{q}_{\in [q]}\right) \in \Delta\left(\left[q\right]\right)$. The definition of $\overline{\Delta}$ yields $\overline{\Delta}\left(\left[q\right]\right) = \Delta\left(\left[q\right]\right)$. Hence, $w \in \Delta\left(\left[q\right]\right) = \overline{\Delta}\left(\left[q\right]\right)$. Thus, $\overline{\Delta}\left(\left[q\right]\right)$ is the C-orbit on $\mathfrak{A}^n$ containing w (since $\overline{\Delta}\left(\left[q\right]\right)$ is a C-orbit on $\mathfrak{A}^n$ (since $\overline{\Delta}\left(\left[q\right]\right)$ is an n-necklace)). In other words, $\overline{\Delta}\left(\left[q\right]\right)$ is the C-orbit of w. Hence, $\overline{\Delta}\left(\left[q\right]\right) = P$ (since P is the C-orbit of w).

But $P = \overline{\Delta}\left(\left[q\right]\right) = \Delta\left(\left[q\right]\right)$ and thus $\left|P\right| = \left|\Delta\left(\left[q\right]\right)\right| = \left|\left[q\right]\right|$ (since Δ is injective), so that $\left|\left[q\right]\right| = \left|P\right| = n/d$. By the definition of the period of an (n/d)-necklace, we see that the period of the (n/d)-necklace $[q]$ is $\left|\left[q\right]\right| = n/d$. In other words, the (n/d)-necklace $[q]$ is aperiodic (by the definition of "aperiodic"). In other words, $[q] \in \operatorname{Aper}\left(n/d\right)$ (since $\operatorname{Aper}\left(n/d\right)$ is the set of all aperiodic (n/d)-necklaces). Now, $P = \overline{\Delta}\left(\underbrace{\left[q\right]}_{\in \operatorname{Aper}(n/d)}\right) \in \overline{\Delta}\left(\operatorname{Aper}\left(n/d\right)\right)$.

Let us now forget that we fixed P. We thus have proven that every $P \in \left\{N \text{ is an } n\text{-necklace} \ \mid \ \left|N\right| = n/d\right\}$ satisfies $P \in \overline{\Delta}\left(\operatorname{Aper}\left(n/d\right)\right)$. In other words, $\left\{N \text{ is an } n\text{-necklace} \ \mid \ \left|N\right| = n/d\right\} \subset \overline{\Delta}\left(\operatorname{Aper}\left(n/d\right)\right)$. This proves (13.147.7).

Combining (13.147.6) with (13.147.7), we obtain

$$\overline{\Delta}\left(\operatorname{Aper}\left(n/d\right)\right) = \left\{N \text{ is an } n\text{-necklace} \ \mid \ \left|N\right| = n/d\right\}.$$

We have $\Delta\left(\underbrace{p}_{\in P}\right) \in \Delta\left(P\right) = \Delta\left(Q\right)$. Hence, there exists some $q \in Q$ such that $\Delta\left(p\right) = \Delta\left(q\right)$. Consider this q. We have $\Delta\left(p\right) = \Delta\left(q\right)$, and thus $p = q$ (since Δ is injective). The element p belongs to $P \cap Q$ (since $p \in P$ and $p = q \in Q$), and thus the two orbits P and Q have an element in common (namely, p). But any two orbits which have an element in common must be identical. Thus, P and Q are identical, i.e., we have $P = Q$.

Let us forget that we fixed P and Q. We thus have shown that any two C-orbits P and Q on $\mathfrak{A}^{n/d}$ which satisfy $\overline{\Delta}\left(P\right) = \overline{\Delta}\left(Q\right)$ must satisfy $P = Q$. In other words, the map $\overline{\Delta}$ is injective.

[1034] *Proof.* Let $M \in \operatorname{Aper}\left(n/d\right)$. Then, M is an aperiodic (n/d)-necklace (since $\operatorname{Aper}\left(n/d\right)$ is the set of all aperiodic (n/d)-necklaces). In other words, M is an (n/d)-necklace whose period is n/d.

The period of M is defined to be $\left|M\right|$. Thus, $\left|M\right|$ is n/d (since the period of M is n/d). In other words, $\left|M\right| = n/d$. But the definition of $\overline{\Delta}$ yields $\overline{\Delta}\left(M\right) = \Delta\left(M\right)$, thus

$$\begin{aligned} \left|\overline{\Delta}\left(M\right)\right| = \left|\Delta\left(M\right)\right| &= \left|M\right| \qquad \left(\text{since } \Delta \text{ is injective}\right) \\ &= n/d. \end{aligned}$$

Hence, $\overline{\Delta}\left(M\right) \in \left\{N \text{ is an } n\text{-necklace} \ \mid \ \left|N\right| = n/d\right\}$ (since $\overline{\Delta}\left(M\right)$ is an n-necklace (because $\overline{\Delta}$ is a map from the set of all (n/d)-necklaces to the set of all n-necklaces)).

Now, let us forget that we fixed M. We thus have proven that $\overline{\Delta}\left(M\right) \in \left\{N \text{ is an } n\text{-necklace} \ \mid \ \left|N\right| = n/d\right\}$ for every $M \in \operatorname{Aper}\left(n/d\right)$. In other words, $\overline{\Delta}\left(\operatorname{Aper}\left(n/d\right)\right) \subset \left\{N \text{ is an } n\text{-necklace} \ \mid \ \left|N\right| = n/d\right\}$. This proves (13.147.6).

[1035] This is a particular case of the following elementary fact: If e is a positive integer, then the only subgroup of C having index e is $\left\langle c^e\right\rangle$. (This is because $C \cong \left(\mathbb{Z}, +\right)$, and because the only subgroup of $\left(\mathbb{Z}, +\right)$ having index e is $e\mathbb{Z}$.)

Hence,

$$\left|\overline{\Delta}\left(\operatorname{Aper}\left(n/d\right)\right)\right| = \left|\left\{N \text{ is an } n\text{-necklace} \ \mid \ \left|N\right| = n/d\right\}\right|,$$

so that

$$\left|\left\{N \text{ is an } n\text{-necklace} \ \mid \ \left|N\right| = n/d\right\}\right| = \left|\overline{\Delta}\left(\operatorname{Aper}\left(n/d\right)\right)\right| = \left|\operatorname{Aper}\left(n/d\right)\right|$$

(since $\overline{\Delta}$ is injective). This proves (13.147.5).

Now, every divisor d of n satisfies

$$\begin{aligned} \left|\left\{N \text{ is an } n\text{-necklace} \ \mid \ \left|N\right| = \underbrace{d}_{=n/(n/d)}\right\}\right| &= \left|\left\{N \text{ is an } n\text{-necklace} \ \mid \ \left|N\right| = n/\left(n/d\right)\right\}\right| \\ &= \left|\operatorname{Aper}\left(\underbrace{n/\left(n/d\right)}_{=d}\right)\right| \\ &\qquad \left(\text{by (13.147.5), applied to } n/d \text{ instead of } d\right) \\ &= \left|\operatorname{Aper}\left(d\right)\right|. \end{aligned} \tag{13.147.8}$$

Now, (13.147.4) becomes

$$\left|\mathfrak{A}\right|^{n} = \sum_{d\mid n} d \underbrace{\left|\left\{N \text{ is an } n\text{-necklace} \ \mid \ \left|N\right| = d\right\}\right|}_{\substack{=\left|\operatorname{Aper}(d)\right| \\ \text{(by (13.147.8))}}} = \sum_{d\mid n} d\left|\operatorname{Aper}\left(d\right)\right| = \sum_{e\mid n} e\left|\operatorname{Aper}\left(e\right)\right| \tag{13.147.9}$$

(here, we renamed the summation index d as e). Now,

$$\begin{aligned} \sum_{d\mid n} \mu\left(d\right) \underbrace{\left|\mathfrak{A}\right|^{n/d}}_{\substack{=\sum_{e\mid n/d} e\left|\operatorname{Aper}(e)\right| \\ \text{(by (13.147.9), applied to} \\ n/d \text{ instead of } n)}} &= \sum_{d\mid n} \mu\left(d\right) \sum_{e\mid n/d} e\left|\operatorname{Aper}\left(e\right)\right| = \underbrace{\sum_{d\mid n} \sum_{e\mid n/d}}_{=\sum_{e\mid n} \sum_{d\mid n/e}} \mu\left(d\right) e\left|\operatorname{Aper}\left(e\right)\right| \\ &= \sum_{e\mid n} \sum_{d\mid n/e} \mu\left(d\right) e\left|\operatorname{Aper}\left(e\right)\right| = \sum_{e\mid n} \underbrace{\left(\sum_{d\mid n/e} \mu\left(d\right)\right)}_{\substack{=\delta_{n/e,1} \\ \text{(by (13.147.1), applied} \\ \text{to } n/e \text{ instead of } n)}} e\left|\operatorname{Aper}\left(e\right)\right| \\ &= \sum_{e\mid n} \underbrace{\delta_{n/e,1}}_{=\delta_{n,e}} e\left|\operatorname{Aper}\left(e\right)\right| = \sum_{e\mid n} \delta_{n,e} e\left|\operatorname{Aper}\left(e\right)\right| = n\left|\operatorname{Aper}\left(n\right)\right|. \end{aligned}$$

Solving this for $\left|\operatorname{Aper}\left(n\right)\right|$, we obtain

$$\left|\operatorname{Aper}\left(n\right)\right| = \frac{1}{n} \sum_{d\mid n} \mu\left(d\right) \left|\mathfrak{A}\right|^{n/d}. \tag{13.147.10}$$

Now, let us recall that $\operatorname{Aper}\left(n\right)$ is the set of all aperiodic n-necklaces. Thus,

$$\left|\operatorname{Aper}\left(n\right)\right| = \left(\text{the number of all aperiodic } n\text{-necklaces}\right).$$

Hence,

$$\left(\text{the number of all aperiodic } n\text{-necklaces}\right) = \left|\operatorname{Aper}\left(n\right)\right| = \frac{1}{n} \sum_{d\mid n} \mu\left(d\right) \left|\mathfrak{A}\right|^{n/d}.$$

This solves Exercise 6.1.34(f).

(g) We shall use the notations we introduced in the solution of Exercise 6.1.34(f).

We have seen (in the solution of Exercise 6.1.34(f)) that for any n-necklace N, the cardinality $|N|$ is a divisor of n. Hence,

$$\begin{aligned}
&(\text{the number of all } n\text{-necklaces}) \\
&= \sum_{d|n} \underbrace{(\text{the number of all } n\text{-necklaces } N \text{ such that } |N| = d)}_{\substack{=|\{N \text{ is an } n\text{-necklace} \ | \ |N|=d\}|=|\operatorname{Aper}(d)| \\ (\text{by } (13.147.8))}} \\
&= \sum_{d|n} |\operatorname{Aper}(d)| = \sum_{e|n} \underbrace{|\operatorname{Aper}(e)|}_{\substack{=\frac{1}{e}\sum_{d|e}\mu(d)|\mathfrak{A}|^{e/d} \\ (\text{by } (13.147.10)\text{, applied to } e \\ \text{instead of } n)}} \qquad (\text{here, we renamed the summation index } d \text{ as } e) \\
&= \sum_{e|n} \frac{1}{e} \sum_{d|e} \mu(d) |\mathfrak{A}|^{e/d} = \underbrace{\sum_{e|n} \sum_{d|e}}_{=\sum_{d|n} \sum_{\substack{e|n; \\ d|e}}} \frac{1}{e} \mu(d) |\mathfrak{A}|^{e/d} = \sum_{d|n} \sum_{\substack{e|n; \\ d|e}} \frac{1}{e} \mu(d) |\mathfrak{A}|^{e/d} \\
&= \underbrace{\sum_{d|n} \sum_{e|n/d}}_{=\sum_{e|n}\sum_{d|n/e}} \frac{1}{de} \mu(d) |\mathfrak{A}|^{e} \qquad (\text{here, we have substituted } de \text{ for } e \text{ in the inner sum})
\end{aligned}$$

(13.147.11)
$$= \sum_{e|n} \sum_{d|n/e} \frac{1}{de} \mu(d) |\mathfrak{A}|^{e}.$$

From (13.147.2), we have

$$\phi(n) = \sum_{d|n} \mu(d) \frac{n}{d}. \qquad (13.147.12)$$

Now,

$$\begin{aligned}
\frac{1}{n} \sum_{d|n} \phi(d) |\mathfrak{A}|^{n/d} &= \frac{1}{n} \sum_{d|n} \phi(n/d) |\mathfrak{A}|^{d} \qquad (\text{here, we have substituted } n/d \text{ for } d \text{ in the sum}) \\
&= \frac{1}{n} \sum_{e|n} \underbrace{\phi(n/e)}_{\substack{=\sum_{d|n/e}\mu(d)\frac{n/e}{d} \\ (\text{by } (13.147.12)\text{, applied to } n/e \\ \text{instead of } n)}} |\mathfrak{A}|^{e} \\
&\qquad (\text{here, we have renamed the summation index } d \text{ as } e) \\
&= \frac{1}{n} \sum_{e|n} \left(\sum_{d|n/e} \mu(d) \frac{n/e}{d} \right) |\mathfrak{A}|^{e} = \frac{1}{n} \sum_{e|n} \sum_{d|n/e} \mu(d) \frac{n/e}{d} |\mathfrak{A}|^{e} = \frac{1}{n} \sum_{e|n} \sum_{d|n/e} \mu(d) \frac{n}{de} |\mathfrak{A}|^{e} \\
&= \sum_{e|n} \sum_{d|n/e} \frac{1}{de} \mu(d) |\mathfrak{A}|^{e}.
\end{aligned}$$

Compared with (13.147.11), this yields

$$(\text{the number of all } n\text{-necklaces}) = \frac{1}{n} \sum_{d|n} \phi(d) |\mathfrak{A}|^{n/d}.$$

This solves Exercise 6.1.34(g).

(h) *Alternative solution of Exercise 6.1.29:* Let n be a positive integer. Exercise 6.1.34(e) yields that the aperiodic n-necklaces are in bijection with Lyndon words of length n. Hence,

$$(\text{the number of all aperiodic } n\text{-necklaces}) = (\text{the number of all Lyndon words of length } n)\,,$$

so that

$$\begin{aligned}&(\text{the number of all Lyndon words of length } n)\\ &= (\text{the number of all aperiodic } n\text{-necklaces}) = \frac{1}{n}\sum_{d\mid n}\mu\left(d\right)\left|\mathfrak{A}\right|^{n/d}\end{aligned}$$

(by Exercise 6.1.34(f)). Thus,

$$(\text{the number of all Lyndon words of length } n) = \frac{1}{n}\sum_{d\mid n}\mu\left(d\right)\underbrace{\left|\mathfrak{A}\right|^{n/d}}_{\substack{=q^{n/d}\\ (\text{since } |\mathfrak{A}|=q)}} = \frac{1}{n}\sum_{d\mid n}\mu\left(d\right)q^{n/d}.$$

This solves Exercise 6.1.29 again. Thus, Exercise 6.1.34(h) is solved.

(i) *First solution of Exercise 6.1.34(i):* Let $q\in\mathbb{Z}$. We need to show that $n \mid \sum_{d\mid n}\mu\left(d\right)q^{n/d}$ and $n \mid \sum_{d\mid n}\phi\left(d\right)q^{n/d}$.

Let r be the remainder of q modulo n. Then, $r\in\{0,1,\ldots,n-1\}\subset\mathbb{N}$. Fix a finite totally ordered set $\mathfrak{A}$ containing r elements. (Such an $\mathfrak{A}$ exists, since $r\in\mathbb{N}$. For example, we can set $\mathfrak{A}=\{1,2,\ldots,r\}$.) Exercise 6.1.34(f) shows that the number of all aperiodic n-necklaces is $\dfrac{1}{n}\sum_{d\mid n}\mu\left(d\right)\left|\mathfrak{A}\right|^{n/d}$. Hence, $\dfrac{1}{n}\sum_{d\mid n}\mu\left(d\right)\left|\mathfrak{A}\right|^{n/d}$ is an integer (since the number of all aperiodic n-necklaces is an integer). In other words, $n\mid\sum_{d\mid n}\mu\left(d\right)\left|\mathfrak{A}\right|^{n/d}$, so that $\sum_{d\mid n}\mu\left(d\right)\left|\mathfrak{A}\right|^{n/d}\equiv 0\mod n$. But r is the remainder of q modulo n. Thus, $r\equiv q\mod n$, so that $q\equiv r=\left|\mathfrak{A}\right|\mod n$ (since the set $\mathfrak{A}$ has r elements). Now, $\sum_{d\mid n}\mu\left(d\right)\underbrace{q^{n/d}}_{\substack{\equiv|\mathfrak{A}|^{n/d}\mod n\\ (\text{since } q\equiv|\mathfrak{A}|\mod n)}}\equiv\sum_{d\mid n}\mu\left(d\right)\left|\mathfrak{A}\right|^{n/d}\equiv 0\mod n$, so that $n\mid\sum_{d\mid n}\mu\left(d\right)q^{n/d}$.

Also, Exercise 6.1.34(g) shows that the number of all n-necklaces is $\dfrac{1}{n}\sum_{d\mid n}\phi\left(d\right)\left|\mathfrak{A}\right|^{n/d}$. Hence, $\dfrac{1}{n}\sum_{d\mid n}\phi\left(d\right)\left|\mathfrak{A}\right|^{n/d}$ is an integer (since the number of all n-necklaces is an integer). In other words, $n\mid\sum_{d\mid n}\phi\left(d\right)\left|\mathfrak{A}\right|^{n/d}$, so that $\sum_{d\mid n}\phi\left(d\right)\left|\mathfrak{A}\right|^{n/d}\equiv 0\mod n$. Now, $\sum_{d\mid n}\phi\left(d\right)\underbrace{q^{n/d}}_{\substack{\equiv|\mathfrak{A}|^{n/d}\mod n\\ (\text{since } q\equiv|\mathfrak{A}|\mod n)}}\equiv\sum_{d\mid n}\phi\left(d\right)\left|\mathfrak{A}\right|^{n/d}\equiv 0\mod n$, so that $n\mid\sum_{d\mid n}\phi\left(d\right)q^{n/d}$. The solution of Exercise 6.1.34(i) is thus complete.

Second solution of Exercise 6.1.34(i): Forget that we fixed n. Let $q\in\mathbb{Z}$. Let A denote the ring $\mathbb{Z}$. For every $n\in\{1,2,3,\ldots\}$, let φ_n denote the identity endomorphism id of A. Exercise 2.9.8 yields (among other things) that the seven equivalent assertions $\mathcal{C}$, $\mathcal{D}$, $\mathcal{E}$, $\mathcal{F}$, $\mathcal{G}$, $\mathcal{H}$ and $\mathcal{J}$ of Exercise 2.9.6 are satisfied for the family $\left(b_n\right)_{n\geq 1}=\left(q^n\right)_{n\geq 1}$. In particular, Assertion $\mathcal{F}$ of Exercise 2.9.6 is satisfied for the family $\left(b_n\right)_{n\geq 1}=\left(q^n\right)_{n\geq 1}$. In other words, every positive integer n satisfies

$$\sum_{d\mid n}\mu\left(d\right)\varphi_d\left(q^{n/d}\right)\in n\mathbb{Z}. \tag{13.147.13}$$

Also, Assertion $\mathcal{G}$ of Exercise 2.9.6 is satisfied for the family $\left(b_n\right)_{n\geq 1}=\left(q^n\right)_{n\geq 1}$ (since the seven equivalent assertions $\mathcal{C}$, $\mathcal{D}$, $\mathcal{E}$, $\mathcal{F}$, $\mathcal{G}$, $\mathcal{H}$ and $\mathcal{J}$ of Exercise 2.9.6 are satisfied for the family $\left(b_n\right)_{n\geq 1}=\left(q^n\right)_{n\geq 1}$). In other words, every positive integer n satisfies

$$\sum_{d\mid n}\phi\left(d\right)\varphi_d\left(q^{n/d}\right)\in n\mathbb{Z}. \tag{13.147.14}$$

Now, fix a positive integer n. We can rewrite (13.147.13) as $n \mid \sum_{d\mid n} \mu(d) \varphi_d\left(q^{n/d}\right)$. Since

$$\sum_{d\mid n} \mu(d) \underbrace{\varphi_d}_{=\mathrm{id}} \left(q^{n/d}\right) = \sum_{d\mid n} \mu(d) q^{n/d},$$

this simplifies to $n \mid \sum_{d\mid n} \mu(d) q^{n/d}$. Also, we can rewrite (13.147.14) as $n \mid \sum_{d\mid n} \phi(d) \varphi_d\left(q^{n/d}\right)$. Since

$$\sum_{d\mid n} \phi(d) \underbrace{\varphi_d}_{=\mathrm{id}} \left(q^{n/d}\right) = \sum_{d\mid n} \phi(d) q^{n/d},$$

this simplifies to $n \mid \sum_{d\mid n} \phi(d) q^{n/d}$. Exercise 6.1.34(i) is thus solved again.

13.148. **Solution to Exercise 6.1.35.** *Solution to Exercise 6.1.35.* There exists a $u \in \mathfrak{A}^*$ such that $w = uv$ (since v is a suffix of w). Consider this u.

The word v is Lyndon and thus nonempty. Hence, $\ell(v) > 0$.

The exercise asks us to prove the logical equivalence

$$(t \text{ is the longest Lyndon suffix of } wt) \iff (\text{we do not have } v < t). \qquad (13.148.1)$$

We shall prove the $\Longrightarrow$ and $\Longleftarrow$ parts of this equivalence separately:

$\Longrightarrow$: Assume that t is the longest Lyndon suffix of wt. We need to prove that we do not have $v < t$.

Assume the contrary. Thus, $v < t$. Now, both v and t are Lyndon words. Hence, Proposition 6.1.16(a) (applied to v and t instead of u and v) yields that the word vt is Lyndon. But vt is a suffix of wt (since v is a suffix of w). Hence, vt is a Lyndon suffix of wt. This Lyndon suffix is clearly longer than t (since $\ell(vt) = \underbrace{\ell(v)}_{>0} + \ell(t) > \ell(t)$), which flies in the face of the fact that t is the longest Lyndon suffix of wt. This contradiction shows that our assumption was wrong. Hence, we have proven that we do not have $v < t$. This proves the $\Longrightarrow$ part of the equivalence (13.148.1).

$\Longleftarrow$: Assume that we do not have $v < t$. We must show that t is the longest Lyndon suffix of wt.

Assume the contrary. Then, t is not the longest Lyndon suffix of wt. Since t is a Lyndon suffix of wt, this means that there exists a Lyndon suffix r of wt which satisfies $\ell(r) > \ell(t)$. Let q be the shortest such suffix.

Thus, q is a Lyndon suffix of wt and satisfies $\ell(q) > \ell(t)$. Moreover, q is the shortest such Lyndon suffix. Hence, if r is a Lyndon suffix of wt which satisfies $\ell(r) > \ell(t)$, then

$$\ell(r) \geq \ell(q). \qquad (13.148.2)$$

Both q and t are suffixes of wt, and the suffix q begins earlier (since $\ell(q) > \ell(t)$). Thus, there exists a nonempty suffix g of w such that $q = gt$. Consider this g.

The word t is nonempty (since it is Lyndon) and a suffix of q (since $q = gt$). Thus, Corollary 6.1.15 (applied to q and t instead of w and v) yields $t \geq q$ (since q is Lyndon).

The word t is Lyndon (by assumption) and a proper suffix of q (since $q = gt$ and since g is nonempty). In other words, t is a Lyndon proper suffix[1036] of q. Moreover, the word t is the longest Lyndon proper suffix of q [1037]. In other words, the word t is the longest proper suffix of q such that t is Lyndon. Also, $\ell(q) > \ell(t) \geq 1$ (since t is nonempty). Thus, q is a Lyndon word of length > 1. Hence, Exercise 6.1.31(a) (applied to q, g and t instead of w, u and v) shows that the words g and t are Lyndon. In particular, g is a Lyndon suffix of w (since g is Lyndon and a suffix of w). Since v is the longest Lyndon suffix of w, this shows that g is at most as long as v. In other words, $\ell(g) \leq \ell(v)$.

Since t is nonempty, we have $g < gt$.

[1036] Of course, a Lyndon proper suffix of q just means a proper suffix z of q such that z is Lyndon.

[1037] *Proof.* Let r be a Lyndon proper suffix of q. We shall prove that $\ell(r) \leq \ell(t)$.

Indeed, assume the contrary (for the sake of contradiction). Thus, $\ell(r) > \ell(t)$. Now, since r is a suffix of q, and since q (in turn) is a suffix of wt, we see that r is a suffix of wt. Thus, r is a Lyndon suffix of wt (since r is Lyndon). Thus, (13.148.2) shows that $\ell(r) \geq \ell(q)$. But since r is a proper suffix of q, we have $\ell(r) < \ell(q)$. This contradicts $\ell(r) \geq \ell(q)$. This contradiction proves that our assumption was wrong. Hence, we have shown that $\ell(r) \leq \ell(t)$.

Now, let us forget that we fixed r. We thus have shown that any Lyndon proper suffix r of q satisfies $\ell(r) \leq \ell(t)$. In other words, any Lyndon proper suffix of q is at most as long as t. Since t is a Lyndon proper suffix of q, this shows that the word t is the **longest** Lyndon proper suffix of q. Qed.

Both g and v are suffixes of w, and the suffix g begins no earlier than v (since $\ell(g) \leq \ell(v)$). Therefore, g is a suffix of v. Hence, Corollary 6.1.15 (applied to v and g instead of w and v) yields $g \geq v$. Hence, $v \leq g < gt = q \leq t$ (since $t \geq q$). This contradicts the fact that we do not have $v < t$. Thus, we have obtained a contradiction. Our assumption was therefore wrong, and we have shown that t is the longest Lyndon suffix of wt. This proves the $\Longleftarrow$ part of the equivalence (13.148.1).

We have now proven both parts of the equivalence (13.148.1), and thus solved Exercise 6.1.35.

[*Remark:* Exercise 6.1.35 still holds in the partial-order setting[1038]. In fact, the solution we have given above still applies in this setting.]

13.149. **Solution to Exercise 6.1.36.** *Solution to Exercise 6.1.36.* Let a be the first letter of the word w. (This is well-defined, since w has length $> 1 > 0$.) We consider a as a one-letter word; thus, $\ell(a) = 1$. Clearly, a is a prefix of w (since a is the first letter of w). Hence, there exists a word w' such that $w = aw'$. Consider this w'. The word w' is nonempty (since w has length > 1).

Now, if h is any word, then

$$\text{the suffixes of } h \text{ are precisely the proper suffixes of } ah \tag{13.149.1}$$

(since a is a single letter). Applying this to $h = w'$, we see that the suffixes of w' are precisely the proper suffixes of aw'. In other words, the suffixes of w' are precisely the proper suffixes of w (since $aw' = w$). Thus, v is the longest Lyndon suffix of w' (since v is the longest Lyndon proper suffix of w).

On the other hand, applying (13.149.1) to $h = w't$, we see that the suffixes of $w't$ are precisely the proper suffixes of $aw't$. In other words, the suffixes of $w't$ are precisely the proper suffixes of wt (since $aw' = w$).

But Exercise 6.1.35 (applied to w' instead of w) shows that

$$(t \text{ is the longest Lyndon suffix of } w't \text{ if and only if we do not have } v < t).$$

Since the suffixes of $w't$ are precisely the proper suffixes of wt, this result rewrites as follows:

$$(t \text{ is the longest Lyndon proper suffix of } wt \text{ if and only if we do not have } v < t).$$

This solves Exercise 6.1.36.

[*Remark:* Exercise 6.1.36 still holds in the partial-order setting[1039]. In fact, the solution we have given above still applies in this setting.]

13.150. **Solution to Exercise 6.1.39.** *Solution to Exercise 6.1.39.* Recall the definition of $\operatorname{stf} w$. It says that $\operatorname{stf} w = (u, v)$, where u and v are defined as follows:

- The word v is defined as the longest proper suffix of w such that v is Lyndon.
- The word u is defined as the nonempty word such that $w = uv$.

Consider these u and v. Thus, $(u, v) = \operatorname{stf} w = (g, h)$. In other words, $u = g$ and $v = h$.

We know that v is the longest proper suffix of w such that v is Lyndon. In other words, v is the longest Lyndon proper suffix of w. In other words, h is the longest Lyndon proper suffix of w (since $v = h$). This solves Exercise 6.1.39(a).

We have $w = \underbrace{u}_{=g}\underbrace{v}_{=h} = gh$. This solves Exercise 6.1.39(b).

From Exercise 6.1.31(b), we conclude that $u < w < v$. Since $u = g$, $w = gh$ and $v = h$, this rewrites as $g < gh < h$. This solves Exercise 6.1.39(c).

From Exercise 6.1.31(a), we conclude that the words u and v are Lyndon. Thus, the word u is Lyndon. In other words, the word g is Lyndon (since $u = g$). This solves solves Exercise 6.1.39(d).

The word v is the longest proper suffix of w such that v is Lyndon. In particular, the word v is Lyndon. In other words, $v \in \mathfrak{L}$.

We have $g \in \mathfrak{L}$ (since the word g is Lyndon) and $h = v \in \mathfrak{L}$. Also, the word g is nonempty (since g is Lyndon), and thus we have $\ell(g) \geq 1$. Furthermore, the word h is Lyndon (since $h \in \mathfrak{L}$) and thus nonempty.

[1038] See Exercise 6.1.33 for an explanation of what the partial-order setting is.

[1039] See Exercise 6.1.33 for an explanation of what the partial-order setting is.

Hence, $\ell(h) \geq 1$. Now, $\ell\left(\underbrace{w}_{=gh}\right) = \ell(gh) = \ell(g) + \underbrace{\ell(h)}_{\geq 1} \geq \ell(g) + 1 > \ell(g)$, so that $\ell(g) < \ell(w)$. Also, $\ell\left(\underbrace{w}_{=gh}\right) = \ell(gh) = \underbrace{\ell(g)}_{\geq 1} + \ell(h) \geq 1 + \ell(h) > \ell(h)$, so that $\ell(g) < \ell(w)$. Thus, Exercise 6.1.39(e) is solved.

(f) Exercise 6.1.36 (applied to $v = h$) shows that t is the longest Lyndon proper suffix of wt if and only if we do not have $h < t$ (since h is the longest Lyndon proper suffix of w). This solves Exercise 6.1.39(f).

13.151. **Solution to Exercise 6.1.40.** *Solution to Exercise 6.1.40.* We define a binary relation $\sim$ on the set $\mathfrak{A}^*$ as follows: If w and w' are two words, then we write $w \sim w'$ if and only if w' is a permutation of the word w (that is, if and only if there exists a permutation $\sigma \in \mathfrak{S}_k$ satisfying $w' = \left(w_{\sigma(1)}, w_{\sigma(2)}, \ldots, w_{\sigma(k)}\right)$, where $k = \ell(w)$). We notice the following properties of the relation $\sim$:

- The relation $\sim$ is an equivalence relation; in other words, it is reflexive, symmetric and transitive.
- If w and w' are two words satisfying $w \sim w'$, then $\ell(w) = \ell(w')$.
- If u, v, u' and v' are four words satisfying $u \sim u'$ and $v \sim v'$, then $uv \sim u'v'$. We shall refer to this fact as the *monoidality* of the relation $\sim$.
- If u and v are any two words, then $uv \sim vu$.

We also recall a fundamental property of Lie algebras (one of the forms of the Jacobi identity):

- Every three elements x, y and z of a Lie algebra $\mathfrak{k}$ satisfy

$$[[x,y],z] = [[x,z],y] - [[y,z],x]. \qquad (13.151.1)$$

We can now finally come to the solution of Exercise 6.1.40.

For every $h \in \mathfrak{A}^*$ and $s \in \mathfrak{A}^*$, we define a subset $\mathfrak{L}_{h,s}$ of $\mathfrak{L}$ by

$$\mathfrak{L}_{h,s} = \{w \in \mathfrak{L} \mid w \sim h \text{ and } w < s\}.$$

For every $h \in \mathfrak{A}^*$ and $s \in \mathfrak{A}^*$, we define a $\mathbf{k}$-submodule $B_{h,s}$ of B by

$$B_{h,s} = \sum_{w \in \mathfrak{L}_{h,s}} \mathbf{k} b_w.$$

(In other words, for every $h \in \mathfrak{A}^*$ and $s \in \mathfrak{A}^*$, we define $B_{h,s}$ as the $\mathbf{k}$-linear span of the elements b_w with $w \in \mathfrak{L}_{h,s}$.) If h, s, g and t are four words satisfying $h \sim g$ and $s < t$, then

$$B_{h,s} \subset B_{g,t} \qquad (13.151.2)$$

[1040].

(a) We claim that

$$[b_p, b_q] \in B_{pq,q} \qquad \text{for every } (p,q) \in \mathfrak{L} \times \mathfrak{L} \text{ satisfying } p < q. \qquad (13.151.3)$$

Proof of (13.151.3): We can WLOG assume that the alphabet $\mathfrak{A}$ is finite[1041]. Assume this.

[1040] *Proof of (13.151.2):* Let h, s, g and t be four words satisfying $h \sim g$ and $s < t$.

Let $v \in \mathfrak{L}_{h,s}$. Thus, $v \in \mathfrak{L}_{h,s} = \{w \in \mathfrak{L} \mid w \sim h \text{ and } w < s\}$ (by the definition of $\mathfrak{L}_{h,s}$). In other words, v is an element of $\mathfrak{L}$ and satisfies $v \sim h$ and $v < s$. From $v \sim h$ and $h \sim g$, we obtain $v \sim g$ (since the relation $\sim$ is transitive). Also, $v < s < t$. Thus, v is an element of $\mathfrak{L}$ and satisfies $v \sim g$ and $v < t$. In other words, $v \in \{w \in \mathfrak{L} \mid w \sim g \text{ and } w < t\} = \mathfrak{L}_{g,t}$ (since $\mathfrak{L}_{g,t}$ is defined to be $\{w \in \mathfrak{L} \mid w \sim g \text{ and } w < t\}$).

Now, let us forget that we fixed v. We thus have proven that every $v \in \mathfrak{L}_{h,s}$ satisfies $v \in \mathfrak{L}_{g,t}$. In other words, $\mathfrak{L}_{h,s} \subset \mathfrak{L}_{g,t}$. Now, the definition of $B_{h,s}$ yields $B_{h,s} = \sum_{w \in \mathfrak{L}_{h,s}} \mathbf{k} b_w \subset \sum_{w \in \mathfrak{L}_{g,t}} \mathbf{k} b_w$ (since $\mathfrak{L}_{h,s} \subset \mathfrak{L}_{g,t}$). Since $B_{g,t} = \sum_{w \in \mathfrak{L}_{g,t}} \mathbf{k} b_w$ (by the definition of $B_{g,t}$), this rewrites as $B_{h,s} \subset B_{g,t}$. This proves (13.151.2).

[1041] In fact, assume that (13.151.3) is proven in the case when the alphabet $\mathfrak{A}$ is finite. Now, let $\mathfrak{A}$ be arbitrary. We must prove (13.151.3) for this $\mathfrak{A}$.

Fix $(p,q) \in \mathfrak{L} \times \mathfrak{L}$ satisfying $p < q$. We need to prove that $[b_p, b_q] \in B_{pq,q}$. Let $\mathfrak{B}$ denote the set of all letters that appear in (at least) one of the words p and q. Then, $\mathfrak{B}$ is a finite subset of $\mathfrak{A}$, and the words p and q belong to $\mathfrak{B}^*$.

Let $\mathfrak{L}'$ denote the set of all Lyndon words over the alphabet $\mathfrak{B}$. Clearly, a word $w \in \mathfrak{B}^*$ is Lyndon as a word over the alphabet $\mathfrak{B}$ if and only if it is Lyndon as a word over the alphabet $\mathfrak{A}$. Thus, $\mathfrak{L}' = \mathfrak{L} \cap \mathfrak{B}^*$, so that p and q belong to $\mathfrak{L}'$. Also, for every given $w \in \mathfrak{L}'$ of length > 1, the pair $\operatorname{stf} w$ does not depend on whether w is considered as a Lyndon word over the

We shall prove (13.151.3) by strong induction over $\ell(pq)$:

Induction step: Let N be a nonnegative integer. Assume that (13.151.3) holds in the case when $\ell(pq) < N$. We need to prove that (13.151.3) holds in the case when $\ell(pq) = N$.

We have assumed that (13.151.3) holds in the case when $\ell(pq) < N$. In other words, we have

$$[b_p, b_q] \in B_{pq,q} \qquad \text{for every } (p,q) \in \mathfrak{L} \times \mathfrak{L} \text{ satisfying } p < q \text{ and } \ell(pq) < N. \tag{13.151.4}$$

We now must prove that (13.151.3) holds in the case when $\ell(pq) = N$. In other words, we must prove that

$$[b_p, b_q] \in B_{pq,q} \qquad \text{for every } (p,q) \in \mathfrak{L} \times \mathfrak{L} \text{ satisfying } p < q \text{ and } \ell(pq) = N. \tag{13.151.5}$$

Let $\mathfrak{G}$ be the set $\bigcup_{i=0}^{N} \mathfrak{A}^i$. This set $\mathfrak{G}$ is finite (since the set $\mathfrak{A}$ is finite) and totally ordered (by the lexicographic order). Thus, for every $w \in \mathfrak{G}$, we can define a nonnegative integer $\rho(w)$ by

$$\rho(w) = |\{g \in \mathfrak{G} \mid g < w\}|.$$

In other words, for every $w \in \mathfrak{G}$, we define $\rho(w)$ to be the number of all $g \in \mathfrak{G}$ which are smaller than w. It is clear that if w and w' are two elements of $\mathfrak{G}$ satisfying $w < w'$, then

$$\rho(w) < \rho(w') \tag{13.151.6}$$

[1042].

For every $(p,q) \in \mathfrak{L} \times \mathfrak{L}$ satisfying $p < q$ and $\ell(pq) = N$, we have $q \in \mathfrak{G}$ [1043], and thus $\rho(q)$ is well-defined. We are thus going to prove (13.151.5) by strong induction over $\rho(q)$ [1044]:

Induction step: Let K be a nonnegative integer. Assume that (13.151.5) holds in the case when $\rho(q) < K$. We need to prove that (13.151.5) holds in the case when $\rho(q) = K$.

We have assumed that (13.151.5) holds in the case when $\rho(q) < K$. In other words, we have

$$[b_p, b_q] \in B_{pq,q} \qquad \text{for every } (p,q) \in \mathfrak{L} \times \mathfrak{L} \text{ satisfying } p < q \text{ and } \ell(pq) = N \text{ and } \rho(q) < K. \tag{13.151.7}$$

Now, let $(p,q) \in \mathfrak{L} \times \mathfrak{L}$ be such that $p < q$ and $\ell(pq) = N$ and $\rho(q) = K$. We are going to prove that $[b_p, b_q] \in B_{pq,q}$.

We have $(p,q) \in \mathfrak{L} \times \mathfrak{L}$. In other words, the words p and q are Lyndon. Proposition 6.1.16(a) (applied to $u = p$ and $v = q$) thus shows that the word pq is Lyndon. In other words, $pq \in \mathfrak{L}$.

Furthermore, Proposition 6.1.16(b) (applied to $u = p$ and $v = q$) shows that $pq < q$.

The definition of $\mathfrak{L}_{pq,q}$ yields $\mathfrak{L}_{pq,q} = \{w \in \mathfrak{L} \mid w \sim pq \text{ and } w < q\}$. The definition of $B_{pq,q}$ yields $B_{pq,q} = \sum_{w \in \mathfrak{L}_{pq,q}} \mathbf{k} b_w$.

The words p and q are nonempty (since they are Lyndon), and thus have length ≥ 1 each. Hence, the word pq has length $\geq 1 + 1 > 1$. Hence, pq is a Lyndon word of length > 1 (since pq is Lyndon and since pq has length $\ell(pq) > 1$). Therefore, $\operatorname{stf}(pq)$ is well-defined. If $\operatorname{stf}(pq) = (p,q)$, then it is easy to see that

alphabet $\mathfrak{A}$ or as a Lyndon word over the alphabet $\mathfrak{B}$ (because the definition of $\operatorname{stf} w$ involves only suffixes of w, and all of these suffixes belong to $\mathfrak{B}^*$).

For every $h \in \mathfrak{B}^*$ and $s \in \mathfrak{B}^*$, let us define the set $\mathfrak{L}'_{h,s}$, the $\mathbf{k}$-module B' and the $\mathbf{k}$-module $B'_{h,s}$ in the same way as we have defined the set $\mathfrak{L}_{h,s}$, the $\mathbf{k}$-module B and the $\mathbf{k}$-module $B_{h,s}$, but using the alphabet $\mathfrak{B}$ instead of $\mathfrak{A}$. (Thus, $\mathfrak{L}'_{h,s} = \{w \in \mathfrak{L}' \mid w \sim h \text{ and } w < s\}$, $B' = \left(\text{the } \mathbf{k}\text{-submodule of } \mathfrak{g} \text{ spanned by the family } (b_w)_{w \in \mathfrak{L}'}\right)$ and $B'_{h,s} = \sum_{w \in \mathfrak{L}'_{h,s}} \mathbf{k} b_w$.)

Now, whenever w is a Lyndon word over $\mathfrak{B}$ of length > 1, we must have $w \in \mathfrak{L}' = \mathfrak{L} \cap \mathfrak{B}^* \subset \mathfrak{L}$ and thus $b_w = [b_u, b_v]$, where $(u,v) = \operatorname{stf} w$ (according to (6.1.2)). Here, when we speak of $\operatorname{stf} w$, we are regarding w as a Lyndon word over $\mathfrak{A}$, but as we have already explained, the result is the same if we regard w as a Lyndon word over $\mathfrak{B}$ instead. Thus, we can apply (13.151.3) to $\mathfrak{B}$ and $\mathfrak{L}'$ instead of $\mathfrak{A}$ and $\mathfrak{L}$ (since we assumed that (13.151.3) is proven in the case when the alphabet $\mathfrak{A}$ is finite), and obtain $[b_p, b_q] \in B'_{pq,q}$.

But we have $\mathfrak{L}' = \mathfrak{L} \cap \mathfrak{B}^* \subset \mathfrak{L}$, thus $\mathfrak{L}'_{pq,q} \subset \mathfrak{L}_{pq,q}$ and therefore $B'_{pq,q} \subset B_{pq,q}$ (actually, a little thought shows that $B'_{pq,q} = B_{pq,q}$), so that we have $[b_p, b_q] \in B'_{pq,q} \subset B_{pq,q}$, and thus (13.151.3) is proven.

[1042] *Proof.* Let w and w' be two elements of $\mathfrak{G}$ satisfying $w < w'$. Then, $\{g \in \mathfrak{G} \mid g < w\}$ is a proper subset of $\{g \in \mathfrak{G} \mid g < w'\}$ (proper because w belongs to the latter set but not to the former set). Hence, $|\{g \in \mathfrak{G} \mid g < w\}| < |\{g \in \mathfrak{G} \mid g < w'\}|$. Since $\rho(w) = |\{g \in \mathfrak{G} \mid g < w\}|$ and $\rho(w') = |\{g \in \mathfrak{G} \mid g < w'\}|$ (similarly), this rewrites as $\rho(w) < \rho(w')$, qed.

[1043] *Proof.* Let $(p,q) \in \mathfrak{L} \times \mathfrak{L}$ be such that $p < q$ and $\ell(pq) = N$. Then, $N = \ell(pq) = \underbrace{\ell(p)}_{\geq 0} + \ell(q) \geq \ell(q)$, so that $\ell(q) \leq N$ and thus $q \in \bigcup_{i=0}^{N} \mathfrak{A}^i = \mathfrak{G}$, qed.

[1044] Of course, this will be an induction within our current induction step, so the reader should try not to confuse the two inductions going on.

$[b_p, b_q] \in B_{pq,q}$ [1045]. Hence, for the rest of the proof of $[b_p, b_q] \in B_{pq,q}$, we WLOG assume that we don't have $\operatorname{stf}(pq) = (p, q)$. Thus, q is not the longest Lyndon proper suffix of pq [1046]. In other words, the statement that q is the longest Lyndon proper suffix of pq is **false**.

We have $\ell(p) > 1$ [1047]. Thus, p is a Lyndon word of length > 1 (since p is Lyndon and since p has length $\ell(p) > 1$). Therefore, $\operatorname{stf} p$ is well-defined. Set $(u, v) = \operatorname{stf} p$. Then, Exercise 6.1.39(a) (applied to $w = p$, $g = u$ and $h = v$) says that v is the longest Lyndon proper suffix of p. In particular, v is a Lyndon proper suffix of p. Moreover, Exercise 6.1.39(b) (applied to $w = p$, $g = u$ and $h = v$) says that we have $p = uv$. Also, Exercise 6.1.39(c) (applied to $w = p$, $g = u$ and $h = v$) says that we have $u < uv < v$. Finally, Exercise 6.1.39(d) (applied to $w = p$, $g = u$ and $h = v$) says that the word u is Lyndon. In other words, $u \in \mathfrak{L}$. The words u and v are nonempty (since they are Lyndon).

Exercise 6.1.36 (applied to $w = p$ and $t = q$) now shows that q is the longest Lyndon proper suffix of pq if and only if we do not have $v < q$. Thus, the statement that we do not have $v < q$ is **false** (because the statement that q is the longest Lyndon proper suffix of pq is **false**). Thus, we have $v < q$. Thus, $u < v < q$.

Recall that $(u, v) = \operatorname{stf} p$. Thus, (6.1.2) (applied to $w = p$) shows that $b_p = [b_u, b_v]$. Thus,

$$\left[\underbrace{b_p}_{=[b_u,b_v]}, b_q\right] = [[b_u, b_v], b_q] = [[b_u, b_q], b_v] - [[b_v, b_q], b_u] \tag{13.151.8}$$

(by (13.151.1), applied to $\mathfrak{k} = \mathfrak{g}$, $x = b_u$, $y = b_v$ and $z = b_q$).

Now,

$$[b_u, b_q] \in B_{uq,q} \tag{13.151.9}$$

[1048] and

$$[b_v, b_q] \in B_{vq,q} \tag{13.151.10}$$

[1045] *Proof.* Assume that $\operatorname{stf}(pq) = (p, q)$. Thus, (6.1.2) (applied to $w = pq$, $u = p$ and $v = q$) shows that $b_{pq} = [b_p, b_q]$. But pq is an element of $\mathfrak{L}$ and satisfies $pq \sim pq$ (since the relation $\sim$ is reflexive) and $pq < q$. In other words, $pq \in \{w \in \mathfrak{L} \mid w \sim pq \text{ and } w < q\} = \mathfrak{L}_{pq,q}$. Thus, $\mathbf{k} b_{pq} \subset \sum_{w \in \mathfrak{L}_{pq,q}} \mathbf{k} b_w$.

Now, from $b_{pq} = [b_p, b_q]$, we obtain

$$[b_p, b_q] = b_{pq} \in \mathbf{k} b_{pq} \subset \sum_{w \in \mathfrak{L}_{pq,q}} \mathbf{k} b_w = B_{pq,q}$$

(since $B_{pq,q} = \sum_{w \in \mathfrak{L}_{pq,q}} \mathbf{k} b_w$), qed.

[1046] *Proof.* Assume the contrary. Thus, q is the longest Lyndon proper suffix of pq.

Let $w = pq$. Recall that q is the longest Lyndon proper suffix of pq. In other words, q is the longest Lyndon proper suffix of w (since $w = pq$).

Let $(g, h) = \operatorname{stf}(pq)$. Then, $(g, h) = \operatorname{stf}\left(\underbrace{pq}_{=w}\right) = \operatorname{stf} w$. Thus, h is the longest Lyndon proper suffix of w (by Exercise 6.1.39(a)). Comparing this with the fact that q is the longest Lyndon proper suffix of w, we obtain that $h = q$.

But Exercise 6.1.39(b) shows that $w = g \underbrace{h}_{=q} = gq$. Thus, $gq = w = pq$. Cancelling q from this equality, we obtain $g = p$. Now, from $(g, h) = \operatorname{stf}(pq)$, we obtain $\operatorname{stf}(pq) = \left(\underbrace{g}_{=p}, \underbrace{h}_{=q}\right) = (p, q)$. This contradicts the fact that we don't have $\operatorname{stf}(pq) = (p, q)$. This contradiction proves that our assumption was wrong, qed.

[1047] *Proof.* Assume the contrary. Thus, $\ell(p) \leq 1$. Since p is nonempty, this shows that $\ell(p) = 1$. Therefore, q is the longest proper suffix of pq. Thus, q is the longest Lyndon proper suffix of pq (since q is Lyndon). This contradicts the fact that q is not the longest Lyndon proper suffix of pq. This contradiction shows that our assumption was wrong, qed.

[1048] *Proof of (13.151.9):* We have $(u, q) \in \mathfrak{L} \times \mathfrak{L}$ (since $u \in \mathfrak{L}$ and $q \in \mathfrak{L}$) and $u < q$. Also, $\ell\left(\underbrace{p}_{=uv}\right) = \ell(uv) = \ell(u) + \underbrace{\ell(v)}_{\substack{>0 \\ (\text{since } v \text{ is nonempty})}} > \ell(u)$. But $\ell(pq) = N$, so that $N = \ell(pq) = \underbrace{\ell(p)}_{>\ell(u)} + \ell(q) > \ell(u) + \ell(q) = \ell(uq)$. Thus, $\ell(uq) < N$. Therefore, (13.151.4) (applied to (u, q) instead of (p, q)) yields $[b_u, b_q] \in B_{uq,q}$. This proves (13.151.9).

[1049]. Hence, (13.151.8) becomes

$$[b_p, b_q] = \left[\underbrace{[b_u, b_q]}_{\substack{\in B_{uq,q} \\ \text{(by (13.151.9))}}}, b_v \right] - \left[\underbrace{[b_v, b_q]}_{\substack{\in B_{vq,q} \\ \text{(by (13.151.10))}}}, b_u \right] \in [B_{uq,q}, b_v] - [B_{vq,q}, b_u] . \tag{13.151.11}$$

On the other hand, if g and h are two Lyndon words satisfying $g < q$, $h < q$ and $gh \sim p$, then

$$(\text{every } r \in \mathfrak{L}_{gq,q} \text{ satisfies } [b_r, b_h] \in B_{pq,q}) \tag{13.151.12}$$

[1050] and therefore

$$[B_{gq,q}, b_h] \subset B_{pq,q} \tag{13.151.13}$$

[1049] *Proof of (13.151.10):* We have $(v, q) \in \mathfrak{L} \times \mathfrak{L}$ (since $v \in \mathfrak{L}$ and $q \in \mathfrak{L}$) and $v < q$. Also, $\ell \left(\underbrace{p}_{=uv} \right) = \ell (uv) = \underbrace{\ell (u)}_{\substack{>0 \\ \text{(since } u \text{ is nonempty)}}} + \ell (v) > \ell (v)$. But $\ell (pq) = N$, so that $N = \ell (pq) = \underbrace{\ell (p)}_{> \ell(v)} + \ell (q) > \ell (v) + \ell (q) = \ell (vq)$. Thus, $\ell (vq) < N$. Therefore, (13.151.4) (applied to (v, q) instead of (p, q)) yields $[b_v, b_q] \in B_{vq,q}$. This proves (13.151.10).

[1050] *Proof of (13.151.12):* Let g and h be two Lyndon words satisfying $g < q$, $h < q$ and $gh \sim p$. Let $r \in \mathfrak{L}_{gq,q}$. We must prove that $[b_r, b_h] \in B_{pq,q}$.

The words g and h are Lyndon. In other words, $g \in \mathfrak{L}$ and $h \in \mathfrak{L}$.

We have $gh \sim p$ and thus $\ell (gh) = \ell (p)$. Hence, $\ell (p) = \ell (gh) = \ell (g) + \ell (h)$.

We have $gh \sim p$ and $q \sim q$ (since the relation $\sim$ is reflexive). Hence, $ghq \sim pq$ (by the monoidality of the relation $\sim$).

We have $N = \ell (pq) = \underbrace{\ell (p)}_{\geq 0} + \ell (q) \geq \ell (q)$, whence $\ell (q) \leq N$ and thus $q \in \bigcup_{i=0}^{N} \mathfrak{A}^i = \mathfrak{G}$.

We have $r \in \mathfrak{L}_{gq,q} = \{ w \in \mathfrak{L} \mid w \sim gq \text{ and } w < q \}$ (by the definition of $\mathfrak{L}_{gq,q}$). In other words, r is an element of $\mathfrak{L}$ and satisfies $r \sim gq$ and $r < q$. From $r \sim gq$, we obtain $\ell (r) = \ell (gq) = \ell (g) + \ell (q)$.

If $r = h$, then $\left[\underbrace{b_r}_{\substack{= b_h \\ \text{(since } r = h)}}, b_h \right] = [b_h, b_h] = 0 \in B_{pq,q}$ (since $B_{pq,q}$ is a $\mathbf{k}$-module). Hence, for the rest of the proof of $[b_r, b_h] \in B_{pq,q}$, we WLOG assume that we don't have $r = h$.

Thus, we have $r \neq h$. Hence, we have either $r < h$ or $r > h$ (since the lexicographic order on $\mathfrak{A}^*$ is a total order). In other words, we are in one of the following two Cases:

Case 1: We have $r < h$.

Case 2: We have $r > h$.

Let us first consider Case 1. In this case, we have $r < h$. We have $\ell (rh) = \underbrace{\ell (r)}_{= \ell(g) + \ell(q)} + \ell (h) = \ell (g) + \ell (q) + \ell (h) = \ell (g) + \ell (h) + \ell (q)$. Comparing this with $\ell (pq) = \underbrace{\ell (p)}_{= \ell(g) + \ell(h)} + \ell (q) = \ell (g) + \ell (h) + \ell (q)$, we obtain $\ell (rh) = \ell (pq) = N$.

We have $r \sim gq$ and $h \sim h$ (since the relation $\sim$ is reflexive). Thus, $rh \sim gqh$ (by the monoidality of the relation $\sim$). But $gqh \sim ghq$ (by the monoidality of the relation $\sim$ again, since $g \sim g$ and $qh \sim hq$). From $rh \sim gqh$ and $gqh \sim ghq$, we obtain $rh \sim ghq$ (since the relation $\sim$ is transitive). From $rh \sim ghq$ and $ghq \sim pq$, we obtain $rh \sim pq$ (since the relation $\sim$ is transitive). Now, (13.151.2) (applied to rh, h, pq and q instead of h, s, g and t) yields $B_{rh,h} \subset B_{pq,q}$ (since $h < q$).

On the other hand, $N = \ell (rh) = \underbrace{\ell (r)}_{\geq 0} + \ell (h) \geq \ell (h)$, whence $\ell (h) \leq N$ and thus $h \in \bigcup_{i=0}^{N} \mathfrak{A}^i = \mathfrak{G}$. Since $h < q$, we have $\rho (h) < \rho (q)$ (by (13.151.6), applied to h and q instead of w and w'), so that $\rho (h) < \rho (q) = K$.

Now, we have $(r, h) \in \mathfrak{L} \times \mathfrak{L}$ (since $r \in \mathfrak{L}$ and $h \in \mathfrak{L}$) and $r < h$ and $\ell (rh) = N$ and $\rho (h) < K$. Therefore, (13.151.7) (applied to r and h instead of p and q) shows that $[b_r, b_h] \in B_{rh,h} \subset B_{pq,q}$. Thus, $[b_r, b_h] \in B_{pq,q}$ is proven in Case 1.

Let us now consider Case 2. In this case, we have $h < r$. We have $\ell (hr) = \ell (h) + \underbrace{\ell (r)}_{= \ell(g) + \ell(q)} = \ell (h) + \ell (g) + \ell (q) = \ell (g) + \ell (h) + \ell (q)$. Comparing this with $\ell (pq) = \underbrace{\ell (p)}_{= \ell(g) + \ell(h)} + \ell (q) = \ell (g) + \ell (h) + \ell (q)$, we obtain $\ell (hr) = \ell (pq) = N$.

We have $h \sim h$ (since the relation $\sim$ is reflexive) and $r \sim gq$. Thus, $hr \sim hgq$ (by the monoidality of the relation $\sim$). But $hgq \sim ghq$ (by the monoidality of the relation $\sim$, since $hg \sim gh$ and $q \sim q$). From $hr \sim hgq$ and $hgq \sim ghq$, we obtain $hr \sim ghq$ (since the relation $\sim$ is transitive). From $hr \sim ghq$ and $ghq \sim pq$, we obtain $hr \sim pq$ (since the relation $\sim$ is transitive). Now, (13.151.2) (applied to hr, r, pq and q instead of h, s, g and t) yields $B_{hr,r} \subset B_{pq,q}$ (since $r < q$).

[1051].

We have $p \sim p$ (since the relation $\sim$ is reflexive). In other words, $uv \sim p$ (since $p = uv$). Also, $vu \sim uv$. Since $p = uv$, this rewrites as $vu \sim p$.

Now, the words u and v are Lyndon and satisfy $u < q$, $v < q$ and $uv \sim p$. Thus, we can apply (13.151.13) to $g = u$ and $h = v$. As a result, we obtain $[B_{uq,q}, b_v] \subset B_{pq,q}$.

Furthermore, the words v and u are Lyndon and satisfy $v < q$, $u < q$ and $vu \sim p$. Thus, we can apply (13.151.13) to $g = v$ and $h = u$. As a result, we obtain $[B_{vq,q}, b_u] \subset B_{pq,q}$.

Now, (13.151.11) becomes

$$[b_p, b_q] \in \underbrace{[B_{uq,q}, b_v]}_{\subset B_{pq,q}} - \underbrace{[B_{vq,q}, b_u]}_{\subset B_{pq,q}} \subset B_{pq,q} - B_{pq,q} \subset B_{pq,q}$$

(since $B_{pq,q}$ is a $\mathbf{k}$-module).

Let us now forget that we fixed (p, q). We thus have proven that

$$[b_p, b_q] \in B_{pq,q} \qquad \text{for every } (p, q) \in \mathfrak{L} \times \mathfrak{L} \text{ satisfying } p < q \text{ and } \ell(pq) = N \text{ and } \rho(q) = K.$$

In other words, (13.151.5) holds in the case when $\rho(q) = K$. This completes the induction step (in the induction proof of (13.151.5)). The induction proof of (13.151.5) is thus finished.

We thus have proven (13.151.5). In other words, (13.151.3) holds in the case when $\ell(pq) = N$. This completes the induction step (in the induction proof of (13.151.3)). The induction proof of (13.151.3) is thus complete.

Now, we recall that B is the $\mathbf{k}$-submodule of $\mathfrak{g}$ spanned by the family $(b_w)_{w \in \mathfrak{L}}$. In other words,

$$B = \sum_{w \in \mathfrak{L}} \mathbf{k} b_w. \tag{13.151.14}$$

Thus,

$$B_{h,s} \subset B \qquad \text{for every } h \in \mathfrak{A}^* \text{ and } s \in \mathfrak{A}^* \tag{13.151.15}$$

[1052].

Now, using (13.151.3), we can easily see the following fact: For any $p \in \mathfrak{L}$ and $q \in \mathfrak{L}$, we have

$$[b_p, b_q] \in B \tag{13.151.16}$$

On the other hand, $N = \ell(hr) = \underbrace{\ell(h)}_{\geq 0} + \ell(r) \geq \ell(r)$, whence $\ell(r) \leq N$ and thus $r \in \bigcup_{i=0}^{N} \mathfrak{A}^i = \mathfrak{G}$. Since $r < q$, we have $\rho(r) < \rho(q)$ (by (13.151.6), applied to r and q instead of w and w'), so that $\rho(r) < \rho(q) = K$.

Now, we have $(h, r) \in \mathfrak{L} \times \mathfrak{L}$ (since $h \in \mathfrak{L}$ and $r \in \mathfrak{L}$) and $h < r$ and $\ell(hr) = N$ and $\rho(r) < K$. Therefore, (13.151.7) (applied to h and r instead of p and q) shows that $[b_h, b_r] \in B_{hr,r} \subset B_{pq,q}$. Thus, $[b_r, b_h] = -\underbrace{[b_h, b_r]}_{\in B_{pq,q}} \in -B_{pq,q} \subset B_{pq,q}$ (since $B_{pq,q}$ is a $\mathbf{k}$-module). Hence, $[b_r, b_h] \in B_{pq,q}$ is proven in Case 2.

Thus, $[b_r, b_h] \in B_{pq,q}$ is proven in each of the two Cases 1 and 2. This completes the proof of (13.151.12).

[1051] *Proof of (13.151.13):* Let g and h be two Lyndon words satisfying $g < q$, $h < q$ and $gh \sim p$. The definition of $B_{gq,q}$ shows that $B_{gq,q} = \sum_{w \in \mathfrak{L}_{gq,q}} \mathbf{k} b_w$. In other words, $B_{gq,q}$ is the $\mathbf{k}$-linear span of the elements b_w with $w \in \mathfrak{L}_{gq,q}$. Hence, in order to prove the relation (13.151.13), it suffices to show that $[b_w, b_h] \in B_{pq,q}$ for every $w \in \mathfrak{L}_{gq,q}$. But this follows from (13.151.12) (applied to $r = w$). This proves (13.151.13).

[1052] *Proof of (13.151.15):* Let $h \in \mathfrak{A}^*$ and $s \in \mathfrak{A}^*$.

Clearly, $\mathfrak{L}_{h,s} \subset \mathfrak{L}$. Thus, $\sum_{w \in \mathfrak{L}_{h,s}} \mathbf{k} b_w \subset \sum_{w \in \mathfrak{L}} \mathbf{k} b_w$. Since $B_{h,s} = \sum_{w \in \mathfrak{L}_{h,s}} \mathbf{k} b_w$ and $B = \sum_{w \in \mathfrak{L}} \mathbf{k} b_w$ (by (13.151.14)), this rewrites as $B_{h,s} \subset B$. This proves (13.151.15).

[1053]. Now, (13.151.14) yields $B = \sum_{w\in\mathfrak{L}} \mathbf{k} b_w = \sum_{p\in\mathfrak{L}} \mathbf{k} b_p$ (here, we renamed the summation index w as p) and $B = \sum_{w\in\mathfrak{L}} \mathbf{k} b_w = \sum_{q\in\mathfrak{L}} \mathbf{k} b_q$ (here, we renamed the summation index w as q). Thus,

$$\begin{aligned}
&\left[\underbrace{B}_{=\sum_{p\in\mathfrak{L}} \mathbf{k} b_p}, \underbrace{B}_{=\sum_{q\in\mathfrak{L}} \mathbf{k} b_q}\right] \\
&= \left[\sum_{p\in\mathfrak{L}} \mathbf{k} b_p, \sum_{q\in\mathfrak{L}} \mathbf{k} b_q\right] = \sum_{p\in\mathfrak{L}} \sum_{q\in\mathfrak{L}} \underbrace{\mathbf{k}\mathbf{k}}_{\subset \mathbf{k}} \underbrace{[b_p, b_q]}_{\substack{\in B \\ \text{(by (13.151.16))}}} \qquad \text{(since the Lie bracket on } \mathfrak{g} \text{ is } \mathbf{k}\text{-bilinear)} \\
&\subset \sum_{p\in\mathfrak{L}} \sum_{q\in\mathfrak{L}} \mathbf{k} B \subset B \qquad \text{(since } B \text{ is a } \mathbf{k}\text{-module)}.
\end{aligned}$$

In other words, B is a Lie subalgebra of $\mathfrak{g}$. This solves Exercise 6.1.40(a).

(b) We shall first prove that

$$f([b_p, b_q]) = [f(b_p), f(b_q)] \qquad \text{for every } (p,q) \in \mathfrak{L}\times\mathfrak{L} \text{ satisfying } p < q. \tag{13.151.17}$$

The proof of (13.151.17) is very similar to our above proof of (13.151.3); it proceeds using the same kind of double induction, with almost the same computations. Here are the details of this proof:

Proof of (13.151.17): We can WLOG assume that the alphabet $\mathfrak{A}$ is finite[1054]. Assume this.

We shall prove (13.151.17) by strong induction over $\ell(pq)$:

Induction step: Let N be a nonnegative integer. Assume that (13.151.17) holds in the case when $\ell(pq) < N$. We need to prove that (13.151.17) holds in the case when $\ell(pq) = N$.

We have assumed that (13.151.17) holds in the case when $\ell(pq) < N$. In other words, we have

$$f([b_p, b_q]) = [f(b_p), f(b_q)] \qquad \text{for every } (p,q) \in \mathfrak{L}\times\mathfrak{L} \text{ satisfying } p < q \text{ and } \ell(pq) < N. \tag{13.151.18}$$

We now must prove that (13.151.17) holds in the case when $\ell(pq) = N$. In other words, we must prove that

$$f([b_p, b_q]) = [f(b_p), f(b_q)] \qquad \text{for every } (p,q) \in \mathfrak{L}\times\mathfrak{L} \text{ satisfying } p < q \text{ and } \ell(pq) = N. \tag{13.151.19}$$

We define a set $\mathfrak{G}$ in the same fashion as in our proof of (13.151.3). Likewise, we define a nonnegative integer $\rho(w)$ for every $w \in \mathfrak{G}$ in the same way as we did in our proof of (13.151.3).

For every $(p,q) \in \mathfrak{L}\times\mathfrak{L}$ satisfying $p < q$ and $\ell(pq) = N$, we have $q \in \mathfrak{G}$ [1055], and thus $\rho(q)$ is well-defined. We are thus going to prove (13.151.19) by strong induction over $\rho(q)$ [1056]:

[1053] *Proof of (13.151.16):* Let $p \in \mathfrak{L}$ and $q \in \mathfrak{L}$. We need to prove that $[b_p, b_q] \in B$. If $p = q$, then this is clear (because if $p = q$, then $\left[\underbrace{b_p}_{\substack{=b_q \\ (\text{since } p=q)}}, b_q\right] = [b_q, b_q] = 0 \in B$ (since B is a $\mathbf{k}$-module)). Thus, for the rest of this proof, we WLOG assume that we don't have $p = q$.

We have $(p,q) \in \mathfrak{L}\times\mathfrak{L}$ (since $p \in \mathfrak{L}$ and $q \in \mathfrak{L}$) and $(q,p) \in \mathfrak{L}\times\mathfrak{L}$ (since $q \in \mathfrak{L}$ and $p \in \mathfrak{L}$). We have $p \neq q$ (since we don't have $p = q$). Thus, we have either $p < q$ or $p > q$ (since the lexicographic order on $\mathfrak{A}^*$ is a total order). In other words, we are in one of the following two Cases:

Case 1: We have $p < q$.

Case 2: We have $p > q$.

Let us first consider Case 1. In this case, we have $p < q$. Thus, from (13.151.3), we obtain $[b_p, b_q] \in B_{pq,q} \subset B$ (by (13.151.15) (applied to $h = pq$ and $s = q$)). Thus, $[b_p, b_q] \in B$ is proven in Case 1.

Let us now consider Case 2. In this case, we have $p > q$. In other words, $q < p$. Thus, (13.151.3) (applied to (q,p) instead of (p,q)) shows that $[b_q, b_p] \in B_{qp,p} \subset B$ (by (13.151.15) (applied to $h = qp$ and $s = p$)). Now, $[b_p, b_q] = -\underbrace{[b_q, b_p]}_{\in B} \in -B \subset B$ (since B is a $\mathbf{k}$-module). Thus, $[b_p, b_q] \in B$ is proven in Case 2.

Now, $[b_p, b_q] \in B$ is proven in each of the two Cases 1 and 2. Thus, (13.151.16) is proven.

[1054] The reasons why this is legitimate are similar to the analogous reasons in the proof of (13.151.3).

[1055] This can be proven in the same way as in our proof of (13.151.3).

[1056] Of course, this will be an induction within our current induction step, so the reader should try not to confuse the two inductions going on.

Induction step: Let K be a nonnegative integer. Assume that (13.151.19) holds in the case when $\rho(q) < K$. We need to prove that (13.151.19) holds in the case when $\rho(q) = K$.

We have assumed that (13.151.19) holds in the case when $\rho(q) < K$. In other words, we have

$$f([b_p, b_q]) = [f(b_p), f(b_q)] \qquad \text{for every } (p,q) \in \mathfrak{L} \times \mathfrak{L} \text{ satisfying } p < q \text{ and } \ell(pq) = N \text{ and } \rho(q) < K. \tag{13.151.20}$$

Now, let $(p,q) \in \mathfrak{L} \times \mathfrak{L}$ be such that $p < q$ and $\ell(pq) = N$ and $\rho(q) = K$. We are going to prove that $f([b_p, b_q]) = [f(b_p), f(b_q)]$.

As in our proof of (13.151.5), we can prove the following facts:

- The words p and q are Lyndon.
- The word pq is Lyndon. In other words, $pq \in \mathfrak{L}$.
- We have $pq < q$.
- The word pq is a Lyndon word of length > 1. Therefore, $\operatorname{stf}(pq)$ is well-defined.

If $\operatorname{stf}(pq) = (p,q)$, then it is easy to see that $f([b_p, b_q]) = [f(b_p), f(b_q)]$ [1057]. Hence, for the rest of the proof of $f([b_p, b_q]) = [f(b_p), f(b_q)]$, we WLOG assume that we don't have $\operatorname{stf}(pq) = (p,q)$. As in our proof of (13.151.5), we can see that p is a Lyndon word of length > 1. Therefore, $\operatorname{stf} p$ is well-defined. Set $(u,v) = \operatorname{stf} p$.

As in our proof of (13.151.5), we can see the following facts:

- The word v is a Lyndon proper suffix of p.
- We have $p = uv$.
- We have $u < uv < v$.
- The word u is Lyndon. In other words, $u \in \mathfrak{L}$.
- The words u and v are nonempty.
- We have $u < v < q$.
- We have $b_p = [b_u, b_v]$.
- The equality (13.151.8) holds.

On the other hand, (6.1.3) (applied to $w = p$) shows that $f([b_u, b_v]) = [f(b_u), f(b_v)]$ (since p is a Lyndon word of length > 1, and since $(u,v) = \operatorname{stf} p$). Now, applying the map f to both sides of the equality $b_p = [b_u, b_v]$, we obtain

$$f(b_p) = f([b_u, b_v]) = [f(b_u), f(b_v)]. \tag{13.151.21}$$

Now,

$$f([b_u, b_q]) = [f(b_u), f(b_q)] \tag{13.151.22}$$

[1058] and

$$f([b_v, b_q]) = [f(b_v), f(b_q)] \tag{13.151.23}$$

[1059].

On the other hand, if g and h are two Lyndon words satisfying $g < q$, $h < q$ and $gh \sim p$, then

$$(\text{every } r \in \mathfrak{L}_{gq,q} \text{ satisfies } f([b_r, b_h]) = [f(b_r), f(b_h)]) \tag{13.151.24}$$

[1057] *Proof.* Assume that $\operatorname{stf}(pq) = (p,q)$. Thus, $(p,q) = \operatorname{stf}(pq)$. Hence, (6.1.3) (applied to $w = pq$, $u = p$ and $v = q$) shows that $f([b_p, b_q]) = [f(b_p), f(b_q)]$, qed.

[1058] *Proof of (13.151.22):* We have $(u,q) \in \mathfrak{L} \times \mathfrak{L}$ (since $u \in \mathfrak{L}$ and $q \in \mathfrak{L}$) and $u < q$. Also, $\ell\left(\underbrace{p}_{=uv}\right) = \ell(uv) = \ell(u) + \underbrace{\ell(v)}_{\substack{>0 \\ (\text{since } v \text{ is nonempty})}} > \ell(u)$. But $\ell(pq) = N$, so that $N = \ell(pq) = \underbrace{\ell(p)}_{>\ell(u)} + \ell(q) > \ell(u) + \ell(q) = \ell(uq)$. Thus, $\ell(uq) < N$. Therefore, (13.151.18) (applied to (u,q) instead of (p,q)) yields $f([b_u, b_q]) = [f(b_u), f(b_q)]$. This proves (13.151.22).

[1059] *Proof of (13.151.23):* We have $(v,q) \in \mathfrak{L} \times \mathfrak{L}$ (since $v \in \mathfrak{L}$ and $q \in \mathfrak{L}$) and $v < q$. Also, $\ell\left(\underbrace{p}_{=uv}\right) = \ell(uv) = \underbrace{\ell(u)}_{\substack{>0 \\ (\text{since } u \text{ is nonempty})}} + \ell(v) > \ell(v)$. But $\ell(pq) = N$, so that $N = \ell(pq) = \underbrace{\ell(p)}_{>\ell(v)} + \ell(q) > \ell(v) + \ell(q) = \ell(vq)$. Thus, $\ell(vq) < N$. Therefore, (13.151.18) (applied to (v,q) instead of (p,q)) yields $f([b_v, b_q]) = [f(b_v), f(b_q)]$. This proves (13.151.23).

[1060] and therefore

$$\text{(every } x \in B_{gq,q} \text{ satisfies } f([x, b_h]) = [f(x), f(b_h)]) \tag{13.151.25}$$

[1061].

Recall that (13.151.3) holds. (This was proven in our solution to Exercise 6.1.40(a).) Now, $(u, q) \in \mathfrak{L} \times \mathfrak{L}$ (since $u \in \mathfrak{L}$ and $q \in \mathfrak{L}$) and $u < q$. Thus, (13.151.3) (applied to (u, q) instead of (p, q)) shows that $[b_u, b_q] \in B_{uq,q}$. Also, $(v, q) \in \mathfrak{L} \times \mathfrak{L}$ (since $v \in \mathfrak{L}$ and $q \in \mathfrak{L}$) and $v < q$. Thus, (13.151.3) (applied to (v, q) instead of (p, q)) shows that $[b_v, b_q] \in B_{vq,q}$.

As in our proof of (13.151.5), we can show that $uv \sim p$ and $vu \sim p$.

Now, the words u and v are Lyndon and satisfy $u < q$, $v < q$, $uv \sim p$ and $[b_u, b_q] \in B_{uq,q}$. Thus, we can apply (13.151.25) to $g = u$, $h = v$ and $x = [b_u, b_q]$. As a result, we obtain

$$f([[b_u, b_q], b_v]) = \left[\underbrace{f([b_u, b_q])}_{\substack{=[f(b_u), f(b_q)] \\ \text{(by (13.151.22))}}}, f(b_v) \right] = [[f(b_u), f(b_q)], f(b_v)]. \tag{13.151.26}$$

Furthermore, the words v and u are Lyndon and satisfy $v < q$, $u < q$, $vu \sim p$ and $[b_v, b_q] \in B_{vq,q}$. Thus, we can apply (13.151.25) to $g = v$, $h = u$ and $x = [b_v, b_q]$. As a result, we obtain

$$f([[b_v, b_q], b_u]) = \left[\underbrace{f([b_v, b_q])}_{\substack{=[f(b_v), f(b_q)] \\ \text{(by (13.151.23))}}}, f(b_u) \right] = [[f(b_v), f(b_q)], f(b_u)]. \tag{13.151.27}$$

Now, applying the map f to both sides of the equality (13.151.8), we obtain

$$\begin{aligned} f([b_p, b_q]) &= f([[b_u, b_q], b_v] - [[b_v, b_q], b_u]) \\ &= \underbrace{f([[b_u, b_q], b_v])}_{\substack{=[[f(b_u), f(b_q)], f(b_v)] \\ \text{(by (13.151.26))}}} - \underbrace{f([[b_v, b_q], b_u])}_{\substack{=[[f(b_v), f(b_q)], f(b_u)] \\ \text{(by (13.151.27))}}} \qquad \text{(since the map } f \text{ is } \mathbf{k}\text{-linear)} \\ &= [[f(b_u), f(b_q)], f(b_v)] - [[f(b_v), f(b_q)], f(b_u)]. \end{aligned}$$

Comparing this with

$$\begin{aligned} \left[\underbrace{f(b_p)}_{\substack{=[f(b_u), f(b_v)] \\ \text{(by (13.151.21))}}}, f(b_q) \right] &= [[f(b_u), f(b_v)], f(b_q)] \\ &= [[f(b_u), f(b_q)], f(b_v)] - [[f(b_v), f(b_q)], f(b_u)]. \\ &\qquad \text{(by (13.151.1), applied to } \mathfrak{k} = \mathfrak{h},\ x = f(b_u),\ y = f(b_v) \text{ and } z = f(b_q)), \end{aligned}$$

we obtain $f([b_p, b_q]) = [f(b_p), f(b_q)]$.

Let us now forget that we fixed (p, q). We thus have proven that

$$f([b_p, b_q]) = [f(b_p), f(b_q)] \qquad \text{for every } (p, q) \in \mathfrak{L} \times \mathfrak{L} \text{ satisfying } p < q \text{ and } \ell(pq) = N \text{ and } \rho(q) = K.$$

In other words, (13.151.19) holds in the case when $\rho(q) = K$. This completes the induction step (in the induction proof of (13.151.19)). The induction proof of (13.151.19) is thus finished.

[1060] The proof of (13.151.24) can be obtained by modifying our proof of (13.151.12) in a straightforward way. (We now must use (13.151.20) instead of (13.151.7).)

[1061] *Proof of (13.151.25):* Let g and h be two Lyndon words satisfying $g < q$, $h < q$ and $gh \sim p$. The definition of $B_{gq,q}$ shows that $B_{gq,q} = \sum_{w \in \mathfrak{L}_{gq,q}} \mathbf{k} b_w$. In other words, $B_{gq,q}$ is the $\mathbf{k}$-linear span of the elements b_w with $w \in \mathfrak{L}_{gq,q}$. Hence, in order to prove the relation (13.151.25), it suffices to show that $f([b_w, b_h]) = [f(b_w), f(b_h)]$ for every $w \in \mathfrak{L}_{gq,q}$. But this follows from (13.151.24) (applied to $r = w$). This proves (13.151.25).

We thus have proven (13.151.19). In other words, (13.151.17) holds in the case when $\ell(pq) = N$. This completes the induction step (in the induction proof of (13.151.17)). The induction proof of (13.151.17) is thus complete.

Using (13.151.17), we can easily see the following fact: For any $p \in \mathfrak{L}$ and $q \in \mathfrak{L}$, we have

$$f([b_p, b_q]) = [f(b_p), f(b_q)] \tag{13.151.28}$$

[1062].

Now, we recall that B is the $\mathbf{k}$-submodule of $\mathfrak{g}$ spanned by the family $(b_w)_{w \in \mathfrak{L}}$. Thus, the family $(b_w)_{w \in \mathfrak{L}}$ spans the $\mathbf{k}$-module B.

Now, we have

$$f([x, y]) = [f(x), f(y)] \qquad \text{for every } x \in B \text{ and } y \in B \tag{13.151.29}$$

[1063]. In other words, the map $f : B \to \mathfrak{h}$ is a Lie algebra homomorphism. This solves Exercise 6.1.40(b).

13.152. **Solution to Exercise 6.1.41.** *Solution to Exercise 6.1.41.* The definition of $\mathfrak{g}$ shows that $\mathfrak{g} = \mathfrak{g}_1 + \mathfrak{g}_2 + \mathfrak{g}_3 + \cdots = \sum_{i \geq 1} \mathfrak{g}_i$. Thus, for every positive integer k, we have

$$\mathfrak{g}_k \subset \mathfrak{g}. \tag{13.152.1}$$

Notice that $\mathfrak{g}_1 = V$, so that

$$V = \mathfrak{g}_1 \subset \mathfrak{g} \qquad (\text{by (13.152.1), applied to } k = 1). \tag{13.152.2}$$

We also recall a fundamental property of Lie algebras (one of the forms of the Jacobi identity):

- Every three elements x, y and z of a Lie algebra $\mathfrak{k}$ satisfy

$$[[x, y], z] = [[x, z], y] - [[y, z], x]. \tag{13.152.3}$$

[1062] *Proof of (13.151.28):* Let $p \in \mathfrak{L}$ and $q \in \mathfrak{L}$. We need to prove that $f([b_p, b_q]) = [f(b_p), f(b_q)]$. If $p = q$, then this is clear (because if $p = q$, then $f\left(\left[\underbrace{b_p}_{\substack{=b_q \\ (\text{since } p=q)}}, b_q\right]\right) = f\left(\underbrace{[b_q, b_q]}_{=0}\right) = f(0) = 0$ and $\left[f\left(\underbrace{b_p}_{\substack{=b_q \\ (\text{since } p=q)}}\right), f(b_q)\right] = [f(b_q), f(b_q)] = 0$, so that both sides of the equality $f([b_p, b_q]) = [f(b_p), f(b_q)]$ vanish). Thus, for the rest of this proof, we WLOG assume that we don't have $p = q$.

We have $(p, q) \in \mathfrak{L} \times \mathfrak{L}$ (since $p \in \mathfrak{L}$ and $q \in \mathfrak{L}$) and $(q, p) \in \mathfrak{L} \times \mathfrak{L}$ (since $q \in \mathfrak{L}$ and $p \in \mathfrak{L}$). We have $p \neq q$ (since we don't have $p = q$). Thus, we have either $p < q$ or $p > q$ (since the lexicographic order on $\mathfrak{A}^*$ is a total order). In other words, we are in one of the following two Cases:

Case 1: We have $p < q$.

Case 2: We have $p > q$.

Let us first consider Case 1. In this case, we have $p < q$. Thus, from (13.151.17), we obtain $f([b_p, b_q]) = [f(b_p), f(b_q)]$. Thus, $f([b_p, b_q]) = [f(b_p), f(b_q)]$ is proven in Case 1.

Let us now consider Case 2. In this case, we have $p > q$. In other words, $q < p$. Thus, (13.151.17) (applied to (q, p) instead of (p, q)) shows that $f([b_q, b_p]) = [f(b_q), f(b_p)]$. Now,

$$f\left(\underbrace{[b_p, b_q]}_{=-[b_q, b_p]}\right) = f(-[b_q, b_p]) = -\underbrace{f([b_q, b_p])}_{=[f(b_q), f(b_p)]} = -[f(b_q), f(b_p)] = [f(b_p), f(b_q)].$$

Thus, $f([b_p, b_q]) = [f(b_p), f(b_q)]$ is proven in Case 2.

Now, $f([b_p, b_q]) = [f(b_p), f(b_q)]$ is proven in each of the two Cases 1 and 2. This completes the proof of (13.151.28).

[1063] *Proof of (13.151.29):* Let $x \in B$ and $y \in B$. We must prove the equality $f([x, y]) = [f(x), f(y)]$. Both sides of this equality are $\mathbf{k}$-linear in each of x and y. Hence, we can WLOG assume that both x and y belong to the family $(b_w)_{w \in \mathfrak{L}}$ (since the family $(b_w)_{w \in \mathfrak{L}}$ spans the $\mathbf{k}$-module B). In other words, we can WLOG assume that there exist $p \in \mathfrak{L}$ and $q \in \mathfrak{L}$ satisfying $x = b_p$ and $y = b_q$. Assume this, and consider these p and q.

From (13.151.28), we obtain $f([b_p, b_q]) = [f(b_p), f(b_q)]$. This rewrites as $f([x, y]) = [f(x), f(y)]$ (since $x = b_p$ and $y = b_q$). This proves (13.151.29).

(a) For every two positive integers i and j, we have

$$[\mathfrak{g}_i, \mathfrak{g}_j] \subset \mathfrak{g}_{i+j}. \tag{13.152.4}$$

Proof of (13.152.4): We shall prove (13.152.4) by induction over i:

Induction base: For every positive integer j, we have $[\mathfrak{g}_1, \mathfrak{g}_j] \subset \mathfrak{g}_{1+j}$ [1064]. In other words, (13.152.4) holds for $i = 1$. This completes the induction base.

Induction step: Let I be a positive integer. Assume that (13.152.4) holds for $i = I$. We must prove that (13.152.4) holds for $i = I + 1$.

We have assumed that (13.152.4) holds for $i = I$. In other words, for every positive integer j, we have

$$[\mathfrak{g}_I, \mathfrak{g}_j] \subset \mathfrak{g}_{I+j}. \tag{13.152.5}$$

Now, let j be a positive integer. Thus, the recursive definition of $\mathfrak{g}_{j+1}$ yields $\mathfrak{g}_{j+1} = \left[V, \underbrace{\mathfrak{g}_{(j+1)-1}}_{=\mathfrak{g}_j}\right] = [V, \mathfrak{g}_j]$. The same argument (applied to I instead of j) shows that $\mathfrak{g}_{I+1} = [V, \mathfrak{g}_I]$.

Also, the recursive definition of $\mathfrak{g}_{I+j+1}$ yields $\mathfrak{g}_{I+j+1} = \left[V, \underbrace{\mathfrak{g}_{(I+j+1)-1}}_{=\mathfrak{g}_{I+j}}\right] = [V, \mathfrak{g}_{I+j}]$.

Now, let $z \in \mathfrak{g}_j$. Then, $z \in \mathfrak{g}_j \subset \mathfrak{g}$ (by (13.152.1), applied to $k = j$).

Also, let $p \in \mathfrak{g}_{I+1}$. We are going to prove that $[p, z] \in \mathfrak{g}_{I+1+j}$.

Recall that $[V, \mathfrak{g}_I]$ is the $\mathbf{k}$-linear span of all elements of the form $[x, y]$ with $(x, y) \in V \times \mathfrak{g}_I$ (indeed, this is how $[V, \mathfrak{g}_I]$ is defined). Thus, p is a $\mathbf{k}$-linear combination of elements of the form $[x, y]$ with $(x, y) \in V \times \mathfrak{g}_I$ (since $p \in \mathfrak{g}_{I+1} = [V, \mathfrak{g}_I]$).

Now, we must prove the relation $[p, z] \in \mathfrak{g}_{I+1+j}$. But this relation is $\mathbf{k}$-linear in p. Thus, we WLOG assume that p is an element of the form $[x, y]$ for with $(x, y) \in V \times \mathfrak{g}_I$ (since p is a $\mathbf{k}$-linear combination of elements of the form $[x, y]$ with $(x, y) \in V \times \mathfrak{g}_I$). In other words, there exists an $(x, y) \in V \times \mathfrak{g}_I$ such that $p = [x, y]$. Consider this (x, y).

[1064] *Proof.* Let j be a positive integer. Thus, the recursive definition of $\mathfrak{g}_{1+j}$ yields $\mathfrak{g}_{1+j} = \left[V, \underbrace{\mathfrak{g}_{(1+j)-1}}_{=\mathfrak{g}_j}\right] = [V, \mathfrak{g}_j]$, so that $[V, \mathfrak{g}_j] = \mathfrak{g}_{1+j}$. Now, $\left[\underbrace{\mathfrak{g}_1}_{=V}, \mathfrak{g}_j\right] = [V, \mathfrak{g}_j] = \mathfrak{g}_{1+j} \subset \mathfrak{g}_{1+j}$, qed.

We have $(x, y) \in V \times \mathfrak{g}_I$. In other words, $x \in V$ and $y \in \mathfrak{g}_I$. Thus, $y \in \mathfrak{g}_I \subset \mathfrak{g}$ (by (13.152.1), applied to $k = I$). Also, $x \in V \subset \mathfrak{g}$ (by (13.152.2)). Now,

$$\begin{aligned} \Bigg[\underbrace{p}_{=[x,y]}, z\Bigg] &= [[x,y],z] = \underbrace{[[x,z],y]}_{=-[y,[x,z]]} - \underbrace{[[y,z],x]}_{=-[x,[y,z]]} \qquad (\text{by } (13.152.3)) \\ &= -[y,[x,z]] - (-[x,[y,z]]) = \Bigg[\underbrace{x}_{\in V}, \Bigg[\underbrace{y}_{\in \mathfrak{g}_I}, \underbrace{z}_{\in\mathfrak{g}_j}\Bigg]\Bigg] - \Bigg[\underbrace{y}_{\in\mathfrak{g}_I}, \Bigg[\underbrace{x}_{\in V}, \underbrace{z}_{\in \mathfrak{g}_j}\Bigg]\Bigg] \\ &\in \Bigg[V, \underbrace{[\mathfrak{g}_I,\mathfrak{g}_j]}_{\substack{\subset \mathfrak{g}_{I+j} \\ (\text{by } (13.152.5))}}\Bigg] - \Bigg[\mathfrak{g}_I, \underbrace{[V,\mathfrak{g}_j]}_{\substack{=\mathfrak{g}_{j+1}\\(\text{since } \mathfrak{g}_{j+1}=[V,\mathfrak{g}_j])}}\Bigg] \\ &\subset \underbrace{[V,\mathfrak{g}_{I+j}]}_{\substack{=\mathfrak{g}_{I+j+1}\\ (\text{since } \mathfrak{g}_{I+j+1}=[V,\mathfrak{g}_{I+j}])}} - \underbrace{[\mathfrak{g}_I,\mathfrak{g}_{j+1}]}_{\substack{\subset \mathfrak{g}_{I+j+1}\\ (\text{by } (13.152.5)\text{, applied to } j+1 \text{ instead of } j)}} \\ &\subset \mathfrak{g}_{I+j+1} - \mathfrak{g}_{I+j+1} \subset \mathfrak{g}_{I+j+1} \qquad (\text{since } \mathfrak{g}_{I+j+1} \text{ is a } \mathbf{k}\text{-module}) \\ &= \mathfrak{g}_{I+1+j}. \end{aligned}$$

Now, let us forget that we fixed z and p. We thus have proven that $[p, z] \in \mathfrak{g}_{I+1+j}$ for any $p \in \mathfrak{g}_{I+1}$ and $z \in \mathfrak{g}_j$. Thus, $[\mathfrak{g}_{I+1}, \mathfrak{g}_j] \subset \mathfrak{g}_{I+1+j}$ (since $\mathfrak{g}_{I+1+j}$ is a $\mathbf{k}$-module).

Let us now forget that we fixed j. We thus have proven that, for every positive integer j, we have $[\mathfrak{g}_{I+1}, \mathfrak{g}_j] \subset \mathfrak{g}_{I+1+j}$. In other words, (13.152.4) holds for $i = I + 1$. This completes the induction step. The induction proof of (13.152.4) is thus complete.

Now, for every two positive integers i and j, we have

$$\begin{aligned} [\mathfrak{g}_i, \mathfrak{g}_j] &\subset \mathfrak{g}_{i+j} \qquad (\text{by } (13.152.4)) \\ &\subset \mathfrak{g} \qquad (\text{by } (13.152.1)\text{, applied to } k = i + j)\,. \end{aligned} \tag{13.152.6}$$

But $\mathfrak{g} = \sum_{i \geq 1} \mathfrak{g}_i = \sum_{j \geq 1} \mathfrak{g}_j$ (here, we have renamed the summation index i as j). Thus,

$$\begin{aligned} \Bigg[\underbrace{\mathfrak{g}}_{=\sum_{i\geq 1}\mathfrak{g}_i}, \underbrace{\mathfrak{g}}_{=\sum_{j\geq 1}\mathfrak{g}_j}\Bigg] &= \left[\sum_{i\geq 1}\mathfrak{g}_i, \sum_{j\geq 1}\mathfrak{g}_j\right] \\ &\subset \sum_{i\geq 1}\sum_{j\geq 1} \underbrace{[\mathfrak{g}_i, \mathfrak{g}_j]}_{\substack{\subset\mathfrak{g} \\ (\text{by } (13.152.6))}} \qquad (\text{since the Lie bracket on } T(V) \text{ is } \mathbf{k}\text{-bilinear}) \\ &\subset \sum_{i\geq 1}\sum_{j\geq 1} \mathfrak{g} \subset \mathfrak{g} \qquad (\text{since } \mathfrak{g} \text{ is a } \mathbf{k}\text{-module})\,. \end{aligned}$$

Thus, $\mathfrak{g}$ is a Lie subalgebra of $T(V)$. This solves Exercise 6.1.41(a).

(b) Let $\mathfrak{k}$ be any Lie subalgebra of $T(V)$ satisfying $V \subset \mathfrak{k}$. We must prove that $\mathfrak{g} \subset \mathfrak{k}$.

We claim that

$$\mathfrak{g}_k \subset \mathfrak{k} \qquad \text{for every positive integer } k. \tag{13.152.7}$$

Proof of (13.152.7): We shall prove (13.152.7) by induction over k:

Induction base: We have $\mathfrak{g}_1 = V \subset \mathfrak{k}$. In other words, (13.152.7) holds for $i = 1$. This completes the induction base.

Induction step: Let K be a positive integer. Assume that (13.152.7) holds for $k = K$. We must prove that (13.152.7) holds for $k = K + 1$.

We have assumed that (13.152.7) holds for $k = K$. In other words, we have $\mathfrak{g}_K \subset \mathfrak{k}$.

But $\mathfrak{k}$ is a Lie subalgebra of $T(V)$. Hence, $[\mathfrak{k}, \mathfrak{k}] \subset \mathfrak{k}$.

Now, the recursive definition of $\mathfrak{g}_{K+1}$ yields $\mathfrak{g}_{K+1} = \left[\underbrace{V}_{\subset \mathfrak{k}}, \underbrace{\mathfrak{g}_{(K+1)-1}}_{=\mathfrak{g}_K \subset \mathfrak{k}}\right] \subset [\mathfrak{k}, \mathfrak{k}] \subset \mathfrak{k}$. In other words, (13.152.7) holds for $k = K+1$. This completes the induction step. The induction proof of (13.152.7) is thus complete.

Now,

$$\mathfrak{g} = \sum_{i \geq 1} \underbrace{\mathfrak{g}_i}_{\substack{\subset \mathfrak{k} \\ \text{(by (13.152.7), applied to } k=i\text{)}}} \subset \sum_{i \geq 1} \mathfrak{k} \subset \mathfrak{k}$$

(since $\mathfrak{k}$ is a $\mathbf{k}$-module (since $\mathfrak{k}$ is a Lie algebra)). This solves Exercise 6.1.41(b).

Before we step to the solution of Exercise 6.1.41(c), we record a simple fact: If $u \in \mathfrak{A}^*$ and $v \in \mathfrak{A}^*$, then

$$x_u x_v = x_{uv} \tag{13.152.8}$$

in the $\mathbf{k}$-algebra $T(V)$ [1065].

(c) We shall solve Exercise 6.1.41(c) by strong induction on $\ell(w)$:

Induction step: Let $N \in \mathbb{N}$. Assume that Exercise 6.1.41(c) holds under the condition that $\ell(w) < N$. We need to show that Exercise 6.1.41(c) also holds under the condition that $\ell(w) = N$.

We have assumed that Exercise 6.1.41(c) holds under the condition that $\ell(w) < N$. In other words, we have

$$b_w \in x_w + \sum_{\substack{v \in \mathfrak{A}^{\ell(w)}; \\ v > w}} \mathbf{k} x_v \qquad \text{for every } w \in \mathfrak{L} \text{ satisfying } \ell(w) < N. \tag{13.152.10}$$

For every $w \in \mathfrak{L}$ satisfying $\ell(w) < N$, we have

$$\begin{aligned} b_w &\in x_w + \sum_{\substack{v \in \mathfrak{A}^{\ell(w)}; \\ v > w}} \mathbf{k} x_v \\ &= x_w + \sum_{\substack{p \in \mathfrak{A}^{\ell(w)}; \\ p > w}} \mathbf{k} x_p \qquad \text{(here, we renamed the summation index } v \text{ as } p) \qquad (13.152.11) \\ &= x_w + \sum_{\substack{q \in \mathfrak{A}^{\ell(w)}; \\ q > w}} \mathbf{k} x_q \qquad \text{(here, we renamed the summation index } p \text{ as } q)\,. \qquad (13.152.12) \end{aligned}$$

[1065] *Proof of (13.152.8):* Let $u \in \mathfrak{A}^*$ and $v \in \mathfrak{A}^*$. We have $uv = \left((uv)_1, (uv)_2, \ldots, (uv)_{\ell(uv)}\right)$ and thus

$$\begin{aligned} \left((uv)_1, (uv)_2, \ldots, (uv)_{\ell(uv)}\right) &= \underbrace{u}_{=(u_1, u_2, \ldots, u_{\ell(u)})} \underbrace{v}_{=(v_1, v_2, \ldots, v_{\ell(v)})} \\ &= \left(u_1, u_2, \ldots, u_{\ell(u)}\right)\left(v_1, v_2, \ldots, v_{\ell(v)}\right) = \left(u_1, u_2, \ldots, u_{\ell(u)}, v_1, v_2, \ldots, v_{\ell(v)}\right) \end{aligned}$$

and therefore

$$x_{(uv)_1} \otimes x_{(uv)_2} \otimes \cdots \otimes x_{(uv)_{\ell(uv)}} = x_{u_1} \otimes x_{u_2} \otimes \cdots \otimes x_{u_{\ell(u)}} \otimes x_{v_1} \otimes x_{v_2} \otimes \cdots \otimes x_{v_{\ell(v)}}.$$

The definition of x_{uv} now yields

$$x_{uv} = x_{(uv)_1} \otimes x_{(uv)_2} \otimes \cdots \otimes x_{(uv)_{\ell(uv)}} = x_{u_1} \otimes x_{u_2} \otimes \cdots \otimes x_{u_{\ell(u)}} \otimes x_{v_1} \otimes x_{v_2} \otimes \cdots \otimes x_{v_{\ell(v)}}. \tag{13.152.9}$$

But the definition of x_u yields $x_u = x_{u_1} \otimes x_{u_2} \otimes \cdots \otimes x_{u_{\ell(u)}}$, and the definition of x_v yields $x_v = x_{v_1} \otimes x_{v_2} \otimes \cdots \otimes x_{v_{\ell(v)}}$. Thus,

$$\begin{aligned} &\underbrace{x_u}_{=x_{u_1} \otimes x_{u_2} \otimes \cdots \otimes x_{u_{\ell(u)}}} \underbrace{x_v}_{=x_{v_1} \otimes x_{v_2} \otimes \cdots \otimes x_{v_{\ell(v)}}} \\ &= \left(x_{u_1} \otimes x_{u_2} \otimes \cdots \otimes x_{u_{\ell(u)}}\right)\left(x_{v_1} \otimes x_{v_2} \otimes \cdots \otimes x_{v_{\ell(v)}}\right) \\ &= x_{u_1} \otimes x_{u_2} \otimes \cdots \otimes x_{u_{\ell(u)}} \otimes x_{v_1} \otimes x_{v_2} \otimes \cdots \otimes x_{v_{\ell(v)}} \\ &= x_{uv} \qquad \text{(by (13.152.9))}\,. \end{aligned}$$

This proves (13.152.8).

From this, it is easy to conclude the following: For every $w \in \mathfrak{L}$ satisfying $\ell(w) < N$, we have

$$b_w \in \sum_{\substack{q \in \mathfrak{A}^{\ell(w)}; \\ q \geq w}} \mathbf{k} x_q \tag{13.152.13}$$

[1066].

Now, let $w \in \mathfrak{L}$ be such that $\ell(w) = N$. We shall prove that $b_w \in x_w + \sum_{\substack{r \in \mathfrak{A}^{\ell(w)}; \\ r > w}} \mathbf{k} x_r$.

If $\ell(w) \leq 1$, then $b_w \in x_w + \sum_{\substack{r \in \mathfrak{A}^{\ell(w)}; \\ r > w}} \mathbf{k} x_r$ holds[1067]. Hence, for the rest of this proof, we can WLOG assume that we don't have $\ell(w) \leq 1$. Assume this.

The word w is Lyndon (since $w \in \mathfrak{L}$) and satisfies $\ell(w) > 1$ (since we don't have $\ell(w) \leq 1$). Thus, w is a Lyndon word of length > 1. Therefore, $\operatorname{stf} w$ is well-defined. Let $(u, v) = \operatorname{stf} w$. The recursive definition of b_w yields $b_w = [b_u, b_v]$ (since $\ell(w) > 1$ and $(u, v) = \operatorname{stf} w$).

From Exercise 6.1.39(a) (applied to $(g, h) = (u, v)$), we see that v is the longest Lyndon proper suffix of w. In particular, v is a Lyndon proper suffix of w. Also, Exercise 6.1.39(d) (applied to $(g, h) = (u, v)$) shows that the word u is Lyndon. From Exercise 6.1.39(c) (applied to $(g, h) = (u, v)$), we obtain $u < uv < v$. Finally, Exercise 6.1.39(b) (applied to $(g, h) = (u, v)$) shows that $w = uv$. Exercise 6.1.39(e) (applied to $(g, h) = (u, v)$) shows that $u \in \mathfrak{L}$, $v \in \mathfrak{L}$, $\ell(u) < \ell(w)$ and $\ell(v) < \ell(w)$.

The word u is Lyndon and thus nonempty. Hence, $\ell(u) \geq 1$. Also, the word v is Lyndon and thus nonempty. Thus, $\ell(v) \geq 1$.

Now, $\ell(u) < \ell(w) = N$. Thus, (13.152.11) (applied to u instead of w) shows that

$$b_u \in x_u + \sum_{\substack{p \in \mathfrak{A}^{\ell(u)}; \\ p > u}} \mathbf{k} x_p. \tag{13.152.14}$$

[1066] *Proof of (13.152.13):* Let $w \in \mathfrak{L}$ be such that $\ell(w) < N$.

Every $q \in \mathfrak{A}^{\ell(w)}$ satisfying $q > w$ must also satisfy $q \geq w$. Thus, $\sum_{\substack{q \in \mathfrak{A}^{\ell(w)}; \\ q > w}} \mathbf{k} x_q \subset \sum_{\substack{q \in \mathfrak{A}^{\ell(w)}; \\ q \geq w}} \mathbf{k} x_q$.

On the other hand, w is an element of $\mathfrak{A}^{\ell(w)}$ (since $\ell(w) = \ell(w)$) and satisfies $w \geq w$. In other words, w is a $q \in \mathfrak{A}^{\ell(w)}$ satisfying $q \geq w$. Thus, $\mathbf{k} x_w$ is an addend of the sum $\sum_{\substack{q \in \mathfrak{A}^{\ell(w)}; \\ q \geq w}} \mathbf{k} x_q$. Hence, $\mathbf{k} x_w \subset \sum_{\substack{q \in \mathfrak{A}^{\ell(w)}; \\ q \geq w}} \mathbf{k} x_q$.

Now, $x_w \in \mathbf{k} x_w \subset \sum_{\substack{q \in \mathfrak{A}^{\ell(w)}; \\ q \geq w}} \mathbf{k} x_q$. But (13.152.12) becomes

$$b_w \in \underbrace{x_w}_{\in \sum_{\substack{q \in \mathfrak{A}^{\ell(w)}; \\ q \geq w}} \mathbf{k} x_q} + \underbrace{\sum_{\substack{q \in \mathfrak{A}^{\ell(w)}; \\ q > w}} \mathbf{k} x_q}_{\subset \sum_{\substack{q \in \mathfrak{A}^{\ell(w)}; \\ q \geq w}} \mathbf{k} x_q} \subset \sum_{\substack{q \in \mathfrak{A}^{\ell(w)}; \\ q \geq w}} \mathbf{k} x_q + \sum_{\substack{q \in \mathfrak{A}^{\ell(w)}; \\ q \geq w}} \mathbf{k} x_q \subset \sum_{\substack{q \in \mathfrak{A}^{\ell(w)}; \\ q \geq w}} \mathbf{k} x_q$$

(since $\sum_{\substack{q \in \mathfrak{A}^{\ell(w)}; \\ q \geq w}} \mathbf{k} x_q$ is a $\mathbf{k}$-module). This proves (13.152.13).

[1067] *Proof.* Assume that $\ell(w) \leq 1$. The word w is Lyndon (since $w \in \mathfrak{L}$) and thus nonempty. Hence, $\ell(w) \geq 1$. Combined with $\ell(w) \leq 1$, this yields $\ell(w) = 1$. In other words, the word w consists of a single letter. In other words, $w = (a)$ for some $a \in \mathfrak{A}$. Consider this a. The definition of b_w then yields $b_w = x_a$ (since $\ell(w) = 1$ and $w = (a)$).

On the other hand, the definition of x_w yields $x_w = x_a$ (since $w = (a)$). Compared with $b_w = x_a$, this yields

$$b_w = x_w = x_w + \underbrace{0}_{\substack{\in \sum_{\substack{r \in \mathfrak{A}^{\ell(w)}; \\ r > w}} \mathbf{k} x_r \\ \left(\text{since } \sum_{\substack{r \in \mathfrak{A}^{\ell(w)}; \\ r > w}} \mathbf{k} x_r \text{ is a } \mathbf{k}\text{-module}\right)}} \in x_w + \sum_{\substack{r \in \mathfrak{A}^{\ell(w)}; \\ r > w}} \mathbf{k} x_r,$$

qed.

Also, (13.152.13) (applied to u instead of w) shows that

$$b_u \in \sum_{\substack{q \in \mathfrak{A}^{\ell(u)}; \\ q \geq u}} \mathbf{k} x_q = \sum_{\substack{p \in \mathfrak{A}^{\ell(u)}; \\ p \geq u}} \mathbf{k} x_p \tag{13.152.15}$$

(here, we renamed the summation index q as p).

Also, $\ell(v) < \ell(w) = N$. Thus, (13.152.12) (applied to v instead of w) shows that

$$b_v \in x_v + \sum_{\substack{q \in \mathfrak{A}^{\ell(v)}; \\ q > v}} \mathbf{k} x_q. \tag{13.152.16}$$

Also, (13.152.13) (applied to v instead of w) shows that

$$b_v \in \sum_{\substack{q \in \mathfrak{A}^{\ell(v)}; \\ q \geq v}} \mathbf{k} x_q. \tag{13.152.17}$$

Now, let $G = \sum\limits_{\substack{r \in \mathfrak{A}^{\ell(w)}; \\ r > w}} \mathbf{k} x_r$. Thus, $G = \sum\limits_{\substack{r \in \mathfrak{A}^{\ell(w)}; \\ r > w}} \mathbf{k} x_r$ is a $\mathbf{k}$-submodule of $T(V)$.

If $p \in \mathfrak{A}^{\ell(u)}$ and $q \in \mathfrak{A}^{\ell(v)}$ are such that $q \geq v$, then

$$x_q x_p \in G \tag{13.152.18}$$

[1068]. Hence,

$$b_v b_u \in G \tag{13.152.19}$$

[1069].

Furthermore, if $p \in \mathfrak{A}^{\ell(u)}$ and $q \in \mathfrak{A}^{\ell(v)}$ are such that $p > u$, then

$$x_p x_q \in G \tag{13.152.20}$$

[1068] *Proof of (13.152.18):* Let $p \in \mathfrak{A}^{\ell(u)}$ and $q \in \mathfrak{A}^{\ell(v)}$ be such that $q \geq v$.

We have $\ell(qp) = \underbrace{\ell(q)}_{\substack{= \ell(v) \\ (\text{since } q \in \mathfrak{A}^{\ell(v)})}} + \underbrace{\ell(p)}_{\substack{= \ell(u) \\ (\text{since } p \in \mathfrak{A}^{\ell(u)})}} = \ell(v) + \ell(u) = \ell(u) + \ell(v)$. Compared with $\ell(w) = \ell(u) + \ell(v)$, this yields $\ell(qp) = \ell(w)$. In other words, $qp \in \mathfrak{A}^{\ell(w)}$.

But the word w is Lyndon and satisfies $w = uv$. Thus, Proposition 6.1.14(a) shows that $v \geq w$ (since v is nonempty). Since $\ell(v) \neq \ell(w)$ (because $\ell(v) < \ell(w)$), we have $v \neq w$. Combined with $v \geq w$, this yields $v > w$. Now, $qp \geq q \geq v > w$.

So we know that $qp \in \mathfrak{A}^{\ell(w)}$ and $qp > w$. In other words, qp is an $r \in \mathfrak{A}^{\ell(w)}$ satisfying $r > w$. Thus, $\mathbf{k} x_{qp}$ is an addend of the sum $\sum\limits_{\substack{r \in \mathfrak{A}^{\ell(w)}; \\ r > w}} \mathbf{k} x_r$. Therefore, $\mathbf{k} x_{qp} \subset \sum\limits_{\substack{r \in \mathfrak{A}^{\ell(w)}; \\ r > w}} \mathbf{k} x_r = G$ (since $G = \sum\limits_{\substack{r \in \mathfrak{A}^{\ell(w)}; \\ r > w}} \mathbf{k} x_r$).

But (13.152.8) (applied to q and p instead of u and v) shows that $x_q x_p = x_{qp} \in \mathbf{k} x_{qp} \subset G$. This proves (13.152.18).

[1069] *Proof of (13.152.19):* We have

$$\begin{aligned} & \underbrace{b_v}_{\substack{\in \sum\limits_{\substack{q \in \mathfrak{A}^{\ell(v)}; \\ q \geq v}} \mathbf{k} x_q \\ (\text{by } (13.152.17))}} \underbrace{b_u}_{\substack{\in \sum\limits_{\substack{p \in \mathfrak{A}^{\ell(u)}; \\ p \geq u}} \mathbf{k} x_p \\ (\text{by } (13.152.15))}} \\ & = \left(\sum_{\substack{q \in \mathfrak{A}^{\ell(v)}; \\ q \geq v}} \mathbf{k} x_q \right) \left(\sum_{\substack{p \in \mathfrak{A}^{\ell(u)}; \\ p \geq u}} \mathbf{k} x_p \right) \subset \sum_{\substack{q \in \mathfrak{A}^{\ell(v)}; \\ q \geq v}} \sum_{\substack{p \in \mathfrak{A}^{\ell(u)}; \\ p \geq u}} \underbrace{\mathbf{k}\mathbf{k}}_{\subset \mathbf{k}} \underbrace{x_q x_p}_{\substack{\in G \\ (\text{by } (13.152.18))}} \subset \sum_{\substack{q \in \mathfrak{A}^{\ell(v)}; \\ q \geq v}} \sum_{\substack{p \in \mathfrak{A}^{\ell(u)}; \\ p \geq u}} \mathbf{k} G \subset G \end{aligned}$$

(since G is a $\mathbf{k}$-module). This proves (13.152.19).

[1070]. Hence,

$$(b_u - x_u)\, b_v \in G \tag{13.152.21}$$

[1071].

Furthermore, if $q \in \mathfrak{A}^{\ell(v)}$ is such that $q > v$, then

$$x_u x_q \in G \tag{13.152.22}$$

[1072]. Hence,

$$x_u\,(b_v - x_v) \in G \tag{13.152.23}$$

[1070] *Proof of (13.152.20):* Let $p \in \mathfrak{A}^{\ell(u)}$ and $q \in \mathfrak{A}^{\ell(v)}$ be such that $p > u$.

We have $\ell(pq) = \underbrace{\ell(p)}_{\substack{=\ell(u) \\ (\text{since } p \in \mathfrak{A}^{\ell(u)})}} + \underbrace{\ell(q)}_{\substack{=\ell(v) \\ (\text{since } q \in \mathfrak{A}^{\ell(v)})}} = \ell(u) + \ell(v)$. Compared with $\ell(w) = \ell(u) + \ell(v)$, this yields $\ell(pq) = \ell(w)$. In other words, $pq \in \mathfrak{A}^{\ell(w)}$.

We have $\ell(p) = \ell(u)$ (since $p \in \mathfrak{A}^{\ell(u)}$), so that $\ell(u) = \ell(p)$.

Assume (for the sake of contradiction) that the word u is a prefix of p. Since the word u has the same length as p (since $\ell(u) = \ell(p)$), this shows that $u = p$. Thus, $u = p > u$, which is absurd. This contradiction shows that our assumption (that the word u is a prefix of p) was false. In other words, the word u is not a prefix of p.

We have $u < p$ (since $p > u$), thus $u \leq p$. Thus, Proposition 6.1.2(d) (applied to u, v, p and $\varnothing$ instead of a, b, c and d) shows that either we have $uv \leq p\varnothing$ or the word u is a prefix of p. Since the word u is not a prefix of p, we thus conclude that $uv \leq p\varnothing$. Hence, $w = uv \leq p\varnothing = p$.

Now, $q \in \mathfrak{A}^{\ell(v)}$, so that $\ell(q) = \ell(v) \geq 1$. The word q is thus nonempty. Hence, $p < pq$. Thus, $w \leq p < pq$, so that $pq > w$.

So we know that $pq \in \mathfrak{A}^{\ell(w)}$ and $pq > w$. In other words, pq is an $r \in \mathfrak{A}^{\ell(w)}$ satisfying $r > w$. Thus, $\mathbf{k}x_{pq}$ is an addend of the sum $\sum_{\substack{r \in \mathfrak{A}^{\ell(w)}; \\ r > w}} \mathbf{k}x_r$. Therefore, $\mathbf{k}x_{pq} \subset \sum_{\substack{r \in \mathfrak{A}^{\ell(w)}; \\ r > w}} \mathbf{k}x_r = G$ (since $G = \sum_{\substack{r \in \mathfrak{A}^{\ell(w)}; \\ r > w}} \mathbf{k}x_r$).

But (13.152.8) (applied to p and q instead of u and v) shows that $x_p x_q = x_{pq} \in \mathbf{k}x_{pq} \subset G$. This proves (13.152.20).

[1071] *Proof of (13.152.21):* Subtracting x_u from both sides of the relation (13.152.14), we obtain

$$b_u - x_u \in \sum_{\substack{p \in \mathfrak{A}^{\ell(u)}; \\ p > u}} \mathbf{k}x_p.$$

Now,

$$\begin{aligned} &\underbrace{(b_u - x_u)}_{\in \sum_{\substack{p \in \mathfrak{A}^{\ell(u)}; \\ p > u}} \mathbf{k}x_p} \underbrace{b_v}_{\substack{\in \sum_{\substack{q \in \mathfrak{A}^{\ell(v)}; \\ q \geq v}} \mathbf{k}x_q \\ (\text{by } (13.152.17))}} \\ &\in \left(\sum_{\substack{p \in \mathfrak{A}^{\ell(u)}; \\ p > u}} \mathbf{k}x_p \right) \left(\sum_{\substack{q \in \mathfrak{A}^{\ell(v)}; \\ q \geq v}} \mathbf{k}x_q \right) \subset \sum_{\substack{p \in \mathfrak{A}^{\ell(u)}; \\ p > u}} \sum_{\substack{q \in \mathfrak{A}^{\ell(v)}; \\ q \geq v}} \underbrace{\mathbf{k}\mathbf{k}}_{\subset \mathbf{k}} \underbrace{x_p x_q}_{\substack{\in G \\ (\text{by } (13.152.20))}} \subset \sum_{\substack{p \in \mathfrak{A}^{\ell(u)}; \\ p > u}} \sum_{\substack{q \in \mathfrak{A}^{\ell(v)}; \\ q \geq v}} \mathbf{k}G \subset G \end{aligned}$$

(since G is a $\mathbf{k}$-module). This proves (13.152.21).

[1072] *Proof of (13.152.22):* Let $q \in \mathfrak{A}^{\ell(v)}$ be such that $q > v$.

We have $\ell(uq) = \ell(u) + \underbrace{\ell(q)}_{\substack{=\ell(v) \\ (\text{since } q \in \mathfrak{A}^{\ell(v)})}} = \ell(u) + \ell(v)$. Compared with $\ell(w) = \ell(u) + \ell(v)$, this yields $\ell(uq) = \ell(w)$. In other words, $uq \in \mathfrak{A}^{\ell(w)}$.

We have $q > v$, thus $v < q$, thus $v \leq q$. Hence, Proposition 6.1.2(b) (applied to u, v and q instead of a, c and d) shows that $uv \leq uq$. If we had $uv = uq$, then we would have $v = q$ (since we could cancel u from the equality $uv = uq$), which would contradict $v < q$. Thus, we cannot have $uv = uq$. Hence, we have $uv \neq uq$. Combined with $uv \leq uq$, this shows that $uv < uq$. In other words, $uq > uv = w$ (since $w = uv$).

So we know that $uq \in \mathfrak{A}^{\ell(w)}$ and $uq > w$. In other words, uq is an $r \in \mathfrak{A}^{\ell(w)}$ satisfying $r > w$. Thus, $\mathbf{k}x_{uq}$ is an addend of the sum $\sum_{\substack{r \in \mathfrak{A}^{\ell(w)}; \\ r > w}} \mathbf{k}x_r$. Therefore, $\mathbf{k}x_{uq} \subset \sum_{\substack{r \in \mathfrak{A}^{\ell(w)}; \\ r > w}} \mathbf{k}x_r = G$ (since $G = \sum_{\substack{r \in \mathfrak{A}^{\ell(w)}; \\ r > w}} \mathbf{k}x_r$).

But (13.152.8) (applied to u and q instead of u and v) shows that $x_u x_q = x_{uq} \in \mathbf{k}x_{uq} \subset G$. This proves (13.152.22).

[1073].

Now,

$$\begin{aligned} b_w &= [b_u, b_v] = b_u b_v - b_v b_u \qquad (\text{by the definition of the Lie bracket on } T(V)) \\ &= \underbrace{(b_u - x_u) b_v}_{\substack{\in G \\ (\text{by (13.152.21)})}} + \underbrace{x_u (b_v - x_v)}_{\substack{\in G \\ (\text{by (13.152.23)})}} + x_u x_v - \underbrace{b_v b_u}_{\substack{\in G \\ (\text{by (13.152.19)})}} \\ &\qquad (\text{by straightforward computation}) \\ &\in G + G + x_u x_v - G = \underbrace{x_u x_v}_{\substack{= x_{uv} \\ (\text{by (13.152.8)})}} + \underbrace{G + G - G}_{\substack{\subset G \\ (\text{since } G \text{ is a } \mathbf{k}\text{-module})}} \subset \underbrace{x_{uv}}_{\substack{= x_w \\ (\text{since } uv = w)}} + \underbrace{G}_{= \sum_{\substack{r \in \mathfrak{A}^{\ell(w)}; \\ r > w}} \mathbf{k} x_r} \\ &= x_w + \sum_{\substack{r \in \mathfrak{A}^{\ell(w)}; \\ r > w}} \mathbf{k} x_r . \end{aligned}$$

Thus, $b_w \in x_w + \sum_{\substack{r \in \mathfrak{A}^{\ell(w)}; \\ r > w}} \mathbf{k} x_r$ is proven.

Let us now forget that we defined (u, v). We have

$$b_w \in x_w + \sum_{\substack{r \in \mathfrak{A}^{\ell(w)}; \\ r > w}} \mathbf{k} x_r = x_w + \sum_{\substack{v \in \mathfrak{A}^{\ell(w)}; \\ v > w}} \mathbf{k} x_v$$

(here, we have renamed the summation index r as v).

Now, let us forget that we fixed w. We thus have proven that

$$b_w \in x_w + \sum_{\substack{v \in \mathfrak{A}^{\ell(w)}; \\ v > w}} \mathbf{k} x_v \qquad \text{for every } w \in \mathfrak{L} \text{ satisfying } \ell(w) = N.$$

In other words, we have proven that Exercise 6.1.41(c) holds under the condition that $\ell(w) = N$. Thus, our induction is complete, and Exercise 6.1.41(c) is solved.

(d) We know that $\mathfrak{g}$ is a Lie subalgebra of $T(V)$ (by Exercise 6.1.41(a)). Thus, $[\mathfrak{g}, \mathfrak{g}] \subset \mathfrak{g}$.

We first notice that

$$(13.152.24) \qquad b_w \in \mathfrak{g} \qquad \text{for every } w \in \mathfrak{L}.$$

Proof of (13.152.24): We shall prove (13.152.24) by strong induction on $\ell(w)$:

Induction step: Let $N \in \mathbb{N}$. Assume that (13.152.24) holds under the condition that $\ell(w) < N$. We need to show that (13.152.24) also holds under the condition that $\ell(w) = N$.

We have assumed that (13.152.24) holds under the condition that $\ell(w) < N$. In other words, we have

$$(13.152.25) \qquad b_w \in \mathfrak{g} \qquad \text{for every } w \in \mathfrak{L} \text{ satisfying } \ell(w) < N.$$

Now, let $w \in \mathfrak{L}$ be such that $\ell(w) = N$. We shall prove that $b_w \in \mathfrak{g}$.

[1073] *Proof of (13.152.23):* Subtracting x_v from both sides of the relation (13.152.16), we obtain

$$b_v - x_v \in \sum_{\substack{q \in \mathfrak{A}^{\ell(v)}; \\ q > v}} \mathbf{k} x_q .$$

Now,

$$x_u \underbrace{(b_v - x_v)}_{\in \sum_{\substack{q \in \mathfrak{A}^{\ell(v)}; \\ q > v}} \mathbf{k} x_q} \in x_u \left(\sum_{\substack{q \in \mathfrak{A}^{\ell(v)}; \\ q > v}} \mathbf{k} x_q \right) \subset \sum_{\substack{q \in \mathfrak{A}^{\ell(v)}; \\ q > v}} \mathbf{k} \underbrace{x_u x_q}_{\substack{\in G \\ (\text{by (13.152.22)})}} \subset \sum_{\substack{q \in \mathfrak{A}^{\ell(v)}; \\ q > v}} \mathbf{k} G \subset G$$

(since G is a $\mathbf{k}$-module). This proves (13.152.23).

If $\ell(w) \leq 1$, then $b_w \in \mathfrak{g}$ holds[1074]. Hence, for the rest of this proof, we can WLOG assume that we don't have $\ell(w) \leq 1$. Assume this.

The word w is Lyndon (since $w \in \mathfrak{L}$) and satisfies $\ell(w) > 1$ (since we don't have $\ell(w) \leq 1$). Thus, w is a Lyndon word of length > 1. Therefore, $\operatorname{stf} w$ is well-defined. Let $(u, v) = \operatorname{stf} w$. The recursive definition of b_w yields $b_w = [b_u, b_v]$ (since $\ell(w) > 1$ and $(u, v) = \operatorname{stf} w$).

Exercise 6.1.39(e) (applied to $(g, h) = (u, v)$) shows that $u \in \mathfrak{L}$, $v \in \mathfrak{L}$, $\ell(u) < \ell(w)$ and $\ell(v) < \ell(w)$.

Now, $\ell(u) < \ell(w) = N$. Thus, (13.152.25) (applied to u instead of w) shows that $b_u \in \mathfrak{g}$.

Also, $\ell(v) < \ell(w) = N$. Thus, (13.152.25) (applied to v instead of w) shows that $b_v \in \mathfrak{g}$.

Now, $b_w = \left[\underbrace{b_u}_{\in \mathfrak{g}}, \underbrace{b_v}_{\in \mathfrak{g}}\right] \in [\mathfrak{g}, \mathfrak{g}] \subset \mathfrak{g}$.

Let us now forget that we fixed w. We thus have shown that $b_w \in \mathfrak{g}$ for every $w \in \mathfrak{L}$ satisfying $\ell(w) = N$. In other words, (13.152.24) holds under the condition that $\ell(w) = N$. This completes the induction step. The induction proof of (13.152.24) is thus complete.

Now, (13.152.24) shows that $(b_w)_{w \in \mathfrak{L}}$ is a family of elements of $\mathfrak{g}$. Using Exercise 6.1.40(a), it is easy to conclude that this family $(b_w)_{w \in \mathfrak{L}}$ spans the $\mathbf{k}$-module $\mathfrak{g}$ [1075].

We shall now show that the family $(b_w)_{w \in \mathfrak{L}}$ is $\mathbf{k}$-linearly independent. In order to do so, we will prove a more general result:

Proposition 13.152.1. *Let $\mathfrak{V}$ be a $\mathbf{k}$-module, and let W be a totally ordered set. Let $(p_w)_{w \in W}$ be a $\mathbf{k}$-linearly independent family of elements of $\mathfrak{V}$. Let L be a subset of W. Let $(s_w)_{w \in L}$ be a family of elements of $\mathfrak{V}$. Assume that every $w \in L$ satisfies*

$$s_w \in p_w + \sum_{\substack{v \in W; \\ v > w}} \mathbf{k} p_v. \tag{13.152.28}$$

(Here, the ">" sign under the sum refers to the total order on W.) Then, the family $(s_w)_{w \in L}$ is $\mathbf{k}$-linearly independent.

Proposition 13.152.1 is actually a standard criterion for linear independence. It is often summarized by the motto "vectors with distinct leading coordinates are linearly independent". For the sake of completeness, we shall nevertheless prove it. First, let us state an obvious lemma:

[1074] *Proof.* Assume that $\ell(w) \leq 1$. The word w is Lyndon (since $w \in \mathfrak{L}$) and thus nonempty. Hence, $\ell(w) \geq 1$. Combined with $\ell(w) \leq 1$, this yields $\ell(w) = 1$. In other words, the word w consists of a single letter. In other words, $w = (a)$ for some $a \in \mathfrak{A}$. Consider this a. The definition of b_w then yields $b_w = x_a$ (since $\ell(w) = 1$ and $w = (a)$).

Now, $b_w = x_a \in V \subset \mathfrak{g}$, qed.

[1075] *Proof.* Let B be the $\mathbf{k}$-submodule of $\mathfrak{g}$ spanned by the family $(b_w)_{w \in \mathfrak{L}}$ (this is well-defined since $(b_w)_{w \in \mathfrak{L}}$ is a family of elements of $\mathfrak{g}$). Then, clearly,

$$b_w \in B \qquad \text{for every } w \in \mathfrak{L}. \tag{13.152.26}$$

Whenever w is a Lyndon word of length > 1, we have

$$b_w = [b_u, b_v], \qquad \text{where } (u, v) = \operatorname{stf} w$$

(due to the recursive definition of b_w). Thus, Exercise 6.1.40(a) shows that B is a Lie subalgebra of $\mathfrak{g}$. Hence, B is a Lie subalgebra of $T(V)$.

Recall that $(x_a)_{a \in \mathfrak{A}}$ is a basis of the $\mathbf{k}$-module V. Hence,

$$\left(\text{the } \mathbf{k}\text{-linear span of the family } (x_a)_{a \in \mathfrak{A}}\right) = V. \tag{13.152.27}$$

Now, let $a \in \mathfrak{A}$. We shall show that $x_a \in B$.

Indeed, let w be the one-letter word (a). Then, w is Lyndon (since w is a one-letter word) and has length 1. Thus, $b_w = x_a$ (by the definition of b_w, since $w = (a)$). Hence, $x_a = b_w \in B$ (by (13.152.26)).

Now, let us forget that we fixed a. We thus have shown that $x_a \in B$ for every $a \in \mathfrak{A}$. Since B is a $\mathbf{k}$-module, this entails that $\left(\text{the } \mathbf{k}\text{-linear span of the family } (x_a)_{a \in \mathfrak{A}}\right) \subset B$. Because of (13.152.27), this rewrites as $V \subset B$.

Now, we know that B is a Lie subalgebra of $T(V)$ satisfying $V \subset B$. Thus, $\mathfrak{g} \subset B$ (by Exercise 6.1.41(b), applied to $\mathfrak{k} = B$). Combined with $B \subset \mathfrak{g}$ (since B is a $\mathbf{k}$-submodule of $\mathfrak{g}$), this yields $\mathfrak{g} = B$. But the family $(b_w)_{w \in \mathfrak{L}}$ spans the $\mathbf{k}$-module B (since B is the $\mathbf{k}$-submodule of $\mathfrak{g}$ spanned by the family $(b_w)_{w \in \mathfrak{L}}$). Since $\mathfrak{g} = B$, this rewrites as follows: The family $(b_w)_{w \in \mathfrak{L}}$ spans the $\mathbf{k}$-module $\mathfrak{g}$. Qed.

Lemma 13.152.2. *Let $\mathfrak{V}$ be a $\mathbf{k}$-module, and let W be a set. Let $(p_w)_{w\in W}$ be a $\mathbf{k}$-linearly independent family of elements of $\mathfrak{V}$. Let $x \in W$ and $\lambda \in \mathbf{k}$ be such that*

$$0 \in \lambda p_x + \sum_{v\in W\setminus\{x\}} \mathbf{k} p_v . \tag{13.152.29}$$

Then, $\lambda = 0$.

Proof of Lemma 13.152.2. The equality (13.152.29) shows that we can write 0 as a $\mathbf{k}$-linear combination of the elements p_v with $v \in W$ in such a way that the coefficient in front of p_x is λ. But since $(p_w)_{w\in W}$ is $\mathbf{k}$-linearly independent, the only such linear combination which gives 0 is the trivial one (i.e., the one where all the coefficients are 0). Hence, all coefficients in our linear combination are 0. In particular, the coefficient in front of p_x is 0. Since this coefficient is λ, this means that $\lambda = 0$. This proves Lemma 13.152.2. □

Proof of Proposition 13.152.1. Let us first prove the following fact: If S is a finite subset of L, then

$$\text{the family } (s_w)_{w\in S} \text{ is } \mathbf{k}\text{-linearly independent.} \tag{13.152.30}$$

Proof of (13.152.30): We shall prove (13.152.30) by induction over $|S|$:

Induction base: If S is a finite subset of L satisfying $|S| = 0$, then the family $(s_w)_{w\in S}$ is $\mathbf{k}$-linearly independent (because $|S| = 0$ entails $S = \varnothing$, and thus the family $(s_w)_{w\in S}$ is empty). In other words, (13.152.30) holds in the case when $|S| = 0$. This completes the induction base.

Induction step: Let K be a positive integer. Assume that (13.152.30) is proven in the case when $|S| = K - 1$. We must prove that (13.152.30) holds in the case when $|S| = K$.

We have assumed that (13.152.30) is proven in the case when $|S| = K - 1$. In other words, if S is a finite subset of L satisfying $|S| = K - 1$, then

$$\text{the family } (s_w)_{w\in S} \text{ is } \mathbf{k}\text{-linearly independent.} \tag{13.152.31}$$

Now, let S be a finite subset of L satisfying $|S| = K$. Let $(\lambda_w)_{w\in S} \in \mathbf{k}^S$ be a family of elements of $\mathbf{k}$ such that $\sum_{w\in S} \lambda_w s_w = 0$. We shall prove that $(\lambda_w)_{w\in S} = (0)_{w\in S}$.

We have $|S| = K > 0$. Thus, the set S is nonempty. Also, $S \subset L \subset W$, and thus the set S is totally ordered (since it is a subset of the totally ordered set W). Therefore, this set S has a smallest element (since every nonempty finite totally ordered set has a smallest element). Let x be this smallest element. Thus,

$$x \leq w \qquad \text{for every } w \in S \tag{13.152.32}$$

(by the definition of the smallest element). Also, $x \in S$ (since x is the smallest element of S), and thus $|S \setminus \{x\}| = \underbrace{|S|}_{=K} - 1 = K - 1$. Hence, we can apply (13.152.31) to $S \setminus \{x\}$ instead of S. As a result, we see that the family $(s_w)_{w\in S\setminus\{x\}}$ is $\mathbf{k}$-linearly independent. In other words, if $(\mu_w)_{w\in S\setminus\{x\}} \in \mathbf{k}^{S\setminus\{x\}}$ is a family of elements of $\mathbf{k}$ such that $\sum_{w\in S\setminus\{x\}} \mu_w s_w = 0$, then

$$(\mu_w)_{w\in S\setminus\{x\}} = (0)_{w\in S\setminus\{x\}} . \tag{13.152.33}$$

On the other hand, every $w \in S \setminus \{x\}$ satisfies

$$s_w \in \sum_{v\in W\setminus\{x\}} \mathbf{k} p_v \tag{13.152.34}$$

[1076]. Also,

$$s_x - p_x \in \sum_{v\in W\setminus\{x\}} \mathbf{k} p_v \tag{13.152.35}$$

[1076] *Proof of (13.152.34):* Let $w \in S \setminus \{x\}$. Thus, $w \in S$ and $w \neq x$.

Let $v \in \{q \in W \mid q > w\}$. We shall show that $v \in W \setminus \{x\}$.

Indeed, we have $v \in \{q \in W \mid q > w\}$. In other words, v is an element of W and satisfies $v > w$. But $w \in S \setminus \{x\} \subset S$, so that (13.152.32) shows that $x \leq w$. Hence, $w \geq x$, so that $v > w \geq x$. Thus, $v \neq x$. Combining $v \in W$ with $v \neq x$, we obtain $v \in W \setminus \{x\}$.

Let us now forget that we fixed v. We thus have proven that every $v \in \{q \in W \mid q > w\}$ satisfies $v \in W \setminus \{x\}$. In other words, $\{q \in W \mid q > w\} \subset W \setminus \{x\}$. Therefore, $\sum_{v\in\{q\in W \mid q>w\}} \mathbf{k} p_v \subset \sum_{v\in W\setminus\{x\}} \mathbf{k} p_v$.

Also, combining $w \in W$ with $w \neq x$, we obtain $w \in W \setminus \{x\}$. Hence, $\mathbf{k} p_w$ is an addend of the sum $\sum_{v\in W\setminus\{x\}} \mathbf{k} p_v$. Thus, $\mathbf{k} p_w \subset \sum_{v\in W\setminus\{x\}} \mathbf{k} p_v$. Now, $p_w \in \mathbf{k} p_w \subset \sum_{v\in W\setminus\{x\}} \mathbf{k} p_v$.

[1077]. Now, $\sum_{w\in S}\lambda_w s_w = 0$, so that

$$\begin{aligned}
0 &= \sum_{w\in S}\lambda_w s_w = \lambda_x \underbrace{s_x}_{=p_x+(s_x-p_x)} + \underbrace{\sum_{\substack{w\in S;\\ w\neq x}}}_{=\sum_{w\in S\setminus\{x\}}} \lambda_w s_w \qquad (\text{since } x\in S)\\
&= \underbrace{\lambda_x\left(p_x+(s_x-p_x)\right)}_{=\lambda_x p_x+\lambda_x(s_x-p_x)} + \sum_{w\in S\setminus\{x\}} \lambda_w \underbrace{s_w}_{\substack{\in\sum_{v\in W\setminus\{x\}}\mathbf{k}p_v\\ (\text{by (13.152.34)})}}\\
&\in \lambda_x p_x + \lambda_x \underbrace{(s_x-p_x)}_{\substack{\in\sum_{v\in W\setminus\{x\}}\mathbf{k}p_v\\ (\text{by (13.152.35)})}} + \underbrace{\sum_{w\in S\setminus\{x\}}\lambda_w \sum_{v\in W\setminus\{x\}}\mathbf{k}p_v}_{\substack{\subset\sum_{v\in W\setminus\{x\}}\mathbf{k}p_v\\ (\text{since } \sum_{v\in W\setminus\{x\}}\mathbf{k}p_v \text{ is a } \mathbf{k}\text{-module})}}\\
&\subset \lambda_x p_x + \underbrace{\lambda_x \sum_{v\in W\setminus\{x\}}\mathbf{k}p_v + \sum_{v\in W\setminus\{x\}}\mathbf{k}p_v}_{\substack{\subset\sum_{v\in W\setminus\{x\}}\mathbf{k}p_v\\ (\text{since } \sum_{v\in W\setminus\{x\}}\mathbf{k}p_v \text{ is a } \mathbf{k}\text{-module})}} \subset \lambda_x p_x + \sum_{v\in W\setminus\{x\}}\mathbf{k}p_v .
\end{aligned}$$

Lemma 13.152.2 (applied to $\lambda=\lambda_x$) thus shows that $\lambda_x = 0$. Thus,

$$0 = \underbrace{\lambda_x}_{=0} s_x + \underbrace{\sum_{\substack{w\in S;\\ w\neq x}}}_{=\sum_{w\in S\setminus\{x\}}} \lambda_w s_w = \underbrace{0 s_x}_{=0} + \sum_{w\in S\setminus\{x\}}\lambda_w s_w = \sum_{w\in S\setminus\{x\}}\lambda_w s_w .$$

Thus, the family $(\lambda_w)_{w\in S\setminus\{x\}} \in \mathbf{k}^{S\setminus\{x\}}$ satisfies $\sum_{w\in S\setminus\{x\}}\lambda_w s_w = 0$. Hence, $(\lambda_w)_{w\in S\setminus\{x\}} = (0)_{w\in S\setminus\{x\}}$ (according to (13.152.33), applied to $(\mu_w)_{w\in S\setminus\{x\}} = (\lambda_w)_{w\in S\setminus\{x\}}$). In other words,

$$\lambda_w = 0 \qquad \text{for every } w\in S\setminus\{x\}. \tag{13.152.36}$$

Combining this with the fact that $\lambda_x = 0$, we conclude that $\lambda_w = 0$ for every $w\in S$. In other words, $(\lambda_w)_{w\in S} = (0)_{w\in S}$.

Now, $w\in S\subset L\subset W$. Hence, (13.152.28) yields

$$\begin{aligned}
s_w &\in \underbrace{p_w}_{\in\sum_{v\in W\setminus\{x\}}\mathbf{k}p_v} + \underbrace{\sum_{\substack{v\in W;\\ v>w}}}_{=\sum_{v\in\{q\in W\ \mid\ q>w\}}}\mathbf{k}p_v \subset \sum_{v\in W\setminus\{x\}}\mathbf{k}p_v + \underbrace{\sum_{v\in\{q\in W\ \mid\ q>w\}}\mathbf{k}p_v}_{\subset\sum_{v\in W\setminus\{x\}}\mathbf{k}p_v}\\
&\subset \sum_{v\in W\setminus\{x\}}\mathbf{k}p_v + \sum_{v\in W\setminus\{x\}}\mathbf{k}p_v \subset \sum_{v\in W\setminus\{x\}}\mathbf{k}p_v
\end{aligned}$$

(since $\sum_{v\in W\setminus\{x\}}\mathbf{k}p_v$ is a $\mathbf{k}$-module). This proves (13.152.34).

[1077] *Proof of (13.152.35):* Let $v\in\{q\in W\ \mid\ q>x\}$. We shall show that $v\in W\setminus\{x\}$.

Indeed, we have $v\in\{q\in W\ \mid\ q>x\}$. In other words, v is an element of W and satisfies $v>x$. Hence, $v\neq x$ (since $v>x$). Combining $v\in W$ with $v\neq x$, we obtain $v\in W\setminus\{x\}$.

Let us now forget that we fixed v. We thus have proven that every $v\in\{q\in W\ \mid\ q>x\}$ satisfies $v\in W\setminus\{x\}$. In other words, $\{q\in W\ \mid\ q>x\}\subset W\setminus\{x\}$. Therefore, $\sum_{v\in\{q\in W\ \mid\ q>x\}}\mathbf{k}p_v \subset \sum_{v\in W\setminus\{x\}}\mathbf{k}p_v$.

Now, $x\in S\subset L\subset W$. Hence, (13.152.28) (applied to $w=x$) yields

$$s_x \in p_x + \sum_{\substack{v\in W;\\ v>x}}\mathbf{k}p_v .$$

Subtracting p_x from both sides of this relation, we obtain

$$s_x - p_x \in \underbrace{\sum_{\substack{v\in W;\\ v>x}}}_{=\sum_{v\in\{q\in W\ \mid\ q>x\}}}\mathbf{k}p_v = \sum_{v\in\{q\in W\ \mid\ q>x\}}\mathbf{k}p_v \subset \sum_{v\in W\setminus\{x\}}\mathbf{k}p_v .$$

This proves (13.152.35).

Now, let us forget that we fixed $(\lambda_w)_{w\in S}$. We thus have shown that if $(\lambda_w)_{w\in S} \in \mathbf{k}^S$ is a family of elements of $\mathbf{k}$ such that $\sum_{w\in S}\lambda_w s_w = 0$, then $(\lambda_w)_{w\in S} = (0)_{w\in S}$. In other words, the family $(s_w)_{w\in S}$ is $\mathbf{k}$-linearly independent.

Let us now forget that we fixed S. We thus have proven that if S is a finite subset of L satisfying $|S| = K$, then the family $(s_w)_{w\in S}$ is $\mathbf{k}$-linearly independent. In other words, (13.152.30) holds in the case when $|S| = K$. This completes the induction step. Thus, we have proven (13.152.30) by induction.

Now, a family $\mathfrak{f}$ of vectors in a $\mathbf{k}$-module is $\mathbf{k}$-linearly independent if every finite subfamily of $\mathfrak{f}$ is $\mathbf{k}$-linearly independent (because every linear dependence relation for $\mathfrak{f}$ has only finitely many nonzero coefficients, and thus can be recast as a linear dependence relation for some finite subfamily of $\mathfrak{f}$). Now, (13.152.30) shows that every finite subfamily of the family $(s_w)_{w\in L}$ is $\mathbf{k}$-linearly independent; therefore, the previous sentence shows that the family $(s_w)_{w\in L}$ is $\mathbf{k}$-linearly independent. This proves Proposition 13.152.1. □

Let us now return to the solution of Exercise 6.1.41(d). The set $\mathfrak{A}^*$ is totally ordered (by the lexicographic order), and the set $\mathfrak{L}$ is a subset of $\mathfrak{A}^*$. The family $(x_w)_{w\in\mathfrak{A}^*}$ is a basis of the $\mathbf{k}$-module $T(V)$, and thus is a $\mathbf{k}$-linearly independent family of elements of this $\mathbf{k}$-module. The family $(b_w)_{w\in\mathfrak{L}}$ is a family of elements of $T(V)$. Every $w \in \mathfrak{L}$ satisfies

$$b_w \in x_w + \underbrace{\sum_{\substack{v\in\mathfrak{A}^{\ell(w)};\\ v>w}} \mathbf{k}x_v}_{\substack{\subset \sum_{\substack{v\in\mathfrak{A}^*;\\ v>w}} \mathbf{k}x_v \\ (\text{since } \mathfrak{A}^{\ell(w)}\subset\mathfrak{A}^*)}} \qquad (\text{by Exercise 6.1.41(c)})$$
$$\subset x_w + \sum_{\substack{v\in\mathfrak{A}^*;\\ v>w}} \mathbf{k}x_v.$$

Hence, Proposition 13.152.1 (applied to $\mathfrak{V} = T(V)$, $W = \mathfrak{A}^*$, $(p_w)_{w\in\mathfrak{A}^*} = (x_w)_{w\in\mathfrak{A}^*}$, $L = \mathfrak{L}$ and $(s_w)_{w\in\mathfrak{L}} = (b_w)_{w\in\mathfrak{L}}$) shows that the family $(b_w)_{w\in\mathfrak{L}}$ is $\mathbf{k}$-linearly independent. Combining this with the fact that this family $(b_w)_{w\in\mathfrak{L}}$ spans the $\mathbf{k}$-module $\mathfrak{g}$, we therefore conclude that the family $(b_w)_{w\in\mathfrak{L}}$ is a basis of the $\mathbf{k}$-module $\mathfrak{g}$. This solves Exercise 6.1.41(d).

(e) Let us first show a simple lemma:

Lemma 13.152.3. *Let $\mathfrak{a}$ and $\mathfrak{b}$ be two $\mathbf{k}$-Lie algebras. Let $p : \mathfrak{a} \to \mathfrak{b}$ and $q : \mathfrak{a} \to \mathfrak{b}$ be two Lie algebra homomorphisms. Then, $\ker(p-q)$ is a Lie subalgebra of $\mathfrak{a}$.*

Proof of Lemma 13.152.3. Let $x \in \ker(p-q)$ and $y \in \ker(p-q)$. Then, $(p-q)(x) = 0$ (since $x \in \ker(p-q)$), so that $0 = (p-q)(x) = p(x) - q(x)$, and thus $p(x) = q(x)$. The same argument (applied to y instead of x) shows that $p(y) = q(y)$.

But p is a Lie algebra homomorphism, and thus satisfies $p([x,y]) = \left[\underbrace{p(x)}_{=q(x)}, \underbrace{p(y)}_{=q(y)}\right] = [q(x), q(y)] = q([x,y])$ (since q is a Lie algebra homomorphism). Now, $(p-q)([x,y]) = \underbrace{p([x,y])}_{=q([x,y])} - q([x,y]) = q([x,y]) - q([x,y]) = 0$. In other words, $[x,y] \in \ker(p-q)$.

Let us now forget that we fixed x and y. We thus have proven that $[x,y] \in \ker(p-q)$ whenever $x \in \ker(p-q)$ and $y \in \ker(p-q)$. In other words, the set $\ker(p-q)$ is closed under the Lie bracket. Also, clearly, $p-q$ is a $\mathbf{k}$-linear map (since both maps p and q are $\mathbf{k}$-linear), and therefore its kernel $\ker(p-q)$ is a $\mathbf{k}$-submodule of $\mathfrak{a}$. Thus, $\ker(p-q)$ is a $\mathbf{k}$-submodule of $\mathfrak{a}$ which is closed under the Lie bracket. In other words, $\ker(p-q)$ is a Lie subalgebra of $\mathfrak{a}$. This proves Lemma 13.152.3. □

Now, it is easy to see that if Ξ_1 and Ξ_2 are two Lie algebra homomorphisms $\Xi : \mathfrak{g} \to \mathfrak{h}$ such that every $a \in \mathfrak{A}$ satisfies $\Xi(x_a) = \xi(a)$, then $\Xi_1 = \Xi_2$ [1078]. In other words, there exists **at most one** Lie algebra

[1078] *Proof.* Let Ξ_1 and Ξ_2 be two Lie algebra homomorphisms $\Xi : \mathfrak{g} \to \mathfrak{h}$ such that every $a \in \mathfrak{A}$ satisfies $\Xi(x_a) = \xi(a)$. We must prove that $\Xi_1 = \Xi_2$.

homomorphism $\Xi : \mathfrak{g} \to \mathfrak{h}$ such that every $a \in \mathfrak{A}$ satisfies $\Xi(x_a) = \xi(a)$. We shall now show that there actually exists such a homomorphism.

For every $w \in \mathfrak{L}$, we define an element z_w of $\mathfrak{h}$ as follows: We define z_w by recursion on the length of w. If the length of w is 1 [1079], then we have $w = (a)$ for some letter $a \in \mathfrak{A}$, and we set $z_w = \xi(a)$ for this letter a. If the length of w is > 1, then we set $z_w = [z_u, z_v]$, where $(u, v) = \operatorname{stf} w$ [1080].

Thus, we have defined a $z_w \in \mathfrak{h}$ for every $w \in \mathfrak{L}$. Now, Exercise 6.1.41(d) shows that the family $(b_w)_{w\in\mathfrak{L}}$ is a basis of the $\mathbf{k}$-module $\mathfrak{g}$. Thus, we can define a $\mathbf{k}$-module homomorphism $f : \mathfrak{g} \to \mathfrak{h}$ by requiring that

$$(f(b_w) = z_w \qquad \text{for every } w \in \mathfrak{L}). \tag{13.152.37}$$

Consider this f. Whenever w is a Lyndon word of length > 1, we have

$$f([b_u, b_v]) = [f(b_u), f(b_v)], \qquad \text{where } (u, v) = \operatorname{stf} w \tag{13.152.38}$$

[1081]. But recall that the family $(b_w)_{w\in\mathfrak{L}}$ is a basis of the $\mathbf{k}$-module $\mathfrak{g}$. Thus, this family spans the $\mathbf{k}$-module $\mathfrak{g}$. In other words, $\mathfrak{g}$ is the $\mathbf{k}$-submodule of $\mathfrak{g}$ spanned by the family $(b_w)_{w\in\mathfrak{L}}$. Thus, Exercise 6.1.40(b) (applied to $B = \mathfrak{g}$) shows that f is a Lie algebra homomorphism.

We notice that

$$\text{every } a \in \mathfrak{A} \text{ satisfies } f(x_a) = \xi(a)$$

[1082]. Thus, f is a Lie algebra homomorphism $\mathfrak{g} \to \mathfrak{h}$ having the property that every $a \in \mathfrak{A}$ satisfies $f(x_a) = \xi(a)$. Therefore, there exists **at least one** Lie algebra homomorphism $\Xi : \mathfrak{g} \to \mathfrak{h}$ such that every $a \in \mathfrak{A}$ satisfies $\Xi(x_a) = \xi(a)$ (namely, $\Xi = f$). Combining this with the fact that there exists **at most one**

Both Ξ_1 and Ξ_2 are $\mathbf{k}$-linear maps, and therefore $\Xi_1 - \Xi_2$ is a $\mathbf{k}$-linear map as well. Hence, its kernel $\ker(\Xi_1 - \Xi_2)$ is a $\mathbf{k}$-submodule of $\mathfrak{g}$.

We know that Ξ_1 is a Lie algebra homomorphism $\Xi : \mathfrak{g} \to \mathfrak{h}$ such that every $a \in \mathfrak{A}$ satisfies $\Xi(x_a) = \xi(a)$. Thus, every $a \in \mathfrak{A}$ satisfies $\Xi_1(x_a) = \xi(a)$. Similarly, every $a \in \mathfrak{A}$ satisfies $\Xi_2(x_a) = \xi(a)$. Hence, every $a \in \mathfrak{A}$ satisfies $(\Xi_1 - \Xi_2)(x_a) = \underbrace{\Xi_1(x_a)}_{=\xi(a)} - \underbrace{\Xi_2(x_a)}_{=\xi(a)} = \xi(a) - \xi(a) = 0$, so that $x_a \in \ker(\Xi_1 - \Xi_2)$. In other words, the set $\ker(\Xi_1 - \Xi_2)$ contains x_a for every $a \in \mathfrak{A}$. Since $\ker(\Xi_1 - \Xi_2)$ is a $\mathbf{k}$-submodule of $\mathfrak{g}$, this shows that

$$\ker(\Xi_1 - \Xi_2) \supset \left(\text{the } \mathbf{k}\text{-linear span of the family } (x_a)_{a\in\mathfrak{A}}\right).$$

Since $\left(\text{the } \mathbf{k}\text{-linear span of the family } (x_a)_{a\in\mathfrak{A}}\right) = V$ (because the family $(x_a)_{a\in\mathfrak{A}}$ is a basis of the $\mathbf{k}$-module V), this rewrites as $\ker(\Xi_1 - \Xi_2) \supset V$. In other words, $V \subset \ker(\Xi_1 - \Xi_2)$.

From Lemma 13.152.3 (applied to $\mathfrak{a} = \mathfrak{g}$, $\mathfrak{b} = \mathfrak{h}$, $p = \Xi_1$ and $q = \Xi_2$), we see that $\ker(\Xi_1 - \Xi_2)$ is a Lie subalgebra of $\mathfrak{g}$. Since $\mathfrak{g}$ (in turn) is a Lie subalgebra of $T(V)$, this shows that $\ker(\Xi_1 - \Xi_2)$ is a Lie subalgebra of $T(V)$.

Thus, we know that $\ker(\Xi_1 - \Xi_2)$ is a Lie subalgebra of $T(V)$ satisfying $V \subset \ker(\Xi_1 - \Xi_2)$. Exercise 6.1.41(b) (applied to $\mathfrak{k} = \ker(\Xi_1 - \Xi_2)$) thus shows that $\mathfrak{g} \subset \ker(\Xi_1 - \Xi_2)$. Hence, $\Xi_1 - \Xi_2 = 0$, so that $\Xi_1 = \Xi_2$. Qed.

[1079] The length of any $w \in \mathfrak{L}$ must be at least 1. (Indeed, if $w \in \mathfrak{L}$, then the word w is Lyndon and thus nonempty, and hence its length must be at least 1.)

[1080] This is well-defined, because z_u and z_v have already been defined. [*Proof.* Let $(u, v) = \operatorname{stf} w$. Then, Exercise 6.1.39(e) (applied to $(g, h) = (u, v)$) shows that $u \in \mathfrak{L}$, $v \in \mathfrak{L}$, $\ell(u) < \ell(w)$ and $\ell(v) < \ell(w)$. Recall that we are defining z_w by recursion on the length of w. Hence, z_p is already defined for every $p \in \mathfrak{L}$ satisfying $\ell(p) < \ell(w)$. Applying this to $p = u$, we see that z_u is already defined (since $u \in \mathfrak{L}$ and $\ell(u) < \ell(w)$). The same argument (but applied to v instead of u) shows that z_v is already defined. Hence, z_u and z_v have already been defined. Thus, z_w is well-defined by $z_w = [z_u, z_v]$, qed.]

[1081] *Proof of (13.152.38):* Let w be a Lyndon word of length > 1. Let $(u, v) = \operatorname{stf} w$. Thus, $(u, v) = \operatorname{stf} w \in \mathfrak{L} \times \mathfrak{L}$.

The recursive definition of z_w shows that $z_w = [z_u, z_v]$ (since w is a Lyndon word of length > 1, and since $(u, v) = \operatorname{stf} w$). The recursive definition of b_w shows that $b_w = [b_u, b_v]$ (for the same reasons). Thus, $[b_u, b_v] = b_w$, so that

$$\begin{aligned} f\left(\underbrace{[b_u, b_v]}_{=b_w}\right) &= f(b_w) = z_w \qquad (\text{by } (13.152.37)) \\ &= [z_u, z_v]. \end{aligned}$$

Comparing this with

$$\left[\underbrace{f(b_u)}_{\substack{=z_u \\ (\text{by } (13.152.37), \text{ applied to } u \text{ instead of } w)}}, \underbrace{f(b_v)}_{\substack{=z_v \\ (\text{by } (13.152.37), \text{ applied to } v \text{ instead of } w)}}\right] = [z_u, z_v],$$

we obtain $f([b_u, b_v]) = [f(b_u), f(b_v)]$. This proves (13.152.38).

[1082] *Proof.* Let $a \in \mathfrak{A}$. Let w be the one-letter word (a). Then, w is a Lyndon word (since w is a one-letter word) and has length 1. Thus, the definition of z_w shows that $z_w = \xi(a)$ (since $w = (a)$).

Lie algebra homomorphism $\Xi : \mathfrak{g} \to \mathfrak{h}$ such that every $a \in \mathfrak{A}$ satisfies $\Xi(x_a) = \xi(a)$, we conclude that there exists a **unique** Lie algebra homomorphism $\Xi : \mathfrak{g} \to \mathfrak{h}$ such that every $a \in \mathfrak{A}$ satisfies $\Xi(x_a) = \xi(a)$. This solves Exercise 6.1.41(e).

13.153. **Solution to Exercise 6.2.7.** *Solution to Exercise 6.2.7.*

Proof of Remark 6.2.6. (a) Let I and J be two nonempty intervals of $\mathbb{Z}$ satisfying $I < J$. Then,

$$\text{every } i \in I \text{ and } j \in J \text{ satisfy } i < j \qquad (13.153.1)$$

(by the definition of $I < J$), since $I < J$.

Let $p \in I \cap J$. Then, $p \in I \cap J \subset I$ and $p \in I \cap J \subset J$. Hence, $p < p$ (by (13.153.1)), which is absurd. Now, let us forget that we fixed p. Thus, we have obtained a contradiction for every $p \in I \cap J$. Hence, there exists no $p \in I \cap J$. In other words, $I \cap J = \varnothing$, so that the sets I and J are disjoint. This proves Remark 6.2.6(a).

(b) Let I and J be two disjoint nonempty intervals of $\mathbb{Z}$. We need to prove that $I < J$ or $J < I$.

Let i_0 be the smallest element of I (this exists, since I is nonempty), and let j_0 be the smallest element of J (this exists, since J is nonempty). We WLOG assume that $i_0 \leq j_0$ (since otherwise, we can simply interchange I with J).

The intervals I and J are disjoint, and thus $I \cap J = \varnothing$.

Now, let $i \in I$ and $j \in J$ be arbitrary. Assume (for the sake of contradiction) that $i \geq j$. Notice that $j \geq j_0$ (since j is an element of J, whereas j_0 is the smallest element of J), so that $i \geq j \geq j_0$, so that $j_0 \leq i$.

Write the interval I in the form $[p : q]^+$ for some $p \in \mathbb{Z}$ and $q \in \mathbb{Z}$. Since $i \in I = [p : q]^+ = \{p+1, p+2, \ldots, q\}$, we have $p < i \leq q$. Since $i_0 \in I = [p : q]^+ = \{p+1, p+2, \ldots, q\}$, we have $p < i_0 \leq q$. Now, $p < i_0 \leq j_0$ and $j_0 \leq i \leq q$. Hence, $p < j_0 \leq q$ and thus $j_0 \in \{p+1, p+2, \ldots, q\} = [p : q]^+ = I$. Combined with $j_0 \in J$ (since j_0 is the smallest element of J), this yields $j_0 \in I \cap J = \varnothing$, which is absurd. Hence, our assumption (that $i \geq j$) was wrong. We thus have $i < j$.

Now, forget that we have fixed i and j. We thus have shown that every $i \in I$ and $j \in J$ satisfy $i < j$. In other words, $I < J$ (by the definition of $I < J$). Hence, $I < J$ or $J < I$. This proves Remark 6.2.6(b).

(c) For any ℓ-tuple $(I_1, I_2, \ldots, I_\ell)$ of nonempty intervals of $\mathbb{Z}$, we can state the following three properties (which might and might not be satisfied):

- *Property C1:* The intervals I_1, I_2, ..., I_ℓ form a set partition of the set $[0 : n]^+$, where $n = |\alpha|$.
- *Property C2:* We have $I_1 < I_2 < \cdots < I_\ell$.
- *Property C3:* We have $|I_i| = \alpha_i$ for every $i \in \{1, 2, \ldots, \ell\}$.

We have to prove that the interval system $\operatorname{intsys} \alpha$ is the unique ℓ-tuple $(I_1, I_2, \ldots, I_\ell)$ of nonempty intervals of $\mathbb{Z}$ satisfying these three properties C1, C2 and C3.

But the definition of b_w yields $b_w = x_a$ (since $\ell(w) = 1$ and $w = (a)$). Hence, $f\left(\underbrace{b_w}_{=x_a}\right) = f(x_a)$, so that $f(x_a) = f(b_w) = z_w$ (by (13.152.37)) and thus $f(x_a) = z_w = \xi(a)$, qed.

First, it is easy to check that the interval system $\operatorname{intsys}\alpha$ is an ℓ-tuple $(I_1, I_2, \ldots, I_\ell)$ of nonempty intervals of $\mathbb{Z}$ satisfying these three properties C1, C2 and C3.[1083] It remains to prove that it is the only such ℓ-tuple.

So, let us fix any ℓ-tuple $(I_1, I_2, \ldots, I_\ell)$ of nonempty intervals of $\mathbb{Z}$ satisfying the three properties C1, C2 and C3. We need to prove that this ℓ-tuple $(I_1, I_2, \ldots, I_\ell)$ is $\operatorname{intsys}\alpha$. In order to prove this, it is enough to show that

$$I_i = \left[\sum_{k=1}^{i-1} \alpha_k : \sum_{k=1}^{i} \alpha_k\right]^+ \qquad \text{for every } i \in \{1, 2, \ldots, \ell\} \tag{13.153.3}$$

(according to the definition of $\operatorname{intsys}\alpha$). So it remains to prove (13.153.3).

Set $n = |\alpha|$. The intervals I_1, I_2, $\ldots$, I_ℓ form a set partition of the set $[0 : n]^+$ (according to Property C1). Thus, the intervals I_1, I_2, $\ldots$, I_ℓ are disjoint, and satisfy $I_1 \cup I_2 \cup \cdots \cup I_\ell = [0 : n]^+$.

Now, let us show that

$$I_1 \cup I_2 \cup \cdots \cup I_u = \left[0 : \sum_{k=1}^{u} \alpha_k\right]^+ \qquad \text{for every } u \in \{0, 1, \ldots, \ell\}. \tag{13.153.4}$$

Proof of (13.153.4): We will prove (13.153.4) by induction over u.

[1083] *Proof.* Let $(I_1, I_2, \ldots, I_\ell)$ denote the interval system $\operatorname{intsys}\alpha$. We need to prove that the Properties C1, C2 and C3 are satisfied.

Let us first recall that

$$I_i = \left[\sum_{k=1}^{i-1} \alpha_k : \sum_{k=1}^{i} \alpha_k\right]^+ \qquad \text{for every } i \in \{1, 2, \ldots, \ell\} \tag{13.153.2}$$

(by the definition of $\operatorname{intsys}\alpha$).

Proof that Property C1 is satisfied: Let $n = |\alpha|$. Then, $\sum_{k=1}^{\ell} \alpha_k = |\alpha| = n$. Clearly, $I_i \subset [0 : n]^+$ for every $i \in \{1, 2, \ldots, \ell\}$. We have

$$0 = \sum_{k=1}^{0} \alpha_k \leq \sum_{k=1}^{1} \alpha_k \leq \cdots \leq \sum_{k=1}^{\ell} \alpha_k = n$$

(since α_1, α_2, $\ldots$, α_ℓ are positive integers). Hence, for every $x \in [0 : n]^+$, there exists precisely one $i \in \{1, 2, \ldots, \ell\}$ satisfying $\sum_{k=1}^{i-1} \alpha_k < x \leq \sum_{k=1}^{i} \alpha_k$. In other words, for every $x \in [0 : n]^+$, there exists precisely one $i \in \{1, 2, \ldots, \ell\}$ satisfying $x \in \left[\sum_{k=1}^{i-1} \alpha_k : \sum_{k=1}^{i} \alpha_k\right]^+$ (since $x \in \left[\sum_{k=1}^{i-1} \alpha_k : \sum_{k=1}^{i} \alpha_k\right]^+$ is equivalent to $\sum_{k=1}^{i-1} \alpha_k < x \leq \sum_{k=1}^{i} \alpha_k$). In other words, for every $x \in [0 : n]^+$, there exists precisely one $i \in \{1, 2, \ldots, \ell\}$ satisfying $x \in I_i$ (since $I_i = \left[\sum_{k=1}^{i-1} \alpha_k : \sum_{k=1}^{i} \alpha_k\right]^+$). In other words, the subsets I_1, I_2, $\ldots$, I_ℓ of $[0 : n]^+$ are disjoint, and their union is $[0 : n]^+$. In other words, the subsets I_1, I_2, $\ldots$, I_ℓ of $[0 : n]^+$ form a set partition of the set $[0 : n]^+$. This proves that Property C1 is satisfied.

Proof that Property C2 is satisfied: Let $u \in \{1, 2, \ldots, \ell - 1\}$. Let $i \in I_u$ and $j \in I_{u+1}$. Since $i \in I_u = \left[\sum_{k=1}^{u-1} \alpha_k : \sum_{k=1}^{u} \alpha_k\right]^+$ (by (13.153.2), applied to u instead of i), we have $\sum_{k=1}^{u-1} \alpha_k < i \leq \sum_{k=1}^{u} \alpha_k$. Since $j \in I_{u+1} = \left[\sum_{k=1}^{(u+1)-1} \alpha_k : \sum_{k=1}^{u+1} \alpha_k\right]^+$ (by (13.153.2), applied to $u + 1$ instead of i), we have $\sum_{k=1}^{(u+1)-1} \alpha_k < j \leq \sum_{k=1}^{u+1} \alpha_k$. Now, $i \leq \sum_{k=1}^{u} \alpha_k = \sum_{k=1}^{(u+1)-1} \alpha_k < j$.

Now, let us forget that we fixed i and j. We thus have proven that every $i \in I_u$ and $j \in I_{u+1}$ satisfy $i < j$. In other words, $I_u < I_{u+1}$ (by the definition of $I_u < I_{u+1}$).

Now, let us forget that we fixed u. We thus have shown that $I_u < I_{u+1}$ for every $u \in \{1, 2, \ldots, \ell - 1\}$. In other words, $I_1 < I_2 < \cdots < I_\ell$. This proves that Property C2 is satisfied.

Proof that Property C3 is satisfied: Let $i \in \{1, 2, \ldots, \ell\}$. We have $\sum_{k=1}^{i} \alpha_k = \sum_{k=1}^{i-1} \alpha_k + \underbrace{\alpha_i}_{>0} > \sum_{k=1}^{i-1} \alpha_k$. Now, (13.153.2) yields

$$|I_i| = \left|\left[\sum_{k=1}^{i-1} \alpha_k : \sum_{k=1}^{i} \alpha_k\right]^+\right| = \sum_{k=1}^{i} \alpha_k - \sum_{k=1}^{i-1} \alpha_k \qquad \left(\text{since } \sum_{k=1}^{i} \alpha_k > \sum_{k=1}^{i-1} \alpha_k\right)$$
$$= \alpha_i.$$

This proves that Property C3 is satisfied.

We thus have shown that the Properties C1, C2 and C3 are satisfied, qed.

Induction base: For $u = 0$, we have $I_1 \cup I_2 \cup \cdots \cup I_u = I_1 \cup I_2 \cup \cdots \cup I_0 = (\text{empty union}) = \varnothing$ and $[0 : \sum_{k=1}^{u} \alpha_k]^+ = \left[0 : \underbrace{\sum_{k=1}^{0} \alpha_k}_{=0}\right]^+ = [0 : 0]^+ = \varnothing$. Hence, for $u = 0$, both sides of the equality (13.153.4) are $\varnothing$. Thus, for $u = 0$, the equality (13.153.4) holds. This completes the induction base.

Induction step: Let $U \in \{1, 2, \ldots, \ell\}$. Assume that (13.153.4) holds for $u = U - 1$. We now need to prove that (13.153.4) holds for $u = U$ as well.

We have assumed that (13.153.4) holds for $u = U - 1$. In other words,

$$I_1 \cup I_2 \cup \cdots \cup I_{U-1} = \left[0 : \sum_{k=1}^{U-1} \alpha_k\right]^+ . \tag{13.153.5}$$

Let $\xi = \sum_{k=1}^{U-1} \alpha_k$. Then, $\xi \geq 0$, so that $\xi + 1 \geq 1$. Then,

$$n = |\alpha| = \sum_{k=1}^{\ell} \alpha_k = \sum_{k=1}^{U-1} \alpha_k + \underbrace{\sum_{k=U}^{\ell} \alpha_k}_{\substack{>0 \\ (\text{since all } \alpha_k \text{ are } >0, \\ \text{and since } U \leq \ell)}} > \sum_{k=1}^{U-1} \alpha_k = \xi .$$

Hence, $n \geq \xi + 1$ (since ξ and n are integers), so that $1 \leq \xi + 1 \leq n$. In other words, $\xi + 1 \in \{1, 2, \ldots, n\} = [0 : n]^+ = I_1 \cup I_2 \cup \cdots \cup I_\ell$ (since $I_1 \cup I_2 \cup \cdots \cup I_\ell = [0 : n]^+$). In other words, there exists some $v \in \{1, 2, \ldots, \ell\}$ such that $\xi + 1 \in I_v$. Consider this v. If we had $v < U$, then we would have

$$\begin{aligned} \xi + 1 \in I_v \subset I_1 \cup I_2 \cup \cdots \cup I_{U-1} \qquad & (\text{since } v \in \{1, 2, \ldots, U - 1\} \ (\text{because } v < U)) \\ = \left[0 : \underbrace{\sum_{k=1}^{U-1} \alpha_k}_{=\xi}\right]^+ = [0 : \xi]^+ = \{1, 2, \ldots, \xi\} , & \end{aligned}$$

which is absurd. Hence, we cannot have $v < U$. Thus, we have $v \geq U$.

Recall that

$$\left[0 : \underbrace{\xi}_{=\sum_{k=1}^{U-1} \alpha_k}\right]^+ = \left[0 : \sum_{k=1}^{U-1} \alpha_k\right]^+ = I_1 \cup I_2 \cup \cdots \cup I_{U-1} \tag{13.153.6}$$

(by (13.153.5)). Using this, it is easy to see that

$$\text{every } p \in I_U \text{ satisfies } p \geq \xi + 1. \tag{13.153.7}$$

[1084]

We now assume (for the sake of contradiction) that $v \neq U$. Then, $v > U$ (since $v \geq U$ and $v \neq U$), whence $I_U < I_v$ (since Property C2 yields $I_1 < I_2 < \cdots < I_\ell$).

The interval I_U is nonempty, and thus there exists some $p \in I_U$. Consider such a p. Recall that $I_U < I_v$. Thus, every $i \in I_U$ and $j \in I_v$ satisfy $i < j$ (by the definition of $I_U < I_v$). Applying this to $i = p$ and $j = \xi + 1$, we obtain $p < \xi + 1$. But (13.153.7) yields $p \geq \xi + 1$, which contradicts $p < \xi + 1$. This contradiction proves that our assumption (that $v \neq U$) was wrong. Hence, we have $v = U$.

[1084] *Proof of (13.153.7):* Let $p \in I_U$. Assume (for the sake of contradiction) that $p \leq \xi$. But $p \in I_U \subset I_1 \cup I_2 \cup \cdots \cup I_\ell = [0 : n]^+ = \{1, 2, \ldots, n\}$, so that $0 < p \leq n$. Since $0 < p \leq \xi$, we have $p \in \{1, 2, \ldots, \xi\} = [0 : \xi]^+ = I_1 \cup I_2 \cup \cdots \cup I_{U-1}$, which shows that there exists some $r \in \{1, 2, \ldots, U - 1\}$ such that $p \in I_r$. Consider this r. Since $r \neq U$ (because $r \in \{1, 2, \ldots, U - 1\}$), the intervals I_r and I_U are disjoint (since the intervals I_1, I_2, ..., I_ℓ are disjoint). In other words, $I_r \cap I_U = \varnothing$. But combining $p \in I_r$ with $p \in I_U$, we obtain $p \in I_r \cap I_U = \varnothing$, which is absurd. This contradiction proves that our assumption (that $p \leq \xi$) was wrong. Hence, we have $p > \xi$. Thus, $p \geq \xi + 1$ (since p and ξ are integers). This proves (13.153.7).

Now, $\xi+1 \in I_v = I_U$ (since $v=U$). Hence, $\xi+1$ is an element of I_U. Due to (13.153.7), this element $\xi+1$ is the smallest element of I_U.

But since property C3 is satisfied, we have $|I_U| = \alpha_U$ (by Property C3, applied to $i=U$). So we know that the interval I_U has length α_U (since $|I_U| = \alpha_U$) and smallest element $\xi+1$. Therefore, $I_U = [\xi : \xi+\alpha_U]^+$ (because the only interval having length α_U and smallest element $\xi+1$ is the interval $[\xi : \xi+\alpha_U]^+$). Now,

$$I_1 \cup I_2 \cup \cdots \cup I_U = \underbrace{(I_1 \cup I_2 \cup \cdots \cup I_{U-1})}_{\substack{=[0:\xi]^+ \\ \text{(by (13.153.6))}}} \cup \underbrace{I_U}_{=[\xi:\xi+\alpha_U]^+}$$
$$= [0:\xi]^+ \cup [\xi : \xi+\alpha_U]^+ = [0 : \xi+\alpha_U]^+$$

(since $0 \le \xi \le \xi+\alpha_U$ (since $\xi \ge 0$ and $\alpha_U \ge 0$)). Since

$$\underbrace{\xi}_{=\sum_{k=1}^{U-1}\alpha_k} + \alpha_U = \sum_{k=1}^{U-1}\alpha_k + \alpha_U = \sum_{k=1}^{U}\alpha_k,$$

this rewrites as $I_1 \cup I_2 \cup \cdots \cup I_U = \left[0 : \sum_{k=1}^{U}\alpha_k\right]^+$. In other words, (13.153.4) holds for $u=U$. This completes the induction step. The induction proof of (13.153.4) is thus complete.

Proof of (13.153.3): Let $i \in \{1,2,\ldots,\ell\}$. Notice that $\sum_{k=1}^{i-1}\alpha_k \ge 0$ (since $\alpha_k > 0$ for every k) and $\sum_{k=1}^{i}\alpha_k = \sum_{k=1}^{i-1}\alpha_k + \underbrace{\alpha_i}_{>0} > \sum_{k=1}^{i-1}\alpha_k$. Hence, $0 \le \sum_{k=1}^{i-1}\alpha_k \le \sum_{k=1}^{i}\alpha_k$.

The intervals $I_1, I_2, \ldots, I_\ell$ are disjoint. Hence, I_i is disjoint from $I_1 \cup I_2 \cup \cdots \cup I_{i-1}$. Thus,

$$I_i = \underbrace{((I_1 \cup I_2 \cup \cdots \cup I_{i-1}) \cup I_i)}_{=I_1\cup I_2\cup\cdots\cup I_i} \setminus (I_1 \cup I_2 \cup \cdots \cup I_{i-1})$$
$$= \underbrace{(I_1 \cup I_2 \cup \cdots \cup I_i)}_{\substack{=\left[0:\sum_{k=1}^{i}\alpha_k\right]^+ \\ \text{(by (13.153.4), applied to } u=i)}} \setminus \underbrace{(I_1 \cup I_2 \cup \cdots \cup I_{i-1})}_{\substack{=\left[0:\sum_{k=1}^{i-1}\alpha_k\right]^+ \\ \text{(by (13.153.4), applied to } u=i-1)}}$$
$$= \left[0 : \sum_{k=1}^{i}\alpha_k\right]^+ \setminus \left[0 : \sum_{k=1}^{i-1}\alpha_k\right]^+ = \left[\sum_{k=1}^{i-1}\alpha_k : \sum_{k=1}^{i}\alpha_k\right]^+$$

(since $0 \le \sum_{k=1}^{i-1}\alpha_k \le \sum_{k=1}^{i}\alpha_k$). This proves (13.153.3). As explained above, this completes our proof of Remark 6.2.6(c). □

13.154. **Solution to Exercise 6.2.9.** *Solution to Exercise 6.2.9.*

Proof of Lemma 6.2.8. We have $\sigma \in \mathrm{Sh}_{n,m}$, so that $\sigma^{-1}(1) < \sigma^{-1}(2) < \cdots < \sigma^{-1}(n)$ and $\sigma^{-1}(n+1) < \sigma^{-1}(n+2) < \cdots < \sigma^{-1}(n+m)$. In other words, the restriction of the map σ^{-1} to the interval $[0:n]^+$ is strictly increasing, and so is the restriction of the map σ^{-1} to the interval $[n:n+m]^+$.

(a) Let I be an interval of $\mathbb{Z}$ such that $I \subset [0:n+m]^+$. We will only show that $\sigma(I) \cap [0:n]^+$ is an interval; the proof that $\sigma(I) \cap [n:n+m]^+$ is an interval is completely analogous.

It is known that

$$\left(\begin{array}{c}\text{if } \mathfrak{R} \text{ is a finite subset of } \mathbb{Z} \text{ such that every } \alpha \in \mathfrak{R},\ \gamma \in \mathfrak{R} \text{ and } \beta \in \mathbb{Z} \\ \text{satisfying } \alpha < \beta < \gamma \text{ satisfy } \beta \in \mathfrak{R}, \text{ then } \mathfrak{R} \text{ is an interval of } \mathbb{Z}\end{array}\right). \tag{13.154.1}$$

(Indeed, it is clear that $\mathfrak{R} = [\min\mathfrak{R} - 1 : \max\mathfrak{R}]^+$ in this case, unless $\mathfrak{R}$ is empty in which case the statement is obvious anyway.)

We now denote $\mathfrak{R} = \sigma(I) \cap [0:n]^+$. Our next goal is to use (13.154.1) to show that $\mathfrak{R}$ is an interval.

Indeed, let $\alpha \in \mathfrak{R}$, $\gamma \in \mathfrak{R}$ and $\beta \in \mathbb{Z}$ be such that $\alpha < \beta < \gamma$. Then, $\alpha \in \mathfrak{R} = \sigma(I) \cap [0:n]^+ \subset [0:n]^+$ and similarly $\gamma \in [0:n]^+$. Combined with $\alpha < \beta < \gamma$, these yield $\beta \in [0:n]^+$. But recall that the restriction of the map σ^{-1} to the interval $[0:n]^+$ is strictly increasing. Hence, from $\alpha < \beta < \gamma$, we obtain

$\sigma^{-1}(\alpha) < \sigma^{-1}(\beta) < \sigma^{-1}(\gamma)$ (since α, β and γ belong to $[0:n]^+$). In other words, the integer $\sigma^{-1}(\beta)$ lies strictly between the integers $\sigma^{-1}(\alpha)$ and $\sigma^{-1}(\gamma)$. Since $\sigma^{-1}(\alpha) \in I$ (because $\alpha \in \sigma(I) \cap [0:n]^+ \subset \sigma(I)$) and $\sigma^{-1}(\gamma) \in I$ (for similar reasons), this entails $\sigma^{-1}(\beta) \in I$ (because I is an interval, and thus any integer lying between two elements of I must also belong to I). Hence, $\beta \in \sigma(I)$. Combined with $\beta \in [0:n]^+$, this yields $\beta \in \sigma(I) \cap [0:n]^+ = \mathfrak{R}$.

Now, forget that we fixed α, γ and β. We thus have shown that every $\alpha \in \mathfrak{R}$, $\gamma \in \mathfrak{R}$ and $\beta \in \mathbb{Z}$ satisfying $\alpha < \beta < \gamma$ satisfy $\beta \in \mathfrak{R}$. Thus, (13.154.1) shows that $\mathfrak{R}$ is an interval of $\mathbb{Z}$. In other words, $\sigma(I) \cap [0:n]^+$ is an interval of $\mathbb{Z}$ (since $\mathfrak{R} = \sigma(I) \cap [0:n]^+$). This completes the proof of Lemma 6.2.8(a).

(b) The intervals K and L both are subsets of $[0:n]^+$. Therefore, from $K < L$, we obtain $\sigma^{-1}(K) < \sigma^{-1}(L)$ (because the restriction of the map σ^{-1} to the interval $[0:n]^+$ is strictly increasing).

Set $\mathcal{K} = \sigma^{-1}(K)$ and $\mathcal{L} = \sigma^{-1}(L)$. Then, $\mathcal{K}$ is a nonempty interval[1085], and thus can be written in the form $\mathcal{K} = [x_\mathcal{K} : y_\mathcal{K}]^+$ for two elements $x_\mathcal{K}$ and $y_\mathcal{K}$ of $[0:n+m]^+$ satisfying $x_\mathcal{K} < y_\mathcal{K}$. Consider these $x_\mathcal{K}$ and $y_\mathcal{K}$. Also, $\mathcal{L}$ is a nonempty interval[1086], and thus can be written in the form $\mathcal{L} = [x_\mathcal{L} : y_\mathcal{L}]^+$ for two elements $x_\mathcal{L}$ and $y_\mathcal{L}$ of $[0:n+m]^+$ satisfying $x_\mathcal{L} < y_\mathcal{L}$. Consider these $x_\mathcal{L}$ and $y_\mathcal{L}$. Since $[x_\mathcal{K} : y_\mathcal{K}]^+ = \mathcal{K} = \sigma^{-1}(K) < \sigma^{-1}(L) = \mathcal{L} = [x_\mathcal{L} : y_\mathcal{L}]^+$, we have $y_\mathcal{K} \leq x_\mathcal{L}$.

Notice that $\sigma(\mathcal{K}) = K$ (since $\mathcal{K} = \sigma^{-1}(K)$) and $\sigma(\mathcal{L}) = L$ (since $\mathcal{L} = \sigma^{-1}(L)$).

If we had $y_\mathcal{K} = x_\mathcal{L}$, then $\underbrace{\sigma^{-1}(K)}_{\substack{=\mathcal{K}=[x_\mathcal{K}:y_\mathcal{K}]^+=[x_\mathcal{K}:x_\mathcal{L}]^+ \\ (\text{since } y_\mathcal{K}=x_\mathcal{L})}} \cup \underbrace{\sigma^{-1}(L)}_{=\mathcal{L}=[x_\mathcal{L}:y_\mathcal{L}]^+} = [x_\mathcal{K} : x_\mathcal{L}]^+ \cup [x_\mathcal{L} : y_\mathcal{L}]^+ = [x_\mathcal{K} : y_\mathcal{L}]^+$ would be an interval, which would contradict the assumption that $\sigma^{-1}(K) \cup \sigma^{-1}(L)$ is not an interval. Hence, we cannot have $y_\mathcal{K} = x_\mathcal{L}$. Thus, $y_\mathcal{K} \neq x_\mathcal{L}$. Therefore, $y_\mathcal{K} < x_\mathcal{L}$ (since $y_\mathcal{K} \leq x_\mathcal{L}$). Hence, we can define a nonempty interval $\mathcal{P} \subset [0:n+m]^+$ by $\mathcal{P} = [y_\mathcal{K} : x_\mathcal{L}]^+$. Consider this $\mathcal{P}$. It satisfies $|\mathcal{P}| \neq 0$ (since it is nonempty) and $\mathcal{K} < \mathcal{P} < \mathcal{L}$ (since $\mathcal{K} = [x_\mathcal{K} : y_\mathcal{K}]^+$, $\mathcal{P} = [y_\mathcal{K} : x_\mathcal{L}]^+$ and $\mathcal{L} = [x_\mathcal{L} : y_\mathcal{L}]^+$), so that the sets $\mathcal{K}$, $\mathcal{P}$ and $\mathcal{L}$ are disjoint. As a consequence, the sets $\sigma(\mathcal{K})$, $\sigma(\mathcal{P})$ and $\sigma(\mathcal{L})$ are disjoint. In other words, the sets K, $\sigma(\mathcal{P})$ and L are disjoint (since $\sigma(\mathcal{K}) = K$ and $\sigma(\mathcal{L}) = L$).

It is easy to see that $\sigma(\mathcal{P}) \subset [n:n+m]^+$ [1087]. Thus, $\sigma(\mathcal{P}) \cap [n:n+m]^+ = \sigma(\mathcal{P})$. But Lemma 6.2.8(a) (applied to $\mathcal{P}$ instead of I) shows that $\sigma(\mathcal{P}) \cap [0:n]^+$ and $\sigma(\mathcal{P}) \cap [n:n+m]^+$ are intervals. In particular, $\sigma(\mathcal{P}) \cap [n:n+m]^+$ is an interval. In other words, $\sigma(\mathcal{P})$ is an interval (since $\sigma(\mathcal{P}) \cap [n:n+m]^+ = \sigma(\mathcal{P})$).

[1085] because $\mathcal{K} = \sigma^{-1}(K)$ and because we know that $\sigma^{-1}(K)$ is an interval and K is nonempty

[1086] because $\mathcal{L} = \sigma^{-1}(L)$ and because we know that $\sigma^{-1}(L)$ is an interval and L is nonempty

[1087] *Proof.* Assume the contrary. Then, there exists some $q \in \sigma(\mathcal{P})$ such that $q \notin [n:n+m]^+$. Consider this q. Since $q \notin [n:n+m]^+$, we must have $q \in [0:n]^+$.

Notice that $q \in \sigma(\mathcal{P})$, so that $\sigma^{-1}(q) \in \mathcal{P}$.

The elements $\max K$ and $\min L$ of K and L are well-defined, since K and L are nonempty.

Now, $\max K \leq q - 1$. (To prove this, assume the contrary. Thus, $\max K > q - 1$, so that $\max K \geq q$. But $\max K \in K \subset [0:n]^+$ and $q \in [0:n]^+$. Therefore, $\sigma^{-1}(\max K) \geq \sigma^{-1}(q)$ (because $\max K \geq q$, and since the restriction of the map σ^{-1} to the interval $[0:n]^+$ is strictly increasing). But since $\mathcal{K} < \mathcal{P}$, we have $\sigma^{-1}(\max K) < \sigma^{-1}(q)$ (since $\sigma^{-1}\left(\underbrace{\max K}_{\in K}\right) \in \sigma^{-1}(K) = \mathcal{K}$ and $\sigma^{-1}(q) \in \mathcal{P}$), which contradicts $\sigma^{-1}(\max K) \geq \sigma^{-1}(q)$. This contradiction shows that our assumption was wrong, and we have shown that $\max K \leq q - 1$.)

Furthermore, $q \leq \min L - 1$. (To prove this, assume the contrary. Thus, $q > \min L - 1$. Hence, $q \geq \min L$. We have $\min L \in L \subset [0:n]^+$ and $q \in [0:n]^+$. Therefore, $\sigma^{-1}(q) \geq \sigma^{-1}(\min L)$ (because $q \geq \min L$, and since the restriction of the map σ^{-1} to the interval $[0:n]^+$ is strictly increasing). But since $\mathcal{P} < \mathcal{L}$, we have $\sigma^{-1}(q) < \sigma^{-1}(\min L)$ (since $\sigma^{-1}(q) \in \mathcal{P}$ and $\sigma^{-1}\left(\underbrace{\min L}_{\in L}\right) \in \sigma^{-1}(L) = \mathcal{L}$), which contradicts $\sigma^{-1}(q) \geq \sigma^{-1}(\min L)$. This contradiction shows that our assumption was wrong, and we have shown that $q \leq \min L - 1$.)

So we have $\max K \leq \underbrace{q}_{\leq \min L - 1} - 1 \leq \min L - 2$. But K and L are disjoint nonempty intervals satisfying $K < L$, and their union is an interval again (because $K \cup L$ is an interval). Hence, the interval L must begin immediately after the end of the interval K; in other words, we must have $\max K = \min L - 1 > \min L - 2$. This contradicts $\max K \leq \min L - 2$. This contradiction completes our proof.

So we know that $\sigma(\mathcal{P}) \subset [n : n+m]^+$ is a nonempty interval[1088], and $\sigma^{-1}(\sigma(\mathcal{P}))$ is also an interval (since $\sigma^{-1}(\sigma(\mathcal{P})) = \mathcal{P}$). Moreover, we have $\sigma^{-1}(K) < \sigma^{-1}(\sigma(\mathcal{P}))$ (since $\sigma^{-1}(K) = \mathcal{K} < \mathcal{P} = \sigma^{-1}(\sigma(\mathcal{P}))$). Also, $\underbrace{\sigma^{-1}(K)}_{=\mathcal{K}=[x_{\mathcal{K}}:y_{\mathcal{K}}]^+} \cup \underbrace{\sigma^{-1}(\sigma(\mathcal{P}))}_{=\mathcal{P}=[y_{\mathcal{K}}:x_{\mathcal{L}}]^+} = [x_{\mathcal{K}} : y_{\mathcal{K}}]^+ \cup [y_{\mathcal{K}} : x_{\mathcal{L}}]^+ = [x_{\mathcal{K}} : x_{\mathcal{L}}]^+$ is an interval. Furthermore, $\sigma^{-1}(\sigma(\mathcal{P})) < \sigma^{-1}(L)$ (since $\sigma^{-1}(\sigma(\mathcal{P})) = \mathcal{P} < \mathcal{L} = \sigma^{-1}(L)$), and the set $\underbrace{\sigma^{-1}(\sigma(\mathcal{P}))}_{=\mathcal{P}=[y_{\mathcal{K}}:x_{\mathcal{L}}]^+} \cup \underbrace{\sigma^{-1}(L)}_{=\mathcal{L}=[x_{\mathcal{L}}:y_{\mathcal{L}}]^+} = [y_{\mathcal{K}} : x_{\mathcal{L}}]^+ \cup [x_{\mathcal{L}} : y_{\mathcal{L}}]^+ = [y_{\mathcal{K}} : y_{\mathcal{L}}]^+$ is an interval.

Altogether, we now know that $\sigma(\mathcal{P}) \subset [n : n+m]^+$ is a nonempty interval such that $\sigma^{-1}(\sigma(\mathcal{P}))$, $\sigma^{-1}(K) \cup \sigma^{-1}(\sigma(\mathcal{P}))$ and $\sigma^{-1}(\sigma(\mathcal{P})) \cup \sigma^{-1}(L)$ are intervals and such that $\sigma^{-1}(K) < \sigma^{-1}(\sigma(\mathcal{P})) < \sigma^{-1}(L)$. Thus, there exists a nonempty interval $P \subset [n : n+m]^+$ such that $\sigma^{-1}(P)$, $\sigma^{-1}(K) \cup \sigma^{-1}(P)$ and $\sigma^{-1}(P) \cup \sigma^{-1}(L)$ are intervals and such that $\sigma^{-1}(K) < \sigma^{-1}(P) < \sigma^{-1}(L)$ (namely, $P = \sigma(\mathcal{P})$). This proves Lemma 6.2.8(b).

(c) The proof of Lemma 6.2.8(c) is analogous to the proof of Lemma 6.2.8(b). □

13.155. **Solution to Exercise 6.2.11.** *Solution to Exercise 6.2.11.*

Proof of Lemma 6.2.10. We have $\sigma \in \operatorname{Sh}_{n,m}$, so that $\sigma^{-1}(1) < \sigma^{-1}(2) < \cdots < \sigma^{-1}(n)$ and $\sigma^{-1}(n+1) < \sigma^{-1}(n+2) < \cdots < \sigma^{-1}(n+m)$. In other words, the restriction of the map σ^{-1} to the interval $[0 : n]^+$ is strictly increasing, and so is the restriction of the map σ^{-1} to the interval $[n : n+m]^+$.

(a) Let I be an interval of $\mathbb{Z}$ satisfying either $I \subset [0 : n]^+$ or $I \subset [n : n+m]^+$. Assume that $\sigma^{-1}(I)$ is an interval.

Recall that the restriction of the map σ^{-1} to the interval $[0 : n]^+$ is strictly increasing, and so is the restriction of the map σ^{-1} to the interval $[n : n+m]^+$. Hence, the restriction of the map σ^{-1} to the interval I is strictly increasing (since either $I \subset [0 : n]^+$ or $I \subset [n : n+m]^+$).

Write the interval I in the form $I = [\alpha : \beta]^+$ where $0 \leq \alpha \leq \beta \leq n+m$. Write the interval $\sigma^{-1}(I)$ in the form $\sigma^{-1}(I) = [a : b]^+$ where $0 \leq a \leq b \leq n+m$. Since σ is bijective, we have $\left|\sigma^{-1}(I)\right| = |I| = \beta - \alpha$ (since $I = [\alpha : \beta]^+$). Compared with $\left|\sigma^{-1}(I)\right| = b - a$ (since $\sigma^{-1}(I) = [a : b]^+$), this yields $\beta - \alpha = b - a$.

The restriction of the map σ^{-1} to the interval I is injective; hence, it can be viewed as a bijection $I \to \sigma^{-1}(I)$. This bijection must be strictly increasing (since the restriction of the map σ^{-1} to the interval I is strictly increasing), and thus is a strictly increasing bijection $[\alpha : \beta]^+ \to [a : b]^+$ (since it goes from $I = [\alpha : \beta]^+$ to $\sigma^{-1}(I) = [a : b]^+$). But the only such bijection is the one sending every $x \in [\alpha : \beta]^+$ to $x - \alpha + a$. Hence, our bijection $I \to \sigma^{-1}(I)$ must be the map sending every $x \in [\alpha : \beta]^+$ to $x - \alpha + a$. Since our bijection comes from restricting the map σ^{-1}, it thus follows that the map σ^{-1} sends every $x \in [\alpha : \beta]^+$ to $x - \alpha + a$. Thus, $\sigma^{-1}(x) = x - \alpha + a$ for every $x \in [\alpha : \beta]^+$. Substituting y for $x - \alpha + a$ in this fact, we conclude that $\sigma^{-1}(y + \alpha - a) = y$ for every $y \in [a : \beta - \alpha + a]^+$. In other words, $\sigma^{-1}(y + \alpha - a) = y$ for every $y \in [a : b]^+$ (since $\beta - \alpha + a = b$ (because $\beta - \alpha = b - a$)). In other words, $\sigma(y) = y + \alpha - a$ for every $y \in [a : b]^+$. Thus,

$$(\sigma(a+1), \sigma(a+2), \ldots, \sigma(b)) = \left(\alpha+1, \alpha+2, \ldots, \underbrace{b + \alpha - a}_{\substack{=\beta \\ (\text{since } \beta-\alpha=b-a)}}\right) = (\alpha+1, \alpha+2, \ldots, \beta).$$

[1088] nonempty because $\mathcal{P}$ is nonempty

But let $w = (w_1, w_2, \ldots, w_{n+m})$ denote the word uv. The definition of $u \underset{\sigma}{\sqcup\!\sqcup} v$ then yields $u \underset{\sigma}{\sqcup\!\sqcup} v = \left(w_{\sigma(1)}, w_{\sigma(2)}, \ldots, w_{\sigma(n+m)}\right)$, so that

$$\begin{aligned}
&\underbrace{\left(u \underset{\sigma}{\sqcup\!\sqcup} v\right)}_{=\left(w_{\sigma(1)},w_{\sigma(2)},\ldots,w_{\sigma(n+m)}\right)} \left[\underbrace{\sigma^{-1}(I)}_{=[a:b]^+}\right] \\
&= \left(w_{\sigma(1)}, w_{\sigma(2)}, \ldots, w_{\sigma(n+m)}\right) \left[[a:b]^+\right] = \left(w_{\sigma(a+1)}, w_{\sigma(a+2)}, \ldots, w_{\sigma(b)}\right) \\
&= \left(w_{\alpha+1}, w_{\alpha+2}, \ldots, w_{\beta}\right) \qquad \left(\text{since } \left(\sigma(a+1), \sigma(a+2), \ldots, \sigma(b)\right) = (\alpha+1, \alpha+2, \ldots, \beta)\right) \\
&= \underbrace{w}_{=uv} \left[\underbrace{[\alpha:\beta]^+}_{=I}\right] = (uv)[I].
\end{aligned}$$

This proves Lemma 6.2.10(a).

(b) Notice that $\sigma^{-1}(I)$ is a nonempty interval (since I is nonempty and $\sigma^{-1}(I)$ is an interval). Hence, we can write the interval $\sigma^{-1}(I)$ in the form $[a:b]^+$ for some elements a and b of $[0:n+m]^+$ satisfying $a<b$. Consider these a and b. Similarly, write $\sigma^{-1}(J)$ in the form $[c:d]^+$ for some elements c and d of $[0:n+m]^+$ satisfying $c<d$. We have $b \leq c$ (since $\sigma^{-1}(I) < \sigma^{-1}(J)$), but we cannot have $b<c$ (because $\underbrace{[a:b]^+}_{=\sigma^{-1}(I)} \cup \underbrace{[c:d]^+}_{=\sigma^{-1}(J)} = \sigma^{-1}(I) \cup \sigma^{-1}(J)$ is an interval). Thus, we have $b=c$. Hence, $[b:d]^+ = [c:d]^+ = \sigma^{-1}(J)$ and $b=c<d$.

Now, let $A = [0:a]^+$ and $Z = [d:n+m]^+$. Then, the interval $[0:n+m]^+$ is the union of the disjoint intervals A, $\sigma^{-1}(I)$, $\sigma^{-1}(J)$ and Z (because

$$[0:n+m]^+ = \underbrace{[0:a]^+}_{=A} \sqcup \underbrace{[a:b]^+}_{=\sigma^{-1}(I)} \sqcup \underbrace{[b:d]^+}_{=\sigma^{-1}(J)} \sqcup \underbrace{[d:n+m]^+}_{=Z} = A \sqcup \sigma^{-1}(I) \sqcup \sigma^{-1}(J) \sqcup Z$$

), and these intervals satisfy $A < \sigma^{-1}(I) < \sigma^{-1}(J) < Z$ (since $A = [0:a]^+$, $\sigma^{-1}(I) = [a:b]^+$, $\sigma^{-1}(J) = [b:d]^+$ and $Z = [d:n+m]^+$). Therefore,

$$p = p[A] \cdot p\left[\sigma^{-1}(I)\right] \cdot p\left[\sigma^{-1}(J)\right] \cdot p[Z] \qquad \text{for every word } p \in \mathfrak{B}^{n+m} \tag{13.155.1}$$

for any alphabet $\mathfrak{B}$. Applying this to $p=\sigma$ and $\mathfrak{B} = \{1, 2, \ldots, n+m\}$, we obtain

$$\sigma = \sigma[A] \cdot \sigma\left[\sigma^{-1}(I)\right] \cdot \sigma\left[\sigma^{-1}(J)\right] \cdot \sigma[Z] \tag{13.155.2}$$

(where we consider the permutation σ as a word in $\{1, 2, \ldots, n+m\}^{n+m}$ by writing it in one-line notation).

Now, define a word τ by

$$\tau = \sigma[A] \cdot \sigma\left[\sigma^{-1}(J)\right] \cdot \sigma\left[\sigma^{-1}(I)\right] \cdot \sigma[Z]. \tag{13.155.3}$$

This word τ is obtained from the word σ by switching the two factors $\sigma\left[\sigma^{-1}(I)\right]$ and $\sigma\left[\sigma^{-1}(J)\right]$ (this is clear by comparing (13.155.3) with (13.155.2)), and thus is a permutation written in one-line notation (because σ is a permutation). In other words, $\tau \in \mathfrak{S}_{n+m}$.

From (13.155.3), we have

$$\begin{aligned}\tau &= \sigma\left[\underbrace{A}_{=[0:a]^+}\right]\cdot\sigma\left[\underbrace{\sigma^{-1}(J)}_{=[b:d]^+}\right]\cdot\sigma\left[\underbrace{\sigma^{-1}(I)}_{=[a:b]^+}\right]\cdot\sigma\left[\underbrace{Z}_{=[d:n+m]^+}\right]\\
&= \underbrace{\sigma\left[[0:a]^+\right]}_{=(\sigma(1),\sigma(2),\ldots,\sigma(a))}\cdot\underbrace{\sigma\left[[b:d]^+\right]}_{=(\sigma(b+1),\sigma(b+2),\ldots,\sigma(d))}\cdot\underbrace{\sigma\left[[a:b]^+\right]}_{=(\sigma(a+1),\sigma(a+2),\ldots,\sigma(b))}\cdot\underbrace{\sigma\left[[d:n+m]^+\right]}_{=(\sigma(d+1),\sigma(d+2),\ldots,\sigma(n+m))}\\
&= (\sigma(1),\sigma(2),\ldots,\sigma(a))\cdot(\sigma(b+1),\sigma(b+2),\ldots,\sigma(d))\\
&\qquad\cdot(\sigma(a+1),\sigma(a+2),\ldots,\sigma(b))\cdot(\sigma(d+1),\sigma(d+2),\ldots,\sigma(n+m))\\
&= (\sigma(1),\sigma(2),\ldots,\sigma(a),\sigma(b+1),\sigma(b+2),\ldots,\sigma(d),\\
&\qquad\sigma(a+1),\sigma(a+2),\ldots,\sigma(b),\sigma(d+1),\sigma(d+2),\ldots,\sigma(n+m)).\end{aligned}$$

Thus,

$$(\tau(1),\tau(2),\ldots,\tau(a)) = (\sigma(1),\sigma(2),\ldots,\sigma(a)); \tag{13.155.4}$$

$$(\tau(a+1),\tau(a+2),\ldots,\tau(a+d-b)) = (\sigma(b+1),\sigma(b+2),\ldots,\sigma(d)); \tag{13.155.5}$$

$$(\tau(a+d-b+1),\tau(a+d-b+2),\ldots,\tau(d)) = (\sigma(a+1),\sigma(a+2),\ldots,\sigma(b)); \tag{13.155.6}$$

$$(\tau(d+1),\tau(d+2),\ldots,\tau(n+m)) = (\sigma(d+1),\sigma(d+2),\ldots,\sigma(n+m)). \tag{13.155.7}$$

From this, it is easy to see that $\tau^{-1}(J) = [a:a+d-b]^+$ [1089] and $\tau^{-1}(I) = [a+d-b:d]^+$ [1090]. In particular, $\tau^{-1}(J)$ and $\tau^{-1}(I)$ are intervals. Now, obviously, the intervals $[0:a]^+$, $[a:a+d-b]^+$, $[a+d-b:d]^+$ and $[d:n+m]^+$ are disjoint intervals having union $[0:n+m]^+$ and satisfying $[0:a]^+ < [a:a+d-b]^+ < [a+d-b:d]^+ < [d:n+m]^+$. Since $A=[0:a]^+$, $\tau^{-1}(J) = [a:a+d-b]^+$, $\tau^{-1}(I) = [a+d-b:d]^+$ and $Z=[d:n+m]^+$, this rewrites as follows: The intervals A, $\tau^{-1}(J)$, $\tau^{-1}(I)$ and Z are disjoint intervals having union $[0:n+m]^+$ and satisfying $A<\tau^{-1}(J)<\tau^{-1}(I)<Z$. Therefore,

$$p = p[A]\cdot p\left[\tau^{-1}(J)\right]\cdot p\left[\tau^{-1}(I)\right]\cdot p[Z] \qquad \text{for every word } p\in\mathfrak{B}^{n+m} \tag{13.155.8}$$

for every alphabet $\mathfrak{B}$.

Next, we claim that τ belongs to $\operatorname{Sh}_{n,m}$. In fact, let i and j be two elements of $\{1,2,\ldots,n\}$ such that $i<j$. Our next goal is to prove that $\tau^{-1}(i)<\tau^{-1}(j)$.

Indeed, assume the contrary. Thus, $\tau^{-1}(i)\geq\tau^{-1}(j)$, so that $\tau^{-1}(i)>\tau^{-1}(j)$ (since τ is a permutation). In other words, the letter i lies further right than the letter j in the word τ. But we also have $\sigma^{-1}(i)<\sigma^{-1}(j)$ (since $\sigma^{-1}(1)<\sigma^{-1}(2)<\cdots<\sigma^{-1}(n)$ and $i<j$ and since $i,j\in\{1,2,\ldots,n\}$), which means that the letter i lies further left than the letter j in the word σ.

So the order in which the letters i and j appear in the word τ is different from that in σ. Since the word τ is obtained from the word σ by switching the two adjacent factors $\sigma\left[\sigma^{-1}(I)\right]$ and $\sigma\left[\sigma^{-1}(J)\right]$, this is only possible if one of the letters i and j is contained in one of these two factors and the other is contained in the other factor. Hence, this is what must be happening. In particular, one of the letters i and j must be contained in the factor $\sigma\left[\sigma^{-1}(J)\right]$. Since all letters of $\sigma\left[\sigma^{-1}(J)\right]$ belong to the set $[n:n+m]^+$ (because the letters of $\sigma\left[\sigma^{-1}(J)\right]$ are precisely the elements of J, but we have $J\subset[n:n+m]^+$), this yields that

[1089] In fact, this follows from

$$\begin{aligned}\tau\left(\underbrace{[a:a+d-b]^+}_{=\{a+1,a+2,\ldots,a+d-b\}}\right) &= \tau(\{a+1,a+2,\ldots,a+d-b\}) = \{\tau(a+1),\tau(a+2),\ldots,\tau(a+d-b)\}\\
&= \{\sigma(b+1),\sigma(b+2),\ldots,\sigma(d)\} \qquad \text{(by (13.155.5))}\\
&= \sigma\left(\underbrace{\{b+1,b+2,\ldots,d\}}_{=[b:d]^+=\sigma^{-1}(J)}\right) = \sigma\left(\sigma^{-1}(J)\right) = J.\end{aligned}$$

[1090] The proof of this is similar to that of $\tau^{-1}(J) = [a:a+d-b]^+$, but we have to use (13.155.6) this time.

one of the letters i and j must belong to the set $[n : n+m]^+$. But this is impossible, since the letters i and j belong to $\{1, 2, \ldots, n\}$ and thus neither of them can be an element of $[n : n+m]^+$ (because $[n : n+m]^+$ is disjoint with $\{1, 2, \ldots, n\}$). This contradiction shows that our assumption was wrong, and so we do have $\tau^{-1}(i) < \tau^{-1}(j)$.

Now, forget that we fixed i and j. We have shown that any two elements i and j of $\{1, 2, \ldots, n\}$ such that $i < j$ must satisfy $\tau^{-1}(i) < \tau^{-1}(j)$. Thus, $\tau^{-1}(1) < \tau^{-1}(2) < \cdots < \tau^{-1}(n)$. Similarly, $\tau^{-1}(n+1) < \tau^{-1}(n+2) < \cdots < \tau^{-1}(n+m)$. These two chains of inequalities, together, yield that $\tau \in \operatorname{Sh}_{n,m}$. Hence, $u \underset{\tau}{\shuffle} v$ is a well-defined element of the multiset $u \shuffle v$. Since $u \underset{\sigma}{\shuffle} v$ is the lexicographically highest element of this multiset, this shows that

$$u \underset{\sigma}{\shuffle} v \geq u \underset{\tau}{\shuffle} v. \tag{13.155.9}$$

Let $\mathfrak{a}$ denote the word $\left(u \underset{\tau}{\shuffle} v\right)[A]$, and let $\mathfrak{z}$ denote the word $\left(u \underset{\tau}{\shuffle} v\right)[Z]$.

We can now apply (13.155.1) to $p = u \underset{\sigma}{\shuffle} v$ and $\mathfrak{B} = \mathfrak{A}$. As a result, we obtain

$$u \underset{\sigma}{\shuffle} v = \left(u \underset{\sigma}{\shuffle} v\right)[A] \cdot \left(u \underset{\sigma}{\shuffle} v\right)\left[\sigma^{-1}(I)\right] \cdot \left(u \underset{\sigma}{\shuffle} v\right)\left[\sigma^{-1}(J)\right] \cdot \left(u \underset{\sigma}{\shuffle} v\right)[Z]. \tag{13.155.10}$$

Let now $(w_1, w_2, \ldots, w_{n+m})$ denote the word uv. The definition of $u \underset{\sigma}{\shuffle} v$ then yields $u \underset{\sigma}{\shuffle} v = \left(w_{\sigma(1)}, w_{\sigma(2)}, \ldots, w_{\sigma(n+m)}\right)$. Now, it is easy to see that

$$\left(u \underset{\sigma}{\shuffle} v\right)[A] = \mathfrak{a}$$

[1091] and

$$\left(u \underset{\sigma}{\shuffle} v\right)[Z] = \mathfrak{z}$$

[1092]. Thus, (13.155.10) becomes

$$\begin{aligned} u \underset{\sigma}{\shuffle} v &= \underbrace{\left(u \underset{\sigma}{\shuffle} v\right)[A]}_{=\mathfrak{a}} \cdot \underbrace{\left(u \underset{\sigma}{\shuffle} v\right)\left[\sigma^{-1}(I)\right]}_{\substack{=(uv)[I] \\ \text{(by (6.2.1))}}} \cdot \underbrace{\left(u \underset{\sigma}{\shuffle} v\right)\left[\sigma^{-1}(J)\right]}_{\substack{=(uv)[J] \\ \text{(by (6.2.1), applied to } J \\ \text{instead of } I)}} \cdot \underbrace{\left(u \underset{\sigma}{\shuffle} v\right)[Z]}_{=\mathfrak{z}} \\ &= \mathfrak{a} \cdot (uv)[I] \cdot (uv)[J] \cdot \mathfrak{z}. \end{aligned}$$

[1091] *Proof.* We have

$$\underbrace{\left(u \underset{\sigma}{\shuffle} v\right)}_{=\left(w_{\sigma(1)}, w_{\sigma(2)}, \ldots, w_{\sigma(n+m)}\right)} \left[\underbrace{A}_{=[0:a]^+}\right] = \left(w_{\sigma(1)}, w_{\sigma(2)}, \ldots, w_{\sigma(n+m)}\right)\left[[0:a]^+\right] = \left(w_{\sigma(1)}, w_{\sigma(2)}, \ldots, w_{\sigma(a)}\right).$$

The same argument, applied to τ instead of σ, shows that $\left(u \underset{\tau}{\shuffle} v\right)[A] = \left(w_{\tau(1)}, w_{\tau(2)}, \ldots, w_{\tau(a)}\right)$. But (13.155.4) yields $\left(w_{\tau(1)}, w_{\tau(2)}, \ldots, w_{\tau(a)}\right) = \left(w_{\sigma(1)}, w_{\sigma(2)}, \ldots, w_{\sigma(a)}\right)$, so that

$$\left(u \underset{\sigma}{\shuffle} v\right)[A] = \left(w_{\sigma(1)}, w_{\sigma(2)}, \ldots, w_{\sigma(a)}\right) = \left(w_{\tau(1)}, w_{\tau(2)}, \ldots, w_{\tau(a)}\right) = \left(u \underset{\tau}{\shuffle} v\right)[A] = \mathfrak{a},$$

qed.

[1092] *Proof.* We have

$$\underbrace{\left(u \underset{\sigma}{\shuffle} v\right)}_{=\left(w_{\sigma(1)}, w_{\sigma(2)}, \ldots, w_{\sigma(n+m)}\right)} \left[\underbrace{Z}_{=[d:n+m]^+}\right] = \left(w_{\sigma(1)}, w_{\sigma(2)}, \ldots, w_{\sigma(n+m)}\right)\left[[d:n+m]^+\right] = \left(w_{\sigma(d+1)}, w_{\sigma(d+2)}, \ldots, w_{\sigma(n+m)}\right).$$

The same argument, applied to τ instead of σ, shows that $\left(u \underset{\tau}{\shuffle} v\right)[Z] = \left(w_{\tau(d+1)}, w_{\tau(d+2)}, \ldots, w_{\tau(n+m)}\right)$. But (13.155.7) yields $\left(w_{\tau(d+1)}, w_{\tau(d+2)}, \ldots, w_{\tau(n+m)}\right) = \left(w_{\sigma(d+1)}, w_{\sigma(d+2)}, \ldots, w_{\sigma(n+m)}\right)$, so that

$$\left(u \underset{\sigma}{\shuffle} v\right)[Z] = \left(w_{\sigma(d+1)}, w_{\sigma(d+2)}, \ldots, w_{\sigma(n+m)}\right) = \left(w_{\tau(d+1)}, w_{\tau(d+2)}, \ldots, w_{\tau(n+m)}\right) = \left(u \underset{\tau}{\shuffle} v\right)[Z] = \mathfrak{z},$$

qed.

Thus,

$$\begin{aligned}
&\mathfrak{a}\cdot(uv)[I]\cdot(uv)[J]\cdot\mathfrak{z}\\
&=u\underset{\sigma}{\sqcup\!\sqcup}v\geq u\underset{\tau}{\sqcup\!\sqcup}v \qquad\qquad (\text{by } (13.155.9))\\
&=\underbrace{\left(u\underset{\tau}{\sqcup\!\sqcup}v\right)[A]}_{=\mathfrak{a}}\cdot\underbrace{\left(u\underset{\tau}{\sqcup\!\sqcup}v\right)\left[\tau^{-1}(J)\right]}_{\substack{=(uv)[J]\\ (\text{by } (6.2.1)\text{, applied to } \tau \text{ and } J\\ \text{instead of } \sigma \text{ and } I)}}\cdot\underbrace{\left(u\underset{\tau}{\sqcup\!\sqcup}v\right)\left[\tau^{-1}(I)\right]}_{\substack{=(uv)[I]\\ (\text{by } (6.2.1)\text{, applied to } \tau\\ \text{instead of } \sigma)}}\cdot\underbrace{\left(u\underset{\tau}{\sqcup\!\sqcup}v\right)[Z]}_{=\mathfrak{z}}\\
&\qquad\qquad\left(\text{by } (13.155.8)\text{, applied to } p=u\underset{\tau}{\sqcup\!\sqcup}v \text{ and } \mathfrak{B}=\mathfrak{A}\right)\\
&=\mathfrak{a}\cdot(uv)[J]\cdot(uv)[I]\cdot\mathfrak{z}.
\end{aligned}$$

In other words, $\mathfrak{a}\cdot(uv)[J]\cdot(uv)[I]\cdot\mathfrak{z}\leq\mathfrak{a}\cdot(uv)[I]\cdot(uv)[J]\cdot\mathfrak{z}$. Hence, Proposition 6.1.2(c) (applied to $\mathfrak{a}$, $(uv)[J]\cdot(uv)[I]\cdot\mathfrak{z}$ and $(uv)[I]\cdot(uv)[J]\cdot\mathfrak{z}$ instead of a, c and d) yields $(uv)[J]\cdot(uv)[I]\cdot\mathfrak{z}\leq(uv)[I]\cdot(uv)[J]\cdot\mathfrak{z}$. Thus, Proposition 6.1.2(d) (applied to $(uv)[J]\cdot(uv)[I]$, $\mathfrak{z}$, $(uv)[I]\cdot(uv)[J]$ and $\mathfrak{z}$ instead of a, b, c and d) yields $(uv)[J]\cdot(uv)[I]\leq(uv)[I]\cdot(uv)[J]$ (since $\ell((uv)[J]\cdot(uv)[I])=\ell((uv)[J])+\ell((uv)[I])=\ell((uv)[I])+\ell((uv)[J])=\ell((uv)[I]\cdot(uv)[J])$). In other words, $(uv)[I]\cdot(uv)[J]\geq(uv)[J]\cdot(uv)[I]$. This proves Lemma 6.2.10(b).

(c) The proof of Lemma 6.2.10(c) is analogous to that of Lemma 6.2.10(b). □

13.156. **Solution to Exercise 6.2.15.** *Solution to Exercise 6.2.15.*

Proof of Proposition 6.2.14. (a) Let $j\in\{1,2,\ldots,\ell\}$. We have $\sigma=\operatorname{iper}(\alpha,\tau)=\overrightarrow{I_{\tau(1)}}\overrightarrow{I_{\tau(2)}}\cdots\overrightarrow{I_{\tau(\ell)}}$ (in one-line notation). Therefore, the word $\overrightarrow{I_{\tau(j)}}$ appears as a factor in this word σ, starting at position $\sum_{k=1}^{j-1}\alpha_{\tau(k)}+1$ and ending at position $\sum_{k=1}^{j}\alpha_{\tau(k)}$. In other words, the letters of the word $\overrightarrow{I_{\tau(j)}}$ are the letters σ_j of σ for $j\in\left[\sum_{k=1}^{j-1}\alpha_{\tau(k)}:\sum_{k=1}^{j}\alpha_{\tau(k)}\right]^+$. Since the letters of the word $\overrightarrow{I_{\tau(j)}}$ are precisely the elements of $I_{\tau(j)}$, this rewrites as follows: The elements of $I_{\tau(j)}$ are the letters σ_j of σ for $j\in\left[\sum_{k=1}^{j-1}\alpha_{\tau(k)}:\sum_{k=1}^{j}\alpha_{\tau(k)}\right]^+$. In other words,

$$I_{\tau(j)}=\left\{\sigma_j\ \mid\ j\in\left[\sum_{k=1}^{j-1}\alpha_{\tau(k)}:\sum_{k=1}^{j}\alpha_{\tau(k)}\right]^+\right\}=\sigma\left(\left[\sum_{k=1}^{j-1}\alpha_{\tau(k)}:\sum_{k=1}^{j}\alpha_{\tau(k)}\right]^+\right).$$

Hence, $\sigma^{-1}\left(I_{\tau(j)}\right)=\left[\sum_{k=1}^{j-1}\alpha_{\tau(k)}:\sum_{k=1}^{j}\alpha_{\tau(k)}\right]^+$, so that Proposition 6.2.14(a) is proven.

(b) Let $j\in\{1,2,\ldots,\ell\}$. We need to show that the restriction of the map σ^{-1} to the interval $I_{\tau(j)}$ is increasing. In other words, we need to prove that the elements of $I_{\tau(j)}$ occur in increasing order in the word σ. But this is clear, because these elements all occur in the factor $\overrightarrow{I_{\tau(j)}}$ of the word σ, and this factor has them in increasing order (by its definition). This proves Proposition 6.2.14(b).

(c) Let $i\in\{1,2,\ldots,\ell\}$. Then, $\sigma^{-1}(I_i)=\sigma^{-1}\left(I_{\tau(\tau^{-1}(i))}\right)=\left[\sum_{k=1}^{\tau^{-1}(i)-1}\alpha_{\tau(k)},\sum_{k=1}^{\tau^{-1}(i)}\alpha_{\tau(k)}\right]^+$ (according to Proposition 6.2.14(a), applied to $j=\tau^{-1}(i)$). Hence, $\sigma^{-1}(I_i)$ is an interval. Furthermore, the restriction of the map σ^{-1} to the interval $I_i=I_{\tau(\tau^{-1}(i))}$ is increasing (according to Proposition 6.2.14(b), applied to $j=\tau^{-1}(i)$).

Now, forget that we fixed i. We thus have proven that every $i\in\{1,2,\ldots,\ell\}$ has the two properties that:

- the set $\sigma^{-1}(I_i)$ is an interval;
- the restriction of the map σ^{-1} to the interval I_i is increasing.

In other words, the permutation σ is α-clumping. Since $\sigma=\operatorname{iper}(\alpha,\tau)$, this shows that $\operatorname{iper}(\alpha,\tau)$ is α-clumping. Proposition 6.2.14(c) is proven.

(d) Let $i \in \{1, 2, \ldots, \ell - 1\}$. Then,

$$\sigma^{-1}\left(I_{\tau(i)}\right) = \left[\sum_{k=1}^{i-1} \alpha_{\tau(k)} : \sum_{k=1}^{i} \alpha_{\tau(k)}\right]^+ \tag{13.156.1}$$

(by Proposition 6.2.14(a), applied to $j = i$), so that $\sigma^{-1}\left(I_{\tau(i)}\right)$ is an interval. Thus, $\sigma^{-1}\left(I_{\tau(i)}\right)$ is a nonempty interval (nonempty because $I_{\tau(i)}$ is nonempty). Similarly, $\sigma^{-1}\left(I_{\tau(i+1)}\right)$ is a nonempty interval.

Also, Proposition 6.2.14(a) (applied to $j = i + 1$) yields

$$\sigma^{-1}\left(I_{\tau(i+1)}\right) = \left[\sum_{k=1}^{i} \alpha_{\tau(k)} : \sum_{k=1}^{i+1} \alpha_{\tau(k)}\right]^+ . \tag{13.156.2}$$

Now,

$$\sigma^{-1}\left(I_{\tau(i)}\right) = \left[\sum_{k=1}^{i-1} \alpha_{\tau(k)} : \sum_{k=1}^{i} \alpha_{\tau(k)}\right]^+ < \left[\sum_{k=1}^{i} \alpha_{\tau(k)} : \sum_{k=1}^{i+1} \alpha_{\tau(k)}\right]^+ = \sigma^{-1}\left(I_{\tau(i+1)}\right).$$

Also,

$$\underbrace{\sigma^{-1}\left(I_{\tau(i)}\right)}_{\substack{=\left[\sum_{k=1}^{i-1} \alpha_{\tau(k)} : \sum_{k=1}^{i} \alpha_{\tau(k)}\right]^+ \\ \text{(by (13.156.1))}}} \cup \underbrace{\sigma^{-1}\left(I_{\tau(i+1)}\right)}_{\substack{=\left[\sum_{k=1}^{i} \alpha_{\tau(k)} : \sum_{k=1}^{i+1} \alpha_{\tau(k)}\right]^+ \\ \text{(by (13.156.2))}}}$$
$$= \left[\sum_{k=1}^{i-1} \alpha_{\tau(k)} : \sum_{k=1}^{i} \alpha_{\tau(k)}\right]^+ \cup \left[\sum_{k=1}^{i} \alpha_{\tau(k)} : \sum_{k=1}^{i+1} \alpha_{\tau(k)}\right]^+ = \left[\sum_{k=1}^{i-1} \alpha_{\tau(k)} : \sum_{k=1}^{i+1} \alpha_{\tau(k)}\right]^+ ,$$

which is obviously an interval. Proposition 6.2.14(d) is proven. □

13.157. **Solution to Exercise 6.2.17.** *Solution to Exercise 6.2.17.*

Proof of Proposition 6.2.16. We shall use Definition 13.126.3 and Proposition 13.126.4.

(a) The interval system corresponding to α is an ℓ-tuple of intervals (since $\ell(\alpha) = \ell$); denote this ℓ-tuple by $(I_1, I_2, \ldots, I_\ell)$. Then, the intervals $I_1, I_2, \ldots, I_\ell$ form a set partition of $[0 : n]^+$ (according to Remark 6.2.6(c)) and are nonempty (also according to Remark 6.2.6(c)).

We define a map $\operatorname{iper}'_\alpha : \{\omega \in \mathfrak{S}_n \mid \omega \text{ is } \alpha\text{-clumping}\} \to \mathfrak{S}_\ell$ as follows: Let ω be an α-clumping element of $\mathfrak{S}_n$. Then, every $i \in \{1, 2, \ldots, \ell\}$ has the property that $\omega^{-1}(I_i)$ is an interval (since ω is α-clumping). These intervals $\omega^{-1}(I_1)$, $\omega^{-1}(I_2)$, $\ldots$, $\omega^{-1}(I_\ell)$ form a set partition of $[0 : n]^+$ (since the intervals I_1, I_2, $\ldots$, I_ℓ form a set partition of $[0 : n]^+$), and thus are disjoint (and nonempty[1093]). Hence, these intervals form a totally ordered set with respect to the relation $<$ (by Remark 6.2.6(b)). Thus, there exists a unique permutation $\tau \in \mathfrak{S}_\ell$ such that $\omega^{-1}\left(I_{\tau(1)}\right) < \omega^{-1}\left(I_{\tau(2)}\right) < \cdots < \omega^{-1}\left(I_{\tau(\ell)}\right)$. We define $\operatorname{iper}'_\alpha(\omega)$ to be this permutation τ.

[1093] Their nonemptiness follows from the fact that the intervals I_1, I_2, $\ldots$, I_ℓ are nonempty.

Thus, we have defined a map $\operatorname{iper}'_\alpha : \{\omega \in \mathfrak{S}_n \mid \omega \text{ is } \alpha\text{-clumping}\} \to \mathfrak{S}_\ell$. It is easy to see that $\operatorname{iper}_\alpha \circ \operatorname{iper}'_\alpha = \operatorname{id}$ [1094] and that $\operatorname{iper}'_\alpha \circ \operatorname{iper}_\alpha = \operatorname{id}$ [1095]. Hence, the maps $\operatorname{iper}_\alpha$ and $\operatorname{iper}'_\alpha$ are mutually inverse, and thus the map $\operatorname{iper}_\alpha$ is bijective. This completes the proof of Proposition 6.2.16(a).

[1094] *Proof.* Let $\omega \in \mathfrak{S}_n$ be α-clumping. We are going to prove that $\left(\operatorname{iper}_\alpha \circ \operatorname{iper}'_\alpha\right)(\omega) = \operatorname{id}(\omega)$.

Indeed, let us recall how $\operatorname{iper}'_\alpha(\omega)$ was defined: Every $i \in \{1, 2, \ldots, \ell\}$ has the property that $\omega^{-1}(I_i)$ is an interval. These intervals $\omega^{-1}(I_1)$, $\omega^{-1}(I_2)$, $\ldots$, $\omega^{-1}(I_\ell)$ form a totally ordered set with respect to the relation $<$. Then, $\operatorname{iper}'_\alpha(\omega)$ is defined as the unique permutation $\tau \in \mathfrak{S}_\ell$ such that $\omega^{-1}\left(I_{\tau(1)}\right) < \omega^{-1}\left(I_{\tau(2)}\right) < \cdots < \omega^{-1}\left(I_{\tau(\ell)}\right)$. Thus, we have $\operatorname{iper}'_\alpha(\omega) \in \mathfrak{S}_\ell$ and

$$\omega^{-1}\left(I_{\left(\operatorname{iper}'_\alpha(\omega)\right)(1)}\right) < \omega^{-1}\left(I_{\left(\operatorname{iper}'_\alpha(\omega)\right)(2)}\right) < \cdots < \omega^{-1}\left(I_{\left(\operatorname{iper}'_\alpha(\omega)\right)(\ell)}\right). \tag{13.157.1}$$

Denote the permutation $\operatorname{iper}'_\alpha(\omega) \in \mathfrak{S}_\ell$ by τ. Then, (13.157.1) rewrites as

$$\omega^{-1}\left(I_{\tau(1)}\right) < \omega^{-1}\left(I_{\tau(2)}\right) < \cdots < \omega^{-1}\left(I_{\tau(\ell)}\right) \tag{13.157.2}$$

(since $\tau = \operatorname{iper}'_\alpha(\omega)$). The definition of $\operatorname{iper}(\alpha, \tau)$ shows that $\operatorname{iper}(\alpha, \tau)$ is the permutation in $\mathfrak{S}_n$ which (in one-line notation) is the word $\overrightarrow{I_{\tau(1)}}\overrightarrow{I_{\tau(2)}}\cdots\overrightarrow{I_{\tau(\ell)}}$ (a concatenation of ℓ words). In other words, $\operatorname{iper}(\alpha, \tau) = \overrightarrow{I_{\tau(1)}}\overrightarrow{I_{\tau(2)}}\cdots\overrightarrow{I_{\tau(\ell)}}$. Denote this permutation $\operatorname{iper}(\alpha, \tau)$ by η. Then, $\eta = \operatorname{iper}(\alpha, \tau) = \overrightarrow{I_{\tau(1)}}\overrightarrow{I_{\tau(2)}}\cdots\overrightarrow{I_{\tau(\ell)}}$ and

$$\begin{aligned}\left(\operatorname{iper}_\alpha \circ \operatorname{iper}'_\alpha\right)(\omega) &= \operatorname{iper}_\alpha\left(\underbrace{\operatorname{iper}'_\alpha(\omega)}_{=\tau}\right) = \operatorname{iper}_\alpha \tau = \operatorname{iper}(\alpha, \tau) \qquad (\text{by the definition of } \operatorname{iper}_\alpha \tau) \\ &= \eta. \end{aligned} \tag{13.157.3}$$

Our next goal is to prove that $\eta = \omega$.

We know that ω is α-clumping. In other words, every $i \in \{1, 2, \ldots, \ell\}$ has the two properties that:

$$\text{the set } \omega^{-1}(I_i) \text{ is an interval,} \tag{13.157.4}$$

and

$$\text{the restriction of the map } \omega^{-1} \text{ to the interval } I_i \text{ is increasing} \tag{13.157.5}$$

(by the definition of "α-clumping").

We shall now prove that $\operatorname{Inv}\left(\eta^{-1}\right) \subset \operatorname{Inv}\left(\omega^{-1}\right)$.

Indeed, let (i, j) be some element of $\operatorname{Inv}\left(\eta^{-1}\right)$. We want to prove that $(i, j) \in \operatorname{Inv}\left(\omega^{-1}\right)$.

We have $(i, j) \in \operatorname{Inv}\left(\eta^{-1}\right)$. By the definition of $\operatorname{Inv}\left(\eta^{-1}\right)$, this yields that (i, j) is an element of $\{1, 2, \ldots, n\}^2$ satisfying $i < j$ and $\eta^{-1}(i) > \eta^{-1}(j)$. In particular, $\eta^{-1}(i) > \eta^{-1}(j)$. In other words, the letter i must appear after the letter j in the word η (where our use of the word "after" does not imply "immediately after"). Since $\eta = \overrightarrow{I_{\tau(1)}}\overrightarrow{I_{\tau(2)}}\cdots\overrightarrow{I_{\tau(\ell)}}$, this rewrites as follows: The letter i must appear after the letter j in the word $\overrightarrow{I_{\tau(1)}}\overrightarrow{I_{\tau(2)}}\cdots\overrightarrow{I_{\tau(\ell)}}$. Thus, we must be in one of the following two cases:

Case 1: There exist some elements $\mathbf{i}$ and $\mathbf{j}$ of $\{1, 2, \ldots, \ell\}$ such that $\mathbf{i} > \mathbf{j}$ and such that the letter i appears in the word $\overrightarrow{I_{\tau(\mathbf{i})}}$, whereas the letter j appears in the word $\overrightarrow{I_{\tau(\mathbf{j})}}$.

Case 2: There exists some element $\mathbf{i}$ of $\{1, 2, \ldots, \ell\}$ such that both letters i and j appear in the word $\overrightarrow{I_{\tau(\mathbf{i})}}$, and the letter i appears after the letter j in this word.

Let us first consider Case 1. In this case, there exist some elements $\mathbf{i}$ and $\mathbf{j}$ of $\{1, 2, \ldots, \ell\}$ such that $\mathbf{i} > \mathbf{j}$ and such that the letter i appears in the word $\overrightarrow{I_{\tau(\mathbf{i})}}$, whereas the letter j appears in the word $\overrightarrow{I_{\tau(\mathbf{j})}}$. Consider these $\mathbf{i}$ and $\mathbf{j}$. The letter i appears in the word $\overrightarrow{I_{\tau(\mathbf{i})}}$, and thus is an element of $I_{\tau(\mathbf{i})}$ (since the letters of the word $\overrightarrow{I_{\tau(\mathbf{i})}}$ are precisely the elements of $I_{\tau(\mathbf{i})}$). In other words, $i \in I_{\tau(\mathbf{i})}$. Hence, $\omega^{-1}(i) \in \omega^{-1}\left(I_{\tau(\mathbf{i})}\right)$. Similarly, $\omega^{-1}(j) \in \omega^{-1}\left(I_{\tau(\mathbf{j})}\right)$. Now, $\mathbf{i} > \mathbf{j}$, so that $\mathbf{j} < \mathbf{i}$ and thus $\omega^{-1}\left(I_{\tau(\mathbf{j})}\right) < \omega^{-1}\left(I_{\tau(\mathbf{i})}\right)$ (due to (13.157.2)). In other words, every $j' \in \omega^{-1}\left(I_{\tau(\mathbf{j})}\right)$ and $i' \in \omega^{-1}\left(I_{\tau(\mathbf{i})}\right)$ satisfy $j' < i'$ (by the definition of $\omega^{-1}\left(I_{\tau(\mathbf{j})}\right) < \omega^{-1}\left(I_{\tau(\mathbf{i})}\right)$). Applying this to $j' = \omega^{-1}(j)$ and $i' = \omega^{-1}(i)$, we obtain $\omega^{-1}(j) < \omega^{-1}(i)$, so that $\omega^{-1}(i) > \omega^{-1}(j)$.

Now, we know that (i, j) is an element of $\{1, 2, \ldots, n\}^2$ satisfying $i < j$ and $\omega^{-1}(i) > \omega^{-1}(j)$. In other words, $(i, j) \in \operatorname{Inv}\left(\omega^{-1}\right)$ (by the definition of $\operatorname{Inv}\left(\omega^{-1}\right)$). Hence, $(i, j) \in \operatorname{Inv}\left(\omega^{-1}\right)$ is proven in Case 1.

Let us now consider Case 2. In this case, there exists some element $\mathbf{i}$ of $\{1, 2, \ldots, \ell\}$ such that both letters i and j appear in the word $\overrightarrow{I_{\tau(\mathbf{i})}}$, and the letter i appears after the letter j in this word. Consider this i. The letters of the word $\overrightarrow{I_{\tau(\mathbf{i})}}$ are in increasing order (since $\overrightarrow{I_{\tau(\mathbf{i})}}$ is defined as the list of all elements of $I_{\tau(\mathbf{i})}$ in increasing order). Since the letter i appears after the letter j in this word, we must therefore have $i > j$. But this contradicts $i < j$. This contradiction shows that Case 2 cannot occur. Hence, the only possible case is Case 1. Since we have proven $(i, j) \in \operatorname{Inv}\left(\omega^{-1}\right)$ in this case, we therefore conclude that $(i, j) \in \operatorname{Inv}\left(\omega^{-1}\right)$ always holds.

Now, let us forget that we fixed (i, j). We thus have proven that $(i, j) \in \operatorname{Inv}\left(\omega^{-1}\right)$ for every $(i, j) \in \operatorname{Inv}\left(\eta^{-1}\right)$. In other words,

$$\operatorname{Inv}\left(\eta^{-1}\right) \subset \operatorname{Inv}\left(\omega^{-1}\right). \tag{13.157.6}$$

(b) Consider the map $\operatorname{iper}_\alpha$ defined in Proposition 6.2.16(a). Since σ is α-clumping, we have $\sigma \in \{\omega \in \mathfrak{S}_n \mid \omega \text{ is } \alpha\text{-clumping}\}$. In other words, σ belongs to the target of the map $\operatorname{iper}_\alpha$. Thus, σ has a unique preimage under the map $\operatorname{iper}_\alpha$ (since the map $\operatorname{iper}_\alpha$ is bijective (by Proposition 6.2.16(a))). In other words, there exists a unique $\tau \in \mathfrak{S}_\ell$ satisfying $\sigma = \operatorname{iper}_\alpha \tau$. Since $\operatorname{iper}_\alpha \tau = \operatorname{iper}(\alpha, \tau)$ for every $\tau \in \mathfrak{S}_{p+q}$

Next, let us prove that $\operatorname{Inv}\left(\omega^{-1}\right) \subset \operatorname{Inv}\left(\eta^{-1}\right)$.

Indeed, let (i, j) be some element of $\operatorname{Inv}\left(\omega^{-1}\right)$. We want to prove that $(i, j) \in \operatorname{Inv}\left(\eta^{-1}\right)$.

We have $(i, j) \in \operatorname{Inv}\left(\omega^{-1}\right)$. By the definition of $\operatorname{Inv}\left(\omega^{-1}\right)$, this yields that (i, j) is an element of $\{1, 2, \ldots, n\}^2$ satisfying $i < j$ and $\omega^{-1}(i) > \omega^{-1}(j)$.

The intervals $I_1, I_2, \ldots, I_\ell$ form a set partition of $[0 : n]^+$ (according to Remark 6.2.6(c)). Hence, there exists some $u \in \{1, 2, \ldots, \ell\}$ such that $i \in I_u$ (because $i \in \{1, 2, \ldots, n\} = [0 : n]^+$). Similarly, there exists some $v \in \{1, 2, \ldots, \ell\}$ such that $j \in I_v$. Consider these u and v.

First, let us assume (for the sake of contradiction) that $u = v$. Then, $i \in I_u = I_v$ (since $u = v$) and $j \in I_v$. But (13.157.5) (applied to v instead of i) yields that the restriction of the map ω^{-1} to the interval I_v is increasing. Hence, $\omega^{-1}(i) \leq \omega^{-1}(j)$ (since the elements i and j both lie in the interval I_v and satisfy $i < j$), which contradicts $\omega^{-1}(i) > \omega^{-1}(j)$. This contradiction shows that our assumption (that $u = v$) was wrong. Hence, we have $u \neq v$. Thus, $\tau^{-1}(u) \neq \tau^{-1}(v)$ (since τ is a permutation).

Define two elements $\mathbf{i}$ and $\mathbf{j}$ of $\{1, 2, \ldots, \ell\}$ by $\mathbf{i} = \tau^{-1}(u)$ and $\mathbf{j} = \tau^{-1}(v)$. Then, $\tau(\mathbf{i}) = u$ (since $\mathbf{i} = \tau^{-1}(u)$), so that $I_{\tau(\mathbf{i})} = I_u$ and thus $i \in I_u = I_{\tau(\mathbf{i})}$ (since $I_{\tau(\mathbf{i})} = I_u$). Similarly, $j \in I_{\tau(\mathbf{j})}$. We have $\mathbf{i} = \tau^{-1}(u) \neq \tau^{-1}(v) = \mathbf{j}$.

Now, let us assume (for the sake of contradiction) that $\mathbf{i} \leq \mathbf{j}$. Combined with $\mathbf{i} \neq \mathbf{j}$, this yields $\mathbf{i} < \mathbf{j}$. Hence, $\omega^{-1}\left(I_{\tau(\mathbf{i})}\right) < \omega^{-1}\left(I_{\tau(\mathbf{j})}\right)$ (due to (13.157.2)). Thus, every $i' \in \omega^{-1}\left(I_{\tau(\mathbf{i})}\right)$ and $j' \in \omega^{-1}\left(I_{\tau(\mathbf{j})}\right)$ satisfy $i' < j'$ (by the definition of $\omega^{-1}\left(I_{\tau(\mathbf{i})}\right) < \omega^{-1}\left(I_{\tau(\mathbf{j})}\right)$). Applying this to $i' = \omega^{-1}(i)$ and $j' = \omega^{-1}(j)$, we obtain $\omega^{-1}(i) < \omega^{-1}(j)$ (since $\omega^{-1}\left(\underbrace{i}_{\in I_{\tau(\mathbf{i})}}\right) \in \omega^{-1}\left(I_{\tau(\mathbf{i})}\right)$ and $\omega^{-1}\left(\underbrace{j}_{\in I_{\tau(\mathbf{j})}}\right) \in \omega^{-1}\left(I_{\tau(\mathbf{j})}\right)$), which contradicts $\omega^{-1}(i) > \omega^{-1}(j)$. This contradiction proves that our assumption (that $\mathbf{i} \leq \mathbf{j}$) was wrong. Hence, we have $\mathbf{i} > \mathbf{j}$.

Now, recall that the word $\overrightarrow{I_{\tau(\mathbf{i})}}$ is defined as the list of all elements of $I_{\tau(\mathbf{i})}$ in increasing order. Hence, i is a letter of the word $\overrightarrow{I_{\tau(\mathbf{i})}}$ (since i is an element of $I_{\tau(\mathbf{i})}$). Similarly, j is a letter of the word $\overrightarrow{I_{\tau(\mathbf{j})}}$. Both words $\overrightarrow{I_{\tau(\mathbf{i})}}$ and $\overrightarrow{I_{\tau(\mathbf{j})}}$ are factors of the concatenation $\overrightarrow{I_{\tau(1)}}\,\overrightarrow{I_{\tau(2)}} \cdots \overrightarrow{I_{\tau(\ell)}}$, with the factor $\overrightarrow{I_{\tau(\mathbf{i})}}$ appearing after the factor $\overrightarrow{I_{\tau(\mathbf{j})}}$ (since $\mathbf{i} > \mathbf{j}$). Thus, if i_0 is any letter of the word $\overrightarrow{I_{\tau(\mathbf{i})}}$, and if j_0 is any letter of the word $\overrightarrow{I_{\tau(\mathbf{j})}}$, then the letter i_0 appears after the letter j_0 in the concatenation $\overrightarrow{I_{\tau(1)}}\,\overrightarrow{I_{\tau(2)}} \cdots \overrightarrow{I_{\tau(\ell)}}$. Applying this to $i_0 = i$ and $j_0 = j$, we conclude that the letter i appears after the letter j in the concatenation $\overrightarrow{I_{\tau(1)}}\,\overrightarrow{I_{\tau(2)}} \cdots \overrightarrow{I_{\tau(\ell)}}$ (since i is a letter of the word $\overrightarrow{I_{\tau(\mathbf{i})}}$, and since j is a letter of the word $\overrightarrow{I_{\tau(\mathbf{j})}}$). In other words, the letter i appears after the letter j in the word η (since $\eta = \overrightarrow{I_{\tau(1)}}\,\overrightarrow{I_{\tau(2)}} \cdots \overrightarrow{I_{\tau(\ell)}}$). In other words, $\eta^{-1}(i) > \eta^{-1}(j)$.

So (i, j) is an element of $\{1, 2, \ldots, n\}^2$ satisfying $i < j$ and $\eta^{-1}(i) > \eta^{-1}(j)$. In other words, $(i, j) \in \operatorname{Inv}\left(\eta^{-1}\right)$ (by the definition of $\operatorname{Inv}\left(\eta^{-1}\right)$).

Now, let us forget that we fixed (i, j). We thus have proven that $(i, j) \in \operatorname{Inv}\left(\eta^{-1}\right)$ for every $(i, j) \in \operatorname{Inv}\left(\omega^{-1}\right)$. In other words,

$$\operatorname{Inv}\left(\omega^{-1}\right) \subset \operatorname{Inv}\left(\eta^{-1}\right).$$

Combined with (13.157.6), this yields $\operatorname{Inv}\left(\omega^{-1}\right) = \operatorname{Inv}\left(\eta^{-1}\right)$. Thus, Proposition 13.126.4 (applied to $\varphi = \omega^{-1}$ and $\psi = \eta^{-1}$) yields $\omega^{-1} = \eta^{-1}$, whence $\omega = \eta = \left(\operatorname{iper}_\alpha \circ \operatorname{iper}'_\alpha\right)(\omega)$ (by (13.157.3)), so that $\left(\operatorname{iper}_\alpha \circ \operatorname{iper}'_\alpha\right)(\omega) = \omega = \operatorname{id}(\omega)$.

Now, let us forget that we fixed ω. We thus have proven that $\left(\operatorname{iper}_\alpha \circ \operatorname{iper}'_\alpha\right)(\omega) = \operatorname{id}(\omega)$ for every α-clumping permutation $\omega \in \mathfrak{S}_n$. In other words, $\operatorname{iper}_\alpha \circ \operatorname{iper}'_\alpha = \operatorname{id}$, qed.

[1095] *Proof.* Let $\pi \in \mathfrak{S}_\ell$. We shall show that that $\left(\operatorname{iper}'_\alpha \circ \operatorname{iper}_\alpha\right)(\pi) = \operatorname{id}(\pi)$.

Indeed, let $\omega = \operatorname{iper}_\alpha \pi$. Then, $\omega = \operatorname{iper}_\alpha \pi = \operatorname{iper}(\alpha, \pi)$ (by the definition of $\operatorname{iper}_\alpha \pi$). The permutation $\omega = \operatorname{iper}(\alpha, \pi)$ is α-clumping (according to Proposition 6.2.14(c), applied to $\tau = \pi$).

Let us now recall how $\operatorname{iper}'_\alpha(\omega)$ was defined: Every $i \in \{1, 2, \ldots, \ell\}$ has the property that $\omega^{-1}(I_i)$ is an interval. These intervals $\omega^{-1}(I_1), \omega^{-1}(I_2), \ldots, \omega^{-1}(I_\ell)$ form a totally ordered set with respect to the relation $<$. Then, $\operatorname{iper}'_\alpha(\omega)$ is defined as the unique permutation $\tau \in \mathfrak{S}_\ell$ such that $\omega^{-1}\left(I_{\tau(1)}\right) < \omega^{-1}\left(I_{\tau(2)}\right) < \cdots < \omega^{-1}\left(I_{\tau(\ell)}\right)$. Hence, if $\tau \in \mathfrak{S}_\ell$ is a permutation satisfying $\omega^{-1}\left(I_{\tau(1)}\right) < \omega^{-1}\left(I_{\tau(2)}\right) < \cdots < \omega^{-1}\left(I_{\tau(\ell)}\right)$, then

$$\tau = \operatorname{iper}'_\alpha(\omega). \tag{13.157.7}$$

We shall now use this to prove that $\pi = \operatorname{iper}'_\alpha(\omega)$.

We have $\omega = \operatorname{iper}(\alpha, \pi) = \overrightarrow{I_{\pi(1)}}\,\overrightarrow{I_{\pi(2)}} \cdots \overrightarrow{I_{\pi(\ell)}}$ (in one-line notation), according to the definition of $\operatorname{iper}(\alpha, \pi)$.

The intervals $\omega^{-1}(I_1), \omega^{-1}(I_2), \ldots, \omega^{-1}(I_\ell)$ form a set partition of $[0 : n]^+$ (since the intervals $I_1, I_2, \ldots, I_\ell$ form a set partition of $[0 : n]^+$), and thus are disjoint. These intervals $\omega^{-1}(I_1), \omega^{-1}(I_2), \ldots, \omega^{-1}(I_\ell)$ are also nonempty (since the intervals $I_1, I_2, \ldots, I_\ell$ are nonempty).

Now, let $u \in \{1, 2, \ldots, \ell - 1\}$. The intervals $\omega^{-1}\left(I_{\pi(u)}\right)$ and $\omega^{-1}\left(I_{\pi(u+1)}\right)$ are nonempty (since the intervals $\omega^{-1}(I_1), \omega^{-1}(I_2), \ldots, \omega^{-1}(I_\ell)$ are nonempty). We will show that $\omega^{-1}\left(I_{\pi(u)}\right) < \omega^{-1}\left(I_{\pi(u+1)}\right)$.

Indeed, let $i \in \omega^{-1}\left(I_{\pi(u)}\right)$ and $j \in \omega^{-1}\left(I_{\pi(u+1)}\right)$ be arbitrary. We shall prove that $i < j$.

(by the definition of $\operatorname{iper}_\alpha$), this rewrites as follows: There exists a unique $\tau \in \mathfrak{S}_\ell$ satisfying $\sigma = \operatorname{iper}(\alpha, \tau)$. This proves Proposition 6.2.16(b). □

13.158. **Solution to Exercise 6.2.19.** *Solution to Exercise 6.2.19.*

Proof of Proposition 6.2.18. We know that $\alpha\beta$ is a composition of $n+m$ having length $\ell(\alpha\beta) = \ell(\alpha) + \ell(\beta) = p+q$. Hence, the interval system corresponding to $\alpha\beta$ is a $(p+q)$-tuple of intervals which covers $[0:n+m]^+$. Denote this $(p+q)$-tuple by $(I_1, I_2, \ldots, I_{p+q})$. It is clear that $I_1 \cup I_2 \cup \cdots \cup I_p = [0:n]^+$ and $I_{p+1} \cup I_{p+2} \cup \cdots \cup I_{p+q} = [n:n+m]^+$ (since the first p parts of the composition $\alpha\beta$ form the composition α of n). Moreover, $I_1 < I_2 < \cdots < I_{p+q}$ (since $(I_1, I_2, \ldots, I_{p+q})$ is the interval system corresponding to $\alpha\beta$).

The definition of $\operatorname{iper}(\alpha\beta, \tau)$ yields that $\operatorname{iper}(\alpha\beta,\tau) = \overrightarrow{I_{\tau(1)}}\,\overrightarrow{I_{\tau(2)}} \cdots \overrightarrow{I_{\tau(p+q)}}$ (in one-line notation). Denote the permutation $\operatorname{iper}(\alpha\beta,\tau)$ by ω; then, this becomes $\omega = \overrightarrow{I_{\tau(1)}}\,\overrightarrow{I_{\tau(2)}} \cdots \overrightarrow{I_{\tau(p+q)}}$. For every $j \in \{1, 2, \ldots, p+q\}$, the restriction of the map ω^{-1} to the interval $I_{\tau(j)}$ is increasing (by Proposition 6.2.14(b), applied to $n+m$, $\alpha\beta$, $p+q$ and ω instead of n, α, ℓ and σ). Substituting k for $\tau(j)$ here, we obtain: For every $k \in \{1, 2, \ldots, p+q\}$, the restriction of the map ω^{-1} to the interval I_k is increasing. Of course, this yields that for every $k \in \{1, 2, \ldots, p+q\}$, the restriction of the map ω^{-1} to the interval I_k is strictly increasing (because ω^{-1} is injective).

Now, we need to prove that $\tau \in \operatorname{Sh}_{p,q}$ if and only if $\operatorname{iper}(\alpha\beta,\tau) \in \operatorname{Sh}_{n,m}$. In other words, we need to prove that $\tau \in \operatorname{Sh}_{p,q}$ if and only if $\omega \in \operatorname{Sh}_{n,m}$ (since $\omega = \operatorname{iper}(\alpha\beta,\tau)$). We shall prove the $\Longrightarrow$ and $\Longleftarrow$ directions of this statement separately:

$\Longrightarrow$: Assume that $\tau \in \operatorname{Sh}_{p,q}$. We need to show that $\omega \in \operatorname{Sh}_{n,m}$.

We have $\tau \in \operatorname{Sh}_{p,q}$. Thus, $\tau^{-1}(1) < \tau^{-1}(2) < \cdots < \tau^{-1}(p)$ and $\tau^{-1}(p+1) < \tau^{-1}(p+2) < \cdots < \tau^{-1}(p+q)$.

Now, let i and j be two elements of $\{1, 2, \ldots, n\}$ such that $i < j$. We are going to prove that $\omega^{-1}(i) < \omega^{-1}(j)$.

Indeed, we have $i \in \{1, 2, \ldots, n\} = [0:n]^+ = I_1 \cup I_2 \cup \cdots \cup I_p$, so that $i \in I_{i'}$ for some $i' \in \{1, 2, \ldots, p\}$. Similarly, $j \in I_{j'}$ for some $j' \in \{1, 2, \ldots, p\}$. Consider these i' and j'. We cannot have $j' < i'$ (because this would entail $I_{j'} < I_{i'}$ (due to $I_1 < I_2 < \cdots < I_{p+q}$), which would lead to $j < i$ (since $j \in I_{j'}$ and $i \in I_{i'}$), contradicting $i < j$). Hence, we must have either $j' = i'$ or $j' > i'$. If $j' = i'$, then i and j lie in one and the same I_k (namely, the one with $k = j' = i'$), and so $\omega^{-1}(i) < \omega^{-1}(j)$ follows from the fact that the map ω^{-1} restricted to I_k is strictly increasing (and from the inequality $i < j$). Hence, it only remains to consider the case when $j' > i'$. Assume WLOG that we are in this case. Then, $i' < j'$, and so $\tau^{-1}(i') < \tau^{-1}(j')$ (since $\tau^{-1}(1) < \tau^{-1}(2) < \cdots < \tau^{-1}(p)$ and since i' and j' belong to $\{1, 2, \ldots, p\}$), and thus the word $\overrightarrow{I_{i'}}$ appears before the word $\overrightarrow{I_{j'}}$ in the concatenation $\overrightarrow{I_{\tau(1)}}\,\overrightarrow{I_{\tau(2)}} \cdots \overrightarrow{I_{\tau(p+q)}}$. Hence, the letter i appears before the letter

---

We have $i \in \omega^{-1}\left(I_{\pi(u)}\right)$, so that $\omega(i) \in I_{\pi(u)}$. Hence, $\omega(i)$ is a letter of the word $\overrightarrow{I_{\pi(u)}}$ (since the word $\overrightarrow{I_{\pi(u)}}$ is defined as the list of all elements of $I_{\pi(u)}$ in increasing order). Similarly, $\omega(j)$ is a letter of the word $\overrightarrow{I_{\pi(u+1)}}$ (since $j \in \omega^{-1}\left(I_{\pi(u+1)}\right)$).

Both words $\overrightarrow{I_{\pi(u)}}$ and $\overrightarrow{I_{\pi(u+1)}}$ are factors of the concatenation $\overrightarrow{I_{\pi(1)}}\,\overrightarrow{I_{\pi(2)}} \cdots \overrightarrow{I_{\pi(\ell)}}$, with the factor $\overrightarrow{I_{\pi(u+1)}}$ appearing after the factor $\overrightarrow{I_{\pi(u)}}$. Thus, if i_0 is any letter of the word $\overrightarrow{I_{\pi(u)}}$, and if j_0 is any letter of the word $\overrightarrow{I_{\pi(u+1)}}$, then the letter j_0 appears after the letter i_0 in the concatenation $\overrightarrow{I_{\pi(1)}}\,\overrightarrow{I_{\pi(2)}} \cdots \overrightarrow{I_{\pi(\ell)}}$. Applying this to $i_0 = \omega(i)$ and $j_0 = \omega(j)$, we conclude that the letter $\omega(j)$ appears after the letter $\omega(i)$ in the concatenation $\overrightarrow{I_{\pi(1)}}\,\overrightarrow{I_{\pi(2)}} \cdots \overrightarrow{I_{\pi(\ell)}}$ (since $\omega(i)$ is a letter of the word $\overrightarrow{I_{\pi(u)}}$, and since $\omega(j)$ is a letter of the word $\overrightarrow{I_{\pi(u+1)}}$). In other words, the letter $\omega(j)$ appears after the letter $\omega(i)$ in the word ω (since $\omega = \overrightarrow{I_{\pi(1)}}\,\overrightarrow{I_{\pi(2)}} \cdots \overrightarrow{I_{\pi(\ell)}}$). In other words, $\omega^{-1}(\omega(j)) > \omega^{-1}(\omega(i))$. In other words, $j > i$, so that $i < j$.

Now, let us forget that we fixed i and j. We thus have proven that every $i \in \omega^{-1}\left(I_{\pi(u)}\right)$ and $j \in \omega^{-1}\left(I_{\pi(u+1)}\right)$ satisfy $i < j$. In other words, $\omega^{-1}\left(I_{\pi(u)}\right) < \omega^{-1}\left(I_{\pi(u+1)}\right)$ (by the definition of $\omega^{-1}\left(I_{\pi(u)}\right) < \omega^{-1}\left(I_{\pi(u+1)}\right)$).

Now, let us forget that we fixed u. We thus have proven that $\omega^{-1}\left(I_{\pi(u)}\right) < \omega^{-1}\left(I_{\pi(u+1)}\right)$ for every $u \in \{1, 2, \ldots, \ell-1\}$. In other words, $\omega^{-1}\left(I_{\pi(1)}\right) < \omega^{-1}\left(I_{\pi(2)}\right) < \cdots < \omega^{-1}\left(I_{\pi(\ell)}\right)$. Therefore, (13.157.7) (applied to $\tau = \pi$) yields $\pi = \operatorname{iper}'_\alpha\left(\underbrace{\omega}_{=\operatorname{iper}_\alpha \pi}\right) = \operatorname{iper}'_\alpha(\operatorname{iper}_\alpha \pi) = (\operatorname{iper}'_\alpha \circ \operatorname{iper}_\alpha)(\pi)$. Hence, $(\operatorname{iper}'_\alpha \circ \operatorname{iper}_\alpha)(\pi) = \pi = \operatorname{id}(\pi)$.

Now, let us forget that we fixed π. We thus have proven that $(\operatorname{iper}'_\alpha \circ \operatorname{iper}_\alpha)(\pi) = \operatorname{id}(\pi)$ for every $\pi \in \mathfrak{S}_\ell$. In other words, $\operatorname{iper}'_\alpha \circ \operatorname{iper}_\alpha = \operatorname{id}$, qed.

j in the concatenation $\overrightarrow{I_{\tau(1)}}\overrightarrow{I_{\tau(2)}}\cdots\overrightarrow{I_{\tau(p+q)}}$ (because the letter i appears in $\overrightarrow{I_{i'}}$, while the letter j appears in $\overrightarrow{I_{j'}}$). Since $\overrightarrow{I_{\tau(1)}}\overrightarrow{I_{\tau(2)}}\cdots\overrightarrow{I_{\tau(p+q)}} = \omega$, this shows that the letter i appears before the letter j in the word ω. In other words, $\omega^{-1}(i) < \omega^{-1}(j)$.

So we have shown that if two elements i and j of $\{1, 2, \ldots, n\}$ satisfy $i < j$, then $\omega^{-1}(i) < \omega^{-1}(j)$. In other words, $\omega^{-1}(1) < \omega^{-1}(2) < \cdots < \omega^{-1}(n)$. Similarly, $\omega^{-1}(n+1) < \omega^{-1}(n+2) < \cdots < \omega^{-1}(n+m)$. Combining these two chains of inequalities, we conclude that $\omega \in \operatorname{Sh}_{n,m}$. This proves the $\Longrightarrow$ direction.

$\Longleftarrow$: Assume that $\omega \in \operatorname{Sh}_{n,m}$. We need to prove that $\tau \in \operatorname{Sh}_{p,q}$.

We have $\omega \in \operatorname{Sh}_{n,m}$. Thus, $\omega^{-1}(1) < \omega^{-1}(2) < \cdots < \omega^{-1}(n)$ and $\omega^{-1}(n+1) < \omega^{-1}(n+2) < \cdots < \omega^{-1}(n+m)$.

Let i' and j' be two elements of $\{1, 2, \ldots, p\}$ such that $i' < j'$. We are going to prove that $\tau^{-1}(i') < \tau^{-1}(j')$.

Indeed, assume the contrary. Then, $\tau^{-1}(i') \geq \tau^{-1}(j')$, thus $\tau^{-1}(i') > \tau^{-1}(j')$ (since τ is a permutation). Hence, the word $\overrightarrow{I_{i'}}$ appears after[1096] the word $\overrightarrow{I_{j'}}$ in the concatenation $\overrightarrow{I_{\tau(1)}}\overrightarrow{I_{\tau(2)}}\cdots\overrightarrow{I_{\tau(p+q)}}$.

Now, fix any $i \in I_{i'}$ (such an i exists since $I_{i'}$ is nonempty) and fix any $j \in I_{j'}$ (this exists for similar reasons). We have $I_{i'} < I_{j'}$ (since $I_1 < I_2 < \cdots < I_{p+q}$ and $i' < j'$), so that $i < j$ (since $i \in I_{i'}$ and $j \in I_{j'}$).

But the letter i appears after the letter j in the concatenation $\overrightarrow{I_{\tau(1)}}\overrightarrow{I_{\tau(2)}}\cdots\overrightarrow{I_{\tau(p+q)}}$ (because the letter i appears in the word $\overrightarrow{I_{i'}}$, whereas the letter j appears in the word $\overrightarrow{I_{j'}}$, and we know that the word $\overrightarrow{I_{i'}}$ appears after the word $\overrightarrow{I_{j'}}$ in the concatenation $\overrightarrow{I_{\tau(1)}}\overrightarrow{I_{\tau(2)}}\cdots\overrightarrow{I_{\tau(p+q)}}$). In other words, the letter i appears after the letter j in the word ω (since $\overrightarrow{I_{\tau(1)}}\overrightarrow{I_{\tau(2)}}\cdots\overrightarrow{I_{\tau(p+q)}} = \omega$). In other words, $\omega^{-1}(i) > \omega^{-1}(j)$.

But $i' \in \{1, 2, \ldots, p\}$, so that $I_{i'} \subset I_1 \cup I_2 \cup \cdots \cup I_p = [0 : n]^+$ and thus $i \in I_{i'} \subset [0 : n]^+$. Similarly, $j \in [0 : n]^+$. Since the map ω^{-1} restricted to $[0 : n]^+$ is strictly increasing (since $\omega^{-1}(1) < \omega^{-1}(2) < \cdots < \omega^{-1}(n)$), we thus have $\omega^{-1}(i) < \omega^{-1}(j)$ (since $i < j$), contradicting $\omega^{-1}(i) > \omega^{-1}(j)$. This contradiction shows that our assumption was wrong, and so we have shown that $\tau^{-1}(i') < \tau^{-1}(j')$.

Thus, we have proven that if two elements i' and j' of $\{1, 2, \ldots, p\}$ satisfy $i' < j'$, then $\tau^{-1}(i') < \tau^{-1}(j')$. In other words, $\tau^{-1}(1) < \tau^{-1}(2) < \cdots < \tau^{-1}(p)$. Similarly, we can show that $\tau^{-1}(p+1) < \tau^{-1}(p+2) < \cdots < \tau^{-1}(p+q)$. The combination of these two chains of inequalities shows that $\tau \in \operatorname{Sh}_{p,q}$. Thus, the $\Longleftarrow$ direction is proven. The proof of Proposition 6.2.18 is thus complete. □

13.159. **Solution to Exercise 6.2.21.** *Solution to Exercise 6.2.21.*

Proof of Lemma 6.2.20. We have $\tau \in \operatorname{Sh}_{p,q}$ if and only if $\operatorname{iper}(\alpha\beta, \tau) \in \operatorname{Sh}_{n,m}$ (by Proposition 6.2.18). Since we know that $\tau \in \operatorname{Sh}_{p,q}$, we thus conclude that $\operatorname{iper}(\alpha\beta, \tau) \in \operatorname{Sh}_{n,m}$. In other words, $\sigma \in \operatorname{Sh}_{n,m}$ (since $\sigma = \operatorname{iper}(\alpha\beta, \tau)$).

Write the composition $\alpha\beta$ in the form $(\gamma_1, \gamma_2, \ldots, \gamma_{p+q})$. For every $j \in \{0, 1, \ldots, p+q\}$, let s_j denote the integer $\sum_{k=1}^{j} \gamma_{\tau(k)}$. Then, $0 = s_0 < s_1 < s_2 < \cdots < s_{p+q} = n + m$.

Let w denote the word $u \underset{\sigma}{\sqcup\!\sqcup} v$.

The first p parts of the composition $\alpha\beta$ form the composition α of n. Hence, the interval system $(I_1, I_2, \ldots, I_{p+q})$ corresponding to $\alpha\beta$ satisfies $I_1 \cup I_2 \cup \cdots \cup I_p = [0 : n]^+$ and $I_{p+1} \cup I_{p+2} \cup \cdots \cup I_{p+q} = [n : n+m]^+$.

Let $i \in \{1, 2, \ldots, p+q\}$. We know that $I_{\tau(i)}$ is an interval of $\mathbb{Z}$ satisfying either $I_{\tau(i)} \subset [0 : n]^+$ or $I_{\tau(i)} \subset [n : n+m]^+$ (in fact, if $\tau(i) \leq p$, then $I_{\tau(i)} \subset I_1 \cup I_2 \cup \cdots \cup I_p = [0 : n]^+$, whereas otherwise, $I_{\tau(i)} \subset I_{p+1} \cup I_{p+2} \cup \cdots \cup I_{p+q} = [n : n+m]^+$). Also, Proposition 6.2.14(a) (applied to $n+m$, $\alpha\beta$, $(\gamma_1, \gamma_2, \ldots, \gamma_{p+q})$,

[1096] "after" does not imply "immediately after".

$p+q$ and i instead of n, α, $(\alpha_1, \alpha_2, \ldots, \alpha_\ell)$, ℓ and j) yields

$$\sigma^{-1}\left(I_{\tau(i)}\right) = \left[\underbrace{\sum_{k=1}^{i-1} \gamma_{\tau(k)}}_{\substack{=s_{i-1} \\ \text{(by the definition of } s_{i-1})}} : \underbrace{\sum_{k=1}^{i} \gamma_{\tau(k)}}_{\substack{=s_i \\ \text{(by the definition of } s_i)}}\right]^+ = [s_{i-1} : s_i]^+ . \tag{13.159.1}$$

Consequently, $\sigma^{-1}\left(I_{\tau(i)}\right)$ is an interval. Therefore, we can apply Lemma 6.2.10(a) to $I_{\tau(i)}$ instead of I. As a result, we obtain

$$\left(u \underset{\sigma}{\sqcup\!\sqcup} v\right)\left[\sigma^{-1}\left(I_{\tau(i)}\right)\right] = (uv)\left[I_{\tau(i)}\right].$$

Hence,

$$(uv)\left[I_{\tau(i)}\right] = \underbrace{\left(u \underset{\sigma}{\sqcup\!\sqcup} v\right)}_{=w}\left[\underbrace{\sigma^{-1}\left(I_{\tau(i)}\right)}_{\substack{=[s_{i-1}:s_i]^+ \\ \text{(by (13.159.1))}}}\right] = w\left[[s_{i-1} : s_i]^+\right]. \tag{13.159.2}$$

Now, forget that we fixed i. We thus have proven (13.159.2) for every $i \in \{1, 2, \ldots, p+q\}$. Now,

$$\begin{aligned}
&\underbrace{(uv)\left[I_{\tau(1)}\right]}_{\substack{=w\left[[s_0:s_1]^+\right] \\ \text{(by (13.159.2))}}} \cdot \underbrace{(uv)\left[I_{\tau(2)}\right]}_{\substack{=w\left[[s_1:s_2]^+\right] \\ \text{(by (13.159.2))}}} \cdot \cdots \cdot \underbrace{(uv)\left[I_{\tau(p+q)}\right]}_{\substack{=w\left[[s_{p+q-1}:s_{p+q}]^+\right] \\ \text{(by (13.159.2))}}} \\
&= w\left[[s_0 : s_1]^+\right] \cdot w\left[[s_1 : s_2]^+\right] \cdot \cdots \cdot w\left[[s_{p+q-1} : s_{p+q}]^+\right] \\
&= w\left[[s_0 : s_{p+q}]^+\right] = w\left[[0 : n+m]^+\right] \qquad (\text{since } s_0 = 0 \text{ and } s_{p+q} = n+m) \\
&= w \qquad \left(\text{since the word } w = u \underset{\sigma}{\sqcup\!\sqcup} v \text{ has length } n+m\right) \\
&= u \underset{\sigma}{\sqcup\!\sqcup} v.
\end{aligned}$$

This proves Lemma 6.2.20. □

13.160. **Solution to Exercise 6.2.24.** *Solution to Exercise 6.2.24.*

Proof of Proposition 6.2.23. We have $h(1) \geq h(2) \geq \cdots \geq h(p)$. Hence, for every $w \in \mathfrak{W}$, the set $\{i \in \{1, 2, \ldots, p\} \mid h(i) = w\}$ is a (possibly empty) interval of $\{1, 2, \ldots, p\}$, and will be denoted by P_w. Similarly, for every $w \in \mathfrak{W}$, the set $\{i \in \{p+1, p+2, \ldots, p+q\} \mid h(i) = w\}$ is a (possibly empty) interval of $\{p+1, p+2, \ldots, p+q\}$, and will be denoted by Q_w. Every $w \in \mathfrak{W}$ satisfies $h^{-1}(w) = P_w \sqcup Q_w$. Notice that $|P_w| = \mathfrak{a}(w)$ and $|Q_w| = \mathfrak{b}(w)$ for all $w \in \mathfrak{W}$.

Let $(g_1, g_2, \ldots, g_{p+q})$ be the result of sorting the list $(h(1), h(2), \ldots, h(p+q))$ in decreasing order. Then, $g_1 \geq g_2 \geq \cdots \geq g_{p+q}$. Hence, for every $w \in \mathfrak{W}$, the set $\{j \in \{1, 2, \ldots, p+q\} \mid g_j = w\}$ is a (possibly empty) interval of $\{1, 2, \ldots, p+q\}$. Denote this interval by I_w. The size of this interval is

$$\begin{aligned}
|I_w| &= |\{j \in \{1, 2, \ldots, p+q\} \mid g_j = w\}| \qquad (\text{since } I_w = \{j \in \{1, 2, \ldots, p+q\} \mid g_j = w\}) \\
&= \left|\underbrace{\{j \in \{1, 2, \ldots, p+q\} \mid h(j) = w\}}_{=h^{-1}(w)=P_w \sqcup Q_w}\right| \qquad \left(\begin{array}{c}\text{since } (g_1, g_2, \ldots, g_{p+q}) \text{ is the result of} \\ \text{sorting the list } (h(1), h(2), \ldots, h(p+q))\end{array}\right) \\
&= |P_w \sqcup Q_w| = |P_w| + |Q_w|.
\end{aligned}$$

Notice that $\{1, 2, \ldots, p+q\} = \bigsqcup_{w \in \mathfrak{W}} I_w$ (by the definition of the I_w).

Fix $w \in \mathfrak{W}$. Let us define an (I_w, P_w, Q_w)*-shuffle* to mean a bijection $\kappa : I_w \to P_w \sqcup Q_w$ having the property that the maps $\kappa^{-1}|_{P_w} : P_w \to I_w$ and $\kappa^{-1}|_{Q_w} : Q_w \to I_w$ are strictly increasing. It is easy to see that such a (I_w, P_w, Q_w)-shuffle κ is uniquely determined by the subset $\kappa^{-1}(P_w)$ of I_w, and that for a given subset U of I_w, such a (I_w, P_w, Q_w)-shuffle κ satisfying $\kappa^{-1}(P_w) = U$ exists if and only if $|U| = |P_w|$. Hence, there are as many (I_w, P_w, Q_w)-shuffles as there are subsets U of I_w satisfying $|U| = |P_w|$. In other words,

$$\begin{aligned}
&(\text{the number of } (I_w, P_w, Q_w)\text{-shuffles}) \\
&= (\text{the number of subsets } U \text{ of } I_w \text{ satisfying } |U| = |P_w|) \\
&= \binom{|I_w|}{|P_w|} = \binom{\mathfrak{a}(w) + \mathfrak{b}(w)}{\mathfrak{a}(w)} \qquad (13.160.1)
\end{aligned}$$

(since $|I_w| = \underbrace{|P_w|}_{=\mathfrak{a}(w)} + \underbrace{|Q_w|}_{=\mathfrak{b}(w)} = \mathfrak{a}(w) + \mathfrak{b}(w)$ and $|P_w| = \mathfrak{a}(w)$).

Now, forget that we fixed w.

Consider any $\tau \in \mathrm{Sh}_{p,q}$ satisfying $h(\tau(1)) \geq h(\tau(2)) \geq \cdots \geq h(\tau(p+q))$. Then, $(h(\tau(1)), h(\tau(2)), \ldots, h(\tau(p+q))) = (g_1, g_2, \ldots, g_{p+q})$ [1097]. In other words, $h(\tau(j)) = g_j$ for every $j \in \{1, 2, \ldots, p+q\}$. Thus,

$$\tau(I_w) = P_w \sqcup Q_w \qquad \text{for every } w \in \mathfrak{W} \qquad (13.160.2)$$

[1098]. As a consequence, for every $w \in \mathfrak{W}$, the bijection $\tau : \{1, 2, \ldots, p+q\} \to \{1, 2, \ldots, p+q\}$ restricts to a bijection $\tau_w : I_w \to P_w \sqcup Q_w$. This bijection τ_w is an (I_w, P_w, Q_w)-shuffle[1099]. Thus, we have obtained a family $(\tau_w)_{w \in \mathfrak{W}}$ of (I_w, P_w, Q_w)-shuffles parametrized over all $w \in \mathfrak{W}$.

Now, let us forget that we fixed τ. We thus have constructed, for every $\tau \in \mathrm{Sh}_{p,q}$ satisfying $h(\tau(1)) \geq h(\tau(2)) \geq \cdots \geq h(\tau(p+q))$, a family $(\tau_w)_{w \in \mathfrak{W}}$ of (I_w, P_w, Q_w)-shuffles parametrized over all $w \in \mathfrak{W}$. We thus obtain a map

$$\begin{aligned}
\{\tau \in \mathrm{Sh}_{p,q} \mid h(\tau(1)) \geq h(\tau(2)) \geq \cdots \geq h(\tau(p+q))\} &\to \prod_{w \in \mathfrak{W}} (\text{the set of all } (I_w, P_w, Q_w)\text{-shuffles}), \\
\tau &\mapsto (\tau_w)_{w \in \mathfrak{W}}.
\end{aligned}$$

This map is injective[1100] and surjective[1101]. Hence, it is bijective, and this yields that the sets

$$\{\tau \in \mathrm{Sh}_{p,q} \mid h(\tau(1)) \geq h(\tau(2)) \geq \cdots \geq h(\tau(p+q))\}$$

[1097] *Proof.* The list $(h(\tau(1)), h(\tau(2)), \ldots, h(\tau(p+q)))$ is a permutation of the list $(h(1), h(2), \ldots, h(p+q))$ (since $\tau \in \mathfrak{S}_{p+q}$), but is weakly decreasing (since $h(\tau(1)) \geq h(\tau(2)) \geq \cdots \geq h(\tau(p+q))$). Hence, the list $(h(\tau(1)), h(\tau(2)), \ldots, h(\tau(p+q)))$ is the result of sorting the list $(h(1), h(2), \ldots, h(p+q))$ in decreasing order. But this yields that $(h(\tau(1)), h(\tau(2)), \ldots, h(\tau(p+q))) = (g_1, g_2, \ldots, g_{p+q})$ (since $(g_1, g_2, \ldots, g_{p+q})$, too, is the result of sorting the list $(h(1), h(2), \ldots, h(p+q))$ in decreasing order), qed.

[1098] *Proof of (13.160.2):* Let $w \in \mathfrak{W}$. Then,

$$\begin{aligned}
I_w &= \left\{ j \in \{1, 2, \ldots, p+q\} \mid \underbrace{g_j}_{\substack{=h(\tau(j)) \\ =(h\circ\tau)(j)}} = w \right\} = \{ j \in \{1, 2, \ldots, p+q\} \mid (h \circ \tau)(j) = w \} \\
&= (h \circ \tau)^{-1}(w) = \tau^{-1}\left(h^{-1}(w)\right).
\end{aligned}$$

Since τ is a permutation, this yields $\tau(I_w) = h^{-1}(w) = P_w \sqcup Q_w$, and thus (13.160.2) is proven.

[1099] *Proof.* We have $\tau \in \mathrm{Sh}_{p,q}$, so that $\tau^{-1}(1) < \tau^{-1}(2) < \cdots < \tau^{-1}(p)$ and $\tau^{-1}(p+1) < \tau^{-1}(p+2) < \cdots < \tau^{-1}(p+q)$. In other words, the restriction of τ^{-1} to $\{1, 2, \ldots, p\}$ is strictly increasing, and the restriction of τ^{-1} to $\{p+1, p+2, \ldots, p+q\}$ is strictly increasing. Since the restriction of τ^{-1} to $\{1, 2, \ldots, p\}$ is strictly increasing, the restriction of τ^{-1} to P_w must also be strictly increasing (because $P_w \subset \{1, 2, \ldots, p\}$). But this restriction is $\tau_w^{-1} \mid_{P_w} : P_w \to I_w$. Hence, $\tau_w^{-1} \mid_{P_w} : P_w \to I_w$ is strictly increasing. Similarly, the same can be said about $\tau_w^{-1} \mid_{Q_w} : Q_w \to I_w$. Since the maps $\tau_w^{-1} \mid_{P_w} : P_w \to I_w$ and $\tau_w^{-1} \mid_{Q_w} : Q_w \to I_w$ are strictly increasing, we conclude that τ_w is an (I_w, P_w, Q_w)-shuffle (by the definition of a (I_w, P_w, Q_w)-shuffle), qed.

[1100] In fact, any $\tau \in \mathrm{Sh}_{p,q}$ is uniquely determined by $(\tau_w)_{w \in \mathfrak{W}}$ (because τ_w is the restriction of τ to I_w (with a restricted codomain, but this doesn't matter right now), and so knowing $(\tau_w)_{w \in \mathfrak{W}}$ means knowing the values τ on each of the intervals I_w; but this means knowing all values of τ, because $\{1, 2, \ldots, p+q\} = \bigsqcup_{w \in \mathfrak{W}} I_w$).

[1101] *Proof.* We need to show that for every

$$(\sigma_w)_{w \in \mathfrak{W}} \in \prod_{w \in \mathfrak{W}} (\text{the set of all } (I_w, P_w, Q_w)\text{-shuffles}),$$

there exists a $\tau \in \mathrm{Sh}_{p,q}$ satisfying $h(\tau(1)) \geq h(\tau(2)) \geq \cdots \geq h(\tau(p+q))$ and $(\tau_w)_{w \in \mathfrak{W}} = (\sigma_w)_{w \in \mathfrak{W}}$.

and

$$\prod_{w\in\mathfrak{W}} (\text{the set of all } (I_w, P_w, Q_w)\text{-shuffles})$$

are in bijection. Thus,

$$\begin{aligned}
&|\{\tau \in \operatorname{Sh}_{p,q} \ \mid \ h(\tau(1)) \geq h(\tau(2)) \geq \cdots \geq h(\tau(p+q))\}| \\
&= \left| \prod_{w\in\mathfrak{W}} (\text{the set of all } (I_w, P_w, Q_w)\text{-shuffles}) \right| \\
&= \prod_{w\in\mathfrak{W}} \underbrace{|(\text{the set of all } (I_w, P_w, Q_w)\text{-shuffles})|}_{\substack{=(\text{the number of } (I_w,P_w,Q_w)\text{-shuffles})=\dbinom{\mathfrak{a}(w)+\mathfrak{b}(w)}{\mathfrak{a}(w)} \\ (\text{by } (13.160.1))}} \\
&= \prod_{w\in\mathfrak{W}} \binom{\mathfrak{a}(w)+\mathfrak{b}(w)}{\mathfrak{a}(w)}.
\end{aligned}$$

This is precisely the statement of Proposition 6.2.23. □

So fix some $(\sigma_w)_{w\in\mathfrak{W}} \in \prod_{w\in\mathfrak{W}} (\text{the set of all } (I_w, P_w, Q_w)\text{-shuffles})$. For every $w \in \mathfrak{W}$, the map σ_w is an (I_w, P_w, Q_w)-shuffle, hence a bijection from I_w to $P_w \sqcup Q_w$. Since $\bigsqcup_{w\in\mathfrak{W}} I_w = \{1, 2, \ldots, p+q\}$ and $\bigsqcup_{w\in\mathfrak{W}} \underbrace{(P_w \sqcup Q_w)}_{=h^{-1}(w)} = \bigsqcup_{w\in\mathfrak{W}} h^{-1}(w) = \{1, 2, \ldots, p+q\}$, we can piece these bijections σ_w together to a bijection

$$\bigsqcup_{w\in\mathfrak{W}} \sigma_w : \{1, 2, \ldots, p+q\} \to \{1, 2, \ldots, p+q\},$$

whose restriction to each interval I_w coincides with the respective σ_w (except that the codomains of the maps are different). Let τ be this bijection $\bigsqcup_{w\in\mathfrak{W}} \sigma_w$. Clearly, $\tau \in \mathfrak{S}_{p+q}$. We will now show that $\tau \in \operatorname{Sh}_{p,q}$, $h(\tau(1)) \geq h(\tau(2)) \geq \cdots \geq h(\tau(p+q))$ and $(\tau_w)_{w\in\mathfrak{W}} = (\sigma_w)_{w\in\mathfrak{W}}$. Once this is done, the required surjectivity will clearly follow.

Since $\tau = \bigsqcup_{w\in\mathfrak{W}} \sigma_w$, we have $\tau\mid_{I_w} = \sigma_w$ for every $w \in \mathfrak{W}$ (up to the fact that the maps $\tau\mid_{I_w}$ and σ_w have different codomains). More precisely, $\tau_w = \sigma_w$ for every $w \in \mathfrak{W}$. Hence, for every $w \in \mathfrak{W}$, the map σ_w is a restriction of τ (with appropriately restricted codomain).

The map $\tau = \bigsqcup_{w\in\mathfrak{W}} \sigma_w$ is pieced together from bijections $\sigma_w : I_w \to P_w \sqcup Q_w$. Thus, $\tau(I_w) = P_w \sqcup Q_w$ for every $w \in \mathfrak{W}$. In other words, $\tau(I_w) = h^{-1}(w)$ for every $w \in \mathfrak{W}$ (since every $w \in \mathfrak{W}$ satisfies $h^{-1}(w) = P_w \sqcup Q_w$). Hence,

$$\text{every } i \in \{1, 2, \ldots, p+q\} \text{ satisfies } h(\tau(i)) = g_i. \tag{13.160.3}$$

[*Proof of (13.160.3):* Let $i \in \{1, 2, \ldots, p+q\}$. Set $w = h(\tau(i))$. Then, $\tau(i) \in h^{-1}(w) = \tau(I_w)$ (since $\tau(I_w) = h^{-1}(w)$) and thus $i \in I_w$ (since τ is a bijection), so that $i \in I_w = \{j \in \{1, 2, \ldots, p+q\} \ \mid \ g_j = w\}$ and thus $g_i = w = h(\tau(i))$. This proves (13.160.3).]

We have $g_1 \geq g_2 \geq \cdots \geq g_{p+q}$. Due to (13.160.3), this rewrites as $h(\tau(1)) \geq h(\tau(2)) \geq \cdots \geq h(\tau(p+q))$.

We now are going to prove that $\tau \in \operatorname{Sh}_{p,q}$. In order to prove this, it is enough to show that $\tau^{-1}(1) < \tau^{-1}(2) < \cdots < \tau^{-1}(p)$ and $\tau^{-1}(p+1) < \tau^{-1}(p+2) < \cdots < \tau^{-1}(p+q)$. We shall only verify $\tau^{-1}(1) < \tau^{-1}(2) < \cdots < \tau^{-1}(p)$ (the proof of $\tau^{-1}(p+1) < \tau^{-1}(p+2) < \cdots < \tau^{-1}(p+q)$ being analogous).

So let i and j be elements of $\{1, 2, \ldots, p\}$ such that $i < j$. We will prove that $\tau^{-1}(i) < \tau^{-1}(j)$. Indeed, assume the contrary. Then, $\tau^{-1}(i) \geq \tau^{-1}(j)$. In other words, $\tau^{-1}(j) \leq \tau^{-1}(i)$. Thus, $h\left(\tau\left(\tau^{-1}(j)\right)\right) \geq h\left(\tau\left(\tau^{-1}(i)\right)\right)$ (since $h(\tau(1)) \geq h(\tau(2)) \geq \cdots \geq h(\tau(p+q))$). In other words, $h(j) \geq h(i)$. Combined with $h(i) \geq h(j)$ (since $h(1) \geq h(2) \geq \cdots \geq h(p)$ and $i < j$), this yields $h(i) = h(j)$. So we can define a $w \in \mathfrak{W}$ by $w = h(i) = h(j)$. Consider this w. We have $h(i) = w$, so that $i \in h^{-1}(w) = P_w \sqcup Q_w$. But we cannot have $i \in Q_w$ (because i lies in the set $\{1, 2, \ldots, p\}$, which is disjoint to Q_w (since $Q_w \subset \{p+1, p+2, \ldots, p+q\}$)). Since we have $i \in P_w \sqcup Q_w$ but not $i \in Q_w$, we must have $i \in P_w$. Similarly, $j \in P_w$. But recall that σ_w is a (I_w, P_w, Q_w)-shuffle. In other words, $\sigma_w : I_w \to P_w \sqcup Q_w$ is a bijection having the property that the maps $\sigma_w^{-1}\mid_{P_w} : P_w \to I_w$ and $\sigma_w^{-1}\mid_{Q_w} : Q_w \to I_w$ are strictly increasing. Since $\sigma_w^{-1}\mid_{P_w} : P_w \to I_w$ is strictly increasing, we have $\left(\sigma_w^{-1}\mid_{P_w}\right)(i) < \left(\sigma_w^{-1}\mid_{P_w}\right)(j)$ (since $i < j$ and $i \in P_w$ and $j \in P_w$). Since $\left(\sigma_w^{-1}\mid_{P_w}\right)(i) = \sigma_w^{-1}(i) = \tau^{-1}(i)$ (because σ_w is a restriction of τ) and $\left(\sigma_w^{-1}\mid_{P_w}\right)(j) = \tau^{-1}(j)$ (similarly), this rewrites as $\tau^{-1}(i) < \tau^{-1}(j)$, which contradicts $\tau^{-1}(i) \geq \tau^{-1}(j)$. This contradiction proves our assumption wrong, and so we have $\tau^{-1}(i) < \tau^{-1}(j)$.

Let us forget that we fixed i and j. We thus have seen that $\tau^{-1}(i) < \tau^{-1}(j)$ for any elements i and j of $\{1, 2, \ldots, p\}$ such that $i < j$. In other words, $\tau^{-1}(1) < \tau^{-1}(2) < \cdots < \tau^{-1}(p)$. Similarly, $\tau^{-1}(p+1) < \tau^{-1}(p+2) < \cdots < \tau^{-1}(p+q)$. Thus, $\tau \in \operatorname{Sh}_{p,q}$.

We now know that $\tau \in \operatorname{Sh}_{p,q}$ and $h(\tau(1)) \geq h(\tau(2)) \geq \cdots \geq h(\tau(p+q))$. Finally, $(\tau_w)_{w\in\mathfrak{W}} = (\sigma_w)_{w\in\mathfrak{W}}$ follows from the very definition of τ_w (since $\tau_w = \sigma_w$ for every $w \in \mathfrak{W}$). This completes the proof.

13.161. **Solution to Exercise 6.2.25.** *Solution to Exercise 6.2.25.* $\Longrightarrow$: Assume that w is Lyndon. We need to prove that for any two nonempty words $u \in \mathfrak{A}^*$ and $v \in \mathfrak{A}^*$ satisfying $w = uv$, there exists at least one $s \in u \shuffle v$ satisfying $s > w$.

Let $u \in \mathfrak{A}^*$ and $v \in \mathfrak{A}^*$ be two nonempty words satisfying $w = uv$. Then, vu is an element of $u \shuffle v$ [1102] and satisfies $vu > uv$ (by Proposition 6.1.14(c)). Thus, $vu > uv = w$. Hence, there exists at least one $s \in u \shuffle v$ satisfying $s > w$ (namely, $s = vu$). Thus, the $\Longrightarrow$ direction of Exercise 6.2.25 is solved.

$\Longleftarrow$: Assume that for any two nonempty words $u \in \mathfrak{A}^*$ and $v \in \mathfrak{A}^*$ satisfying $w = uv$,

$$\text{there exists at least one } s \in u \shuffle v \text{ satisfying } s > w. \tag{13.161.1}$$

We need to prove that w is Lyndon.

In fact, assume the contrary. Thus, w is not Lyndon. Let $(a_1, a_2, \ldots, a_p)$ be the CFL factorization of w; then, $a_1 \geq a_2 \geq \cdots \geq a_p$ and $a_1 a_2 \cdots a_p = w$. Also, $p \neq 0$ (since $a_1 a_2 \cdots a_p = w$ is nonempty). Since a_1 is Lyndon but w is not, we have $a_1 \neq w$. Thus, $p \neq 1$ (because otherwise, we would have $a_1 = a_1 a_2 \cdots a_p = w$, contradicting $a_1 \neq w$). Combined with $p \neq 0$, this yields $p \geq 2$.

Let $u = a_1$ and $v = a_2 a_3 \cdots a_p$. Then, u is Lyndon (since $u = a_1$), thus nonempty. Also, v is a nonempty product of Lyndon words (nonempty because $p \geq 2$), and hence nonempty itself (since Lyndon words are nonempty). Clearly, $\underbrace{u}_{=a_1} \underbrace{v}_{=a_2 a_3 \cdots a_p} = a_1 (a_2 a_3 \cdots a_p) = a_1 a_2 \cdots a_p = w$. Thus, (13.161.1) yields that there exists at least one $s \in u \shuffle v$ satisfying $s > w$. Thus, w is not the lexicographically highest element of the multiset $u \shuffle v$.

Now, notice that a_2, a_3, $\ldots$, a_p are Lyndon words satisfying $a_2 \geq a_3 \geq \cdots \geq a_p$ (since $a_1 \geq a_2 \geq \cdots \geq a_p$) and $a_2 a_3 \cdots a_p = v$. Hence, $(a_2, a_3, \ldots, a_p)$ is the CFL factorization of v. We have $u = a_1 \geq a_{j+1}$ for every $j \in \{1, 2, \ldots, p-1\}$ (since $a_1 \geq a_2 \geq \cdots \geq a_p$). Thus, Theorem 6.2.2(e) (applied to $p-1$ and $(a_2, a_3, \ldots, a_p)$ instead of q and $(b_1, b_2, \ldots, b_q)$) yields that the lexicographically highest element of the multiset $u \shuffle v$ is uv, and the multiplicity with which this word uv appears in the multiset $u \shuffle v$ is $\operatorname{mult}_u v + 1$.

Now, we know that the lexicographically highest element of the multiset $u \shuffle v$ is $uv = w$. This contradicts the fact that w is not the lexicographically highest element of the multiset $u \shuffle v$. This contradiction shows that our assumption was wrong. Thus, w is Lyndon. This completes the solution of the $\Longleftarrow$ direction of Exercise 6.2.25.

Thus, both the $\Longrightarrow$ and $\Longleftarrow$ directions of Exercise 6.2.25 are proven. Exercise 6.2.25 is thus solved.

[*Remark:* Exercise 6.2.25 still holds in the partial-order setting[1103]. To prove this, we can use Proposition 13.146.3 along with the fact that Exercise 6.2.25 holds in the total-order setting. Here are the details:

Solution to Exercise 6.2.25 in the partial-order setting. $\Longrightarrow$: In the partial-order setting, the $\Longrightarrow$ direction of Exercise 6.2.25 can be proved in the same way as it was proven in the total-order setting.

$\Longleftarrow$: Assume that for any two nonempty words $u \in \mathfrak{A}^*$ and $v \in \mathfrak{A}^*$ satisfying $w = uv$,

$$\text{there exists at least one } s \in u \shuffle v \text{ satisfying } s > w. \tag{13.161.2}$$

We need to prove that w is Lyndon.

Let $\mathfrak{B}$ be any linear extension of the alphabet $\mathfrak{A}$. Then, $\mathfrak{B}^*$ is an extension of $\mathfrak{A}^*$ (according to Proposition 13.146.3(a)). Thus, $\mathfrak{B}^* = \mathfrak{A}^*$ as sets. It is easy to see that for any two nonempty words $u \in \mathfrak{B}^*$ and $v \in \mathfrak{B}^*$ satisfying $w = uv$,

$$\text{there exists at least one } s \in u \shuffle v \text{ satisfying } s > w \text{ in } \mathfrak{B}^*. \tag{13.161.3}$$

[1104]

[1102] *Proof.* Let $n = \ell(u)$ and $m = \ell(v)$. Then, the permutation in $\mathfrak{S}_{n+m}$ which is written as $(n+1, n+2, \ldots, n+m, 1, 2, \ldots, n)$ in one-line notation belongs to $\operatorname{Sh}_{n,m}$. If we denote this permutation by σ, then $vu = u \underset{\sigma}{\shuffle} v \in u \shuffle v$, qed.

[1103] See Exercise 6.1.33 for an explanation of what the partial-order setting is.

[1104] *Proof of (13.161.3):* Let $u \in \mathfrak{B}^*$ and $v \in \mathfrak{B}^*$. Then, $u \in \mathfrak{B}^* = \mathfrak{A}^*$ and $v \in \mathfrak{B}^* = \mathfrak{A}^*$. Hence, (13.161.2) shows that there exists at least one $s \in u \shuffle v$ satisfying $s > w$ in $\mathfrak{A}^*$. Let t be such an s. Then, t is an element of $u \shuffle v$ satisfying $t > w$ in $\mathfrak{A}^*$. Hence, $w < t$ in $\mathfrak{A}^*$ (since $t > w$ in $\mathfrak{A}^*$), so that $w \leq t$ in $\mathfrak{A}^*$.

Now, we can apply Exercise 6.2.25 to $\mathfrak{B}$ instead of $\mathfrak{A}$ (since $\mathfrak{B}$ is totally ordered). As a consequence, we see that w is Lyndon as a word in $\mathfrak{B}^*$ if and only if for any two nonempty words $u \in \mathfrak{B}^*$ and $v \in \mathfrak{B}^*$ satisfying $w = uv$, there exists at least one $s \in u \sqcup\!\sqcup v$ satisfying $s > w$ in $\mathfrak{B}^*$. Thus, w is Lyndon as a word in $\mathfrak{B}^*$ (because we know that for any two nonempty words $u \in \mathfrak{B}^*$ and $v \in \mathfrak{B}^*$ satisfying $w = uv$, there exists at least one $s \in u \sqcup\!\sqcup v$ satisfying $s > w$ in $\mathfrak{B}^*$).

Now, the definition of a Lyndon word shows the following: The word w (as a word in $\mathfrak{B}^*$) is Lyndon if and only if it is nonempty and satisfies the following property:

(13.161.4) Every nonempty proper suffix v of w satisfies $v > w$ in $\mathfrak{B}^*$.

Since the word w is Lyndon, this shows that the word w is nonempty and satisfies the property (13.161.4).

Let now v be a nonempty proper suffix of w (as a word in $\mathfrak{A}^*$). Then, v is a nonempty proper suffix of w (as a word in $\mathfrak{B}^*$). Thus, (13.161.4) shows that v satisfies $v > w$ in $\mathfrak{B}^*$.

Now, let us forget that we fixed $\mathfrak{B}$ and v. We thus have shown that if v is any nonempty proper suffix of w and if $\mathfrak{B}$ is any linear extension of the alphabet $\mathfrak{A}$, then

(13.161.5) $$v > w \text{ in } \mathfrak{B}^*.$$

But the definition of a Lyndon word shows the following: The word w (as a word in $\mathfrak{A}^*$) is Lyndon if and only if it is nonempty and satisfies the following property:

(13.161.6) Every nonempty proper suffix v of w satisfies $v > w$ (in $\mathfrak{A}^*$).

We already know that w is nonempty. We are now going to prove (13.161.6):

Let v be a nonempty proper suffix of w. We have $v > w$ in $\mathfrak{B}^*$ for every linear extension $\mathfrak{B}$ of $\mathfrak{A}$ (according to (13.161.5)). In other words, $w < v$ in $\mathfrak{B}^*$ for every linear extension $\mathfrak{B}$ of $\mathfrak{A}$.

But Proposition 13.146.3(b) (applied to w instead of u) yields that $w < v$ holds in $\mathfrak{A}^*$ if and only if we have

$$(w < v \text{ in } \mathfrak{B}^* \text{ for every linear extension } \mathfrak{B} \text{ of } \mathfrak{A}).$$

We thus conclude that $w < v$ holds in $\mathfrak{A}^*$ (since we know that $w < v$ in $\mathfrak{B}^*$ for every linear extension $\mathfrak{B}$ of $\mathfrak{A}$). In other words, $v > w$ in $\mathfrak{A}^*$. This proves (13.161.6).

So we know that the word w is nonempty and satisfies the property (13.161.6). Thus, the word w (as a word in $\mathfrak{A}^*$) is Lyndon (since the word w (as a word in $\mathfrak{A}^*$) is Lyndon if and only if it is nonempty and satisfies the property (13.161.6)). Thus, the $\Longleftarrow$ direction of Exercise 6.2.25 is proven in the partial-order setting.

Thus, both the $\Longrightarrow$ and $\Longleftarrow$ directions of Exercise 6.2.25 are proven in the partial-order setting. Hence, Exercise 6.2.25 is solved in the partial-order setting. □

]

13.162. **Solution to Exercise 6.3.3.** *Solution to Exercise 6.3.3.*

Proof of Remark 6.3.2. Let $(w_1, w_2, \ldots, w_{n+m})$ denote the concatenation $u \cdot v = (u_1, u_2, \ldots, u_n, v_1, v_2, \ldots, v_m)$. Then, $(u_1, u_2, \ldots, u_n, v_1, v_2, \ldots, v_m) = (w_1, w_2, \ldots, w_{n+m})$, so that $(u_1, u_2, \ldots, u_n) = (w_1, w_2, \ldots, w_n)$ and $(v_1, v_2, \ldots, v_m) = (w_{n+1}, w_{n+2}, \ldots, w_{n+m})$.

By the definition of b_u, we have $b_u = b_{u_1} b_{u_2} \cdots b_{u_n} = b_{w_1} b_{w_2} \cdots b_{w_n}$ (since $(u_1, u_2, \ldots, u_n) = (w_1, w_2, \ldots, w_n)$). By the definition of b_v, we have $b_v = b_{v_1} b_{v_2} \cdots b_{v_m} = b_{w_{n+1}} b_{w_{n+2}} \cdots b_{w_{n+m}}$ (since $(v_1, v_2, \ldots, v_m) = (w_{n+1}, w_{n+2}, \ldots, w_{n+m})$). Now,

$$\underbrace{b_u}_{= b_{w_1} b_{w_2} \cdots b_{w_n}} \sqcup\!\sqcup \underbrace{b_v}_{= b_{w_{n+1}} b_{w_{n+2}} \cdots b_{w_{n+m}}} = \left(b_{w_1} b_{w_2} \cdots b_{w_n}\right) \sqcup\!\sqcup \left(b_{w_{n+1}} b_{w_{n+2}} \cdots b_{w_{n+m}}\right)$$

(13.162.1) $$= \sum_{\sigma \in \operatorname{Sh}_{n,m}} b_{w_{\sigma(1)}} b_{w_{\sigma(2)}} \cdots b_{w_{\sigma(n+m)}}$$

We know that $\mathfrak{B}^*$ is an extension of $\mathfrak{A}^*$. Thus, any two elements a and b of $\mathfrak{A}^*$ satisfying $a \leq b$ in $\mathfrak{A}^*$ satisfy $a \leq b$ in $\mathfrak{B}^*$ (according to the definition of an "extension"). Applying this to $a = w$ and $b = t$, we obtain $w \leq t$ in $\mathfrak{B}^*$. Combined with $w \neq t$ (since $w < t$ in $\mathfrak{A}^*$), this yields $w < t$ in $\mathfrak{B}^*$. In other words, $t > w$ in $\mathfrak{B}^*$.

Hence, there exists at least one $s \in u \sqcup\!\sqcup v$ satisfying $s > w$ in $\mathfrak{B}^*$ (namely, $s = t$). This proves (13.161.3).

(by the definition of $\shuffle$).

However, for every $\sigma \in \operatorname{Sh}_{n,m}$, we have $u \underset{\sigma}{\shuffle} v = \left(w_{\sigma(1)}, w_{\sigma(2)}, \ldots, w_{\sigma(n+m)}\right)$ (by the definition of $u \underset{\sigma}{\shuffle} v$), and therefore $b_{u \underset{\sigma}{\shuffle} v} = b_{w_{\sigma(1)}} b_{w_{\sigma(2)}} \cdots b_{w_{\sigma(n+m)}}$ (by the definition of $b_{u \underset{\sigma}{\shuffle} v}$). Hence, $\sum_{\sigma \in \operatorname{Sh}_{n,m}} \underbrace{b_{u \underset{\sigma}{\shuffle} v}}_{= b_{w_{\sigma(1)}} b_{w_{\sigma(2)}} \cdots b_{w_{\sigma(n+m)}}} = \sum_{\sigma \in \operatorname{Sh}_{n,m}} b_{w_{\sigma(1)}} b_{w_{\sigma(2)}} \cdots b_{w_{\sigma(n+m)}}$. Compared with (13.162.1), this yields $b_u \shuffle b_v = \sum_{\sigma \in \operatorname{Sh}_{n,m}} b_{u \underset{\sigma}{\shuffle} v}$. This proves Remark 6.3.2. □

13.163. **Solution to Exercise 6.3.8.** *Solution to Exercise 6.3.8.*

Proof of Lemma 6.3.7. (a) Let $\mathfrak{M}$ denote the set of all finite multisets of Lyndon words over $\mathfrak{A}$. In other words, $\mathfrak{M}$ is the set of all finite multisets of elements of $\mathfrak{L}$ (since the elements of $\mathfrak{L}$ are the Lyndon words over $\mathfrak{A}$). Define two maps $\mathbf{m} : \mathfrak{M} \to \mathfrak{A}^*$ and $\mathbf{n} : \mathfrak{A}^* \to \mathfrak{M}$ as in Proposition 13.143.1. Then, Proposition 13.143.1 shows that these two maps $\mathbf{m}$ and $\mathbf{n}$ are mutually inverse bijections. Hence, the map $\mathbf{m}$ is a bijection.

On the other hand, for every $M \in \mathfrak{M}$, let us define an element $\mathbf{b}_M$ of A by setting $\mathbf{b}_M = b_{a_1} b_{a_2} \cdots b_{a_k}$, where $a_1, a_2, \ldots, a_k$ denote the elements of M listed in decreasing order. Then, $(b_w)_{w \in \mathfrak{L}}$ is an algebraically independent generating set of the $\mathbf{k}$-algebra A if and only if $(\mathbf{b}_M)_{M \in \mathfrak{M}}$ is a basis of the $\mathbf{k}$-module A [1105].

[1105] *Proof.* Let $\mathfrak{N}$ be the set of all families $(k_w)_{w \in \mathfrak{L}} \in \mathbb{N}^{\mathfrak{L}}$ of nonnegative integers (indexed by the Lyndon words) such that all but finitely many $w \in \mathfrak{L}$ satisfy $k_w = 0$. Proposition 13.143.2 (applied to $S = \mathfrak{L}$) shows that the map $\operatorname{mult} : \mathfrak{M} \to \mathfrak{N}$ that sends each multiset $M \in \mathfrak{M}$ to the family

$$((\text{multiplicity of } w \text{ in the multiset } M))_{w \in S} \in \mathfrak{N}$$

is well-defined and is a bijection. Consider this map mult.

For every family $f = (k_w)_{w \in \mathfrak{L}} \in \mathfrak{N}$, we can define an element $\mathbf{b}_f$ of A by $\mathbf{b}_f = \prod_{w \in \mathfrak{L}} (b_w)^{k_w}$. (This is well-defined, since A is commutative and since all but finitely many $w \in \mathfrak{L}$ satisfy $k_w = 0$). Then, $\left(\mathbf{b}_f\right)_{f \in \mathfrak{N}}$ is the family of all possible monomials in the "variables" b_w (that is, of all possible finite products of elements of the family $(b_w)_{w \in \mathfrak{L}}$, with multiplicities allowed). Hence,

$$\left(\begin{array}{c} (b_w)_{w \in \mathfrak{L}} \text{ is an algebraically independent generating set of the } \mathbf{k}\text{-algebra } A \\ \text{if and only if } \left(\mathbf{b}_f\right)_{f \in \mathfrak{N}} \text{ is a basis of the } \mathbf{k}\text{-module } A \end{array} \right). \tag{13.163.1}$$

Now, we claim that the family $\left(\mathbf{b}_f\right)_{f \in \mathfrak{N}}$ is a reindexing of the family $(\mathbf{b}_M)_{M \in \mathfrak{M}}$. Indeed, since mult is a bijection, it is clear that the family $\left(\mathbf{b}_f\right)_{f \in \mathfrak{N}}$ is a reindexing of the family $(\mathbf{b}_{\operatorname{mult} M})_{M \in \mathfrak{M}}$. We now will prove that every $M \in \mathfrak{M}$ satisfies $\mathbf{b}_{\operatorname{mult} M} = \mathbf{b}_M$.

Indeed, let $M \in \mathfrak{M}$. Let $a_1, a_2, \ldots, a_k$ denote the elements of M listed in decreasing order. Then, every $w \in \mathfrak{L}$ satisfies

$$(\text{multiplicity of } w \text{ in the multiset } M) = (\text{the number of } i \in \{1, 2, \ldots, k\} \text{ satisfying } a_i = w). \tag{13.163.2}$$

But by the definition of $\operatorname{mult} M$, we have

$$\begin{aligned} \operatorname{mult} M &= \left(\underbrace{(\text{multiplicity of } w \text{ in the multiset } M)}_{\substack{= (\text{the number of } i \in \{1,2,\ldots,k\} \text{ satisfying } a_i = w) \\ (\text{by } (13.163.2))}} \right)_{w \in \mathfrak{L}} \\ &= ((\text{the number of } i \in \{1, 2, \ldots, k\} \text{ satisfying } a_i = w))_{w \in \mathfrak{L}}, \end{aligned}$$

so that the definition of $\mathbf{b}_{\operatorname{mult} M}$ becomes

$$\begin{aligned} \mathbf{b}_{\operatorname{mult} M} &= \prod_{w \in \mathfrak{L}} \underbrace{(b_w)^{(\text{the number of } i \in \{1,2,\ldots,k\} \text{ satisfying } a_i = w)}}_{= \prod_{\substack{i \in \{1,2,\ldots,k\}; \\ a_i = w}} b_w} = \prod_{w \in \mathfrak{L}} \prod_{\substack{i \in \{1,2,\ldots,k\}; \\ a_i = w}} \underbrace{b_w}_{\substack{= b_{a_i} \\ (\text{since } w = a_i)}} = \prod_{w \in \mathfrak{L}} \prod_{\substack{i \in \{1,2,\ldots,k\}; \\ a_i = w}} b_{a_i} \\ &= \prod_{i \in \{1,2,\ldots,k\}} b_{a_i} = b_{a_1} b_{a_2} \cdots b_{a_k} = \mathbf{b}_M \end{aligned}$$

(since $\mathbf{b}_M$ was defined as $b_{a_1} b_{a_2} \cdots b_{a_k}$).

Forget now that we fixed M. We thus have shown that every $M \in \mathfrak{M}$ satisfies $\mathbf{b}_{\operatorname{mult} M} = \mathbf{b}_M$. Hence, $(\mathbf{b}_{\operatorname{mult} M})_{M \in \mathfrak{M}} = (\mathbf{b}_M)_{M \in \mathfrak{M}}$. But we know that the family $\left(\mathbf{b}_f\right)_{f \in \mathfrak{N}}$ is a reindexing of the family $(\mathbf{b}_{\operatorname{mult} M})_{M \in \mathfrak{M}}$. Since $(\mathbf{b}_{\operatorname{mult} M})_{M \in \mathfrak{M}} = (\mathbf{b}_M)_{M \in \mathfrak{M}}$, this rewrites as follows: The family $\left(\mathbf{b}_f\right)_{f \in \mathfrak{N}}$ is a reindexing of the family $(\mathbf{b}_M)_{M \in \mathfrak{M}}$. Hence, $\left(\mathbf{b}_f\right)_{f \in \mathfrak{N}}$ is a basis

Hence, in order to prove Lemma 6.3.7(a), it remains to prove that $(\mathbf{b}_M)_{M\in\mathfrak{M}}$ is a basis of the $\mathbf{k}$-module A if and only if $(\mathbf{b}_u)_{u\in\mathfrak{A}^*}$ is a basis of the $\mathbf{k}$-module A.

Let us show that

$$\mathbf{b}_{\mathbf{m}(M)} = \mathbf{b}_M \qquad \text{for every } M \in \mathfrak{M}. \tag{13.163.3}$$

Proof of (13.163.3): Let $M \in \mathfrak{M}$. Let $a_1, a_2, \ldots, a_k$ denote the elements of M listed in decreasing order. Then, $\mathbf{m}(M) = a_1 a_2 \cdots a_k$ (by the definition of $\mathbf{m}(M)$). Combining this with the fact that a_1, a_2, $\ldots$, a_k are Lyndon words (since they are elements of M, which is a multiset of Lyndon words) and satisfy $a_1 \geq a_2 \geq \cdots \geq a_k$ (since $a_1, a_2, \ldots, a_k$ are the elements of M listed in decreasing order), we thus conclude that $(a_1, a_2, \ldots, a_k)$ is the CFL factorization of $\mathbf{m}(M)$. Hence, the definition of $\mathbf{b}_{\mathbf{m}(M)}$ says that $\mathbf{b}_{\mathbf{m}(M)} = b_{a_1} b_{a_2} \cdots b_{a_k}$. Compared with $\mathbf{b}_M = b_{a_1} b_{a_2} \cdots b_{a_k}$ (which is just the definition of $\mathbf{b}_M$), this yields $\mathbf{b}_{\mathbf{m}(M)} = \mathbf{b}_M$. This proves (13.163.3).

Now, the family $\left(\mathbf{b}_{\mathbf{m}(M)}\right)_{M\in\mathfrak{M}}$ is a reindexing of the family $(\mathbf{b}_u)_{u\in\mathfrak{A}^*}$ (since $\mathbf{m}$ is a bijection). Since $\left(\mathbf{b}_{\mathbf{m}(M)}\right)_{M\in\mathfrak{M}} = (\mathbf{b}_M)_{M\in\mathfrak{M}}$ (by (13.163.3)), this rewrites as follows: The family $(\mathbf{b}_M)_{M\in\mathfrak{M}}$ is a reindexing of the family $(\mathbf{b}_u)_{u\in\mathfrak{A}^*}$. Hence, $(\mathbf{b}_M)_{M\in\mathfrak{M}}$ is a basis of the $\mathbf{k}$-module A if and only if $(\mathbf{b}_u)_{u\in\mathfrak{A}^*}$ is a basis of the $\mathbf{k}$-module A. As we said above, this completes our proof of Lemma 6.3.7(a).

(b) The proof of Lemma 6.3.7(b) is analogous to the above proof of Lemma 6.3.7(a).

(c) We assumed that the family $(b_w)_{w\in\mathfrak{L}}$ generates the $\mathbf{k}$-algebra A. By Lemma 6.3.7(b), this yields that the family $(\mathbf{b}_u)_{u\in\mathfrak{A}^*}$ spans the $\mathbf{k}$-module A. Recall also that the family $(g_u)_{u\in\mathfrak{A}^*}$ is a basis of the $\mathbf{k}$-module A, and thus spans this $\mathbf{k}$-module.

We need to prove that the family $(b_w)_{w\in\mathfrak{L}}$ is an algebraically independent generating set of the $\mathbf{k}$-algebra A. According to Lemma 6.3.7(a), this is equivalent to proving that the family $(\mathbf{b}_u)_{u\in\mathfrak{A}^*}$ is a basis of the $\mathbf{k}$-module A. We are going to prove the latter statement.

Let us first notice that $\mathbf{b}_u$ is a homogeneous element of A of degree $\operatorname{Wt}(u)$ for every $u \in \mathfrak{A}^*$ [1106].

Now, let $n \in \mathbb{N}$. It is easy to see that the family $(\mathbf{b}_u)_{u\in\operatorname{Wt}^{-1}(n)}$ spans the $\mathbf{k}$-module A_n (that is, the n-th homogeneous component of the $\mathbf{k}$-module A) [1107]. The same argument (but with $\mathbf{b}_u$ replaced by g_u) shows that the family $(g_u)_{u\in\operatorname{Wt}^{-1}(n)}$ spans the $\mathbf{k}$-module A_n. Since this family $(g_u)_{u\in\operatorname{Wt}^{-1}(n)}$ is linearly

of the $\mathbf{k}$-module A if and only if $(\mathbf{b}_M)_{M\in\mathfrak{M}}$ is a basis of the $\mathbf{k}$-module A. Combined with (13.163.1), this yields that $(b_w)_{w\in\mathfrak{L}}$ is an algebraically independent generating set of the $\mathbf{k}$-algebra A if and only if $(\mathbf{b}_M)_{M\in\mathfrak{M}}$ is a basis of the $\mathbf{k}$-module A, qed.

[1106] *Proof.* Let $u \in \mathfrak{A}^*$. Let $(a_1, a_2, \ldots, a_p)$ be the CFL factorization of u. Then, $a_1, a_2, \ldots, a_p$ are Lyndon words and satisfy $u = a_1 a_2 \cdots a_p$.

The definition of Wt easily yields that

$$\operatorname{Wt}(s_1 s_2 \cdots s_k) = \operatorname{Wt}(s_1) + \operatorname{Wt}(s_2) + \cdots + \operatorname{Wt}(s_k) \qquad \text{for any } k \in \mathbb{N} \text{ and any } k \text{ words } s_1, s_2, \ldots, s_k \text{ in } \mathfrak{A}^*.$$

Applying this to $k = p$ and $s_i = a_i$, we obtain $\operatorname{Wt}(a_1 a_2 \cdots a_p) = \operatorname{Wt}(a_1) + \operatorname{Wt}(a_2) + \cdots + \operatorname{Wt}(a_p)$, so that

$$\operatorname{Wt}(a_1) + \operatorname{Wt}(a_2) + \cdots + \operatorname{Wt}(a_p) = \operatorname{Wt}\left(\underbrace{a_1 a_2 \cdots a_p}_{=u}\right) = \operatorname{Wt}(u).$$

But the definition of $\mathbf{b}_u$ yields $\mathbf{b}_u = b_{a_1} b_{a_2} \cdots b_{a_p}$. Thus, the element $\mathbf{b}_u$ is homogeneous of degree $\operatorname{Wt}(a_1) + \operatorname{Wt}(a_2) + \cdots + \operatorname{Wt}(a_p)$ (since for every $w \in \mathfrak{L}$, the element b_w of A is homogeneous of degree $\operatorname{Wt}(w)$). Since $\operatorname{Wt}(a_1) + \operatorname{Wt}(a_2) + \cdots + \operatorname{Wt}(a_p) = \operatorname{Wt}(u)$, this rewrites as follows: The element $\mathbf{b}_u$ is homogeneous of degree $\operatorname{Wt}(u)$, qed.

[1107] *Proof.* For every $u \in \operatorname{Wt}^{-1}(n)$, the element $\mathbf{b}_u$ is a homogeneous element of A of degree $\operatorname{Wt}(u) = n$ (since $u \in \operatorname{Wt}^{-1}(n)$). In other words, for every $u \in \operatorname{Wt}^{-1}(n)$, we have $\mathbf{b}_u \in A_n$.

Let $\xi \in A_n$. Then, $\xi \in A_n \subset A$, so that ξ is a $\mathbf{k}$-linear combination of the elements $\mathbf{b}_u$ for $u \in \mathfrak{A}^*$ (since the family $(\mathbf{b}_u)_{u\in\mathfrak{A}^*}$ spans the $\mathbf{k}$-module A). In other words, $\xi = \sum_{u\in\mathfrak{A}^*} \lambda_u \mathbf{b}_u$ for some family $(\lambda_u)_{u\in\mathfrak{A}^*} \in \mathbf{k}^{\mathfrak{A}^*}$ of elements of $\mathbf{k}$ such that all but finitely many $u \in \mathfrak{A}^*$ satisfy $\lambda_u = 0$. Consider this $(\lambda_u)_{u\in\mathfrak{A}^*}$.

What happens if we apply the canonical projection $A \to A_n$ (which projects A onto its n-th homogeneous component A_n, annihilating all other components) to both sides of the equality $\xi = \sum_{u\in\mathfrak{A}^*} \lambda_u \mathbf{b}_u$? The left hand side remains ξ (since $\xi \in A_n$). On the right hand side, all addends of the sum in which the u satisfies $\operatorname{Wt}(u) = n$ stay fixed (because we know that for each such u, the element $\mathbf{b}_u$ is a homogeneous element of A of degree $\operatorname{Wt}(u) = n$), whereas all other addends become 0 (for a similar reason); therefore, the sum on the right hand side becomes $\sum_{\substack{u\in\mathfrak{A}^*;\\ \operatorname{Wt}(u)=n}} \lambda_u \mathbf{b}_u$. Thus, the equality becomes $\xi = \sum_{\substack{u\in\mathfrak{A}^*;\\ \operatorname{Wt}(u)=n}} \lambda_u \mathbf{b}_u$. Thus,

$$\xi = \sum_{\substack{u\in\mathfrak{A}^*;\\ \operatorname{Wt}(u)=n}} \lambda_u \mathbf{b}_u = \sum_{u\in\operatorname{Wt}^{-1}(n)} \lambda_u \mathbf{b}_u.$$

Hence, ξ is a $\mathbf{k}$-linear combination of the elements $\mathbf{b}_u$ with $u \in \operatorname{Wt}^{-1}(n)$.

independent (because it is a subfamily of the basis $(g_u)_{u\in\mathfrak{A}^*}$ of A), this shows that the family $(g_u)_{u\in\operatorname{Wt}^{-1}(n)}$ is a basis of the $\mathbf{k}$-module A_n.

But the set $\operatorname{Wt}^{-1}(n)$ is finite[1108]. Hence, A_n is a finite free $\mathbf{k}$-module (since $(g_u)_{u\in\operatorname{Wt}^{-1}(n)}$ is a basis of the $\mathbf{k}$-module A_n), and Exercise 2.5.18(b) (applied to A_n, $\operatorname{Wt}^{-1}(n)$, u, g_u and $\mathbf{b}_u$ instead of A, I, i, γ_i and β_i) yields that $(\mathbf{b}_u)_{u\in\operatorname{Wt}^{-1}(n)}$ is a $\mathbf{k}$-basis of A_n.

Now, let us forget that we fixed n. We thus have shown that for every $n\in\mathbb{N}$, the family $(\mathbf{b}_u)_{u\in\operatorname{Wt}^{-1}(n)}$ is a $\mathbf{k}$-basis of A_n. Hence, the family $(\mathbf{b}_u)_{u\in\mathfrak{A}^*}$ (being the disjoint union of the families $(\mathbf{b}_u)_{u\in\operatorname{Wt}^{-1}(n)}$ over all $n\in\mathbb{N}$) is a $\mathbf{k}$-basis of $\bigoplus_{n\in\mathbb{N}} A_n = A$. In other words, the family $(\mathbf{b}_u)_{u\in\mathfrak{A}^*}$ is a basis of the $\mathbf{k}$-module A. As we know, this completes the proof of Lemma 6.3.7(c). □

13.164. **Solution to Exercise 6.3.11.** *Solution to Exercise 6.3.11.*

Proof of Lemma 6.3.10. (a) Let u, v and v' be three words satisfying $\ell(u)=n$, $\ell(v)=m$, $\ell(v')=m$ and $v'<v$. We must show that $u \underset{\sigma}{\sqcup\!\sqcup} v' < u \underset{\sigma}{\sqcup\!\sqcup} v$.

We have $v'<v$, thus $v'\leq v$. By the definition of the relation $\leq$, this means that

$$\textbf{either} \text{ there exists an } i\in\{1,2,\ldots,\min\{\ell(v'),\ell(v)\}\}$$
$$\text{such that } \left((v')_i < v_i, \text{ and every } j\in\{1,2,\ldots,i-1\} \text{ satisfies } (v')_j = v_j\right),$$
$$\textbf{or} \text{ the word } v' \text{ is a prefix of } v.$$

Since the word v' cannot be a prefix of v (because this would entail that $v'=v$ (because $\ell(v')=m=\ell(v)$, so that v' has the same length as v), which would contradict $v'<v$), this shows that there exists an $i\in\{1,2,\ldots,\min\{\ell(v'),\ell(v)\}\}$ such that $\left((v')_i<v_i, \text{ and every } j\in\{1,2,\ldots,i-1\} \text{ satisfies } (v')_j=v_j\right)$. Denote this i by k. Thus, $k\in\{1,2,\ldots,\min\{\ell(v'),\ell(v)\}\}$ has the property that $(v')_k<v_k$, and

$$\text{every } j\in\{1,2,\ldots,k-1\} \text{ satisfies } (v')_j=v_j. \tag{13.164.1}$$

Now, let us forget that we fixed ξ. We thus have shown that every $\xi\in A_n$ is a $\mathbf{k}$-linear combination of the elements $\mathbf{b}_u$ with $u\in\operatorname{Wt}^{-1}(n)$. Since all these elements $\mathbf{b}_u$ belong to A_n (because for every $u\in\operatorname{Wt}^{-1}(n)$, we have $\mathbf{b}_u\in A_n$), this shows that the family $(\mathbf{b}_u)_{u\in\operatorname{Wt}^{-1}(n)}$ spans the $\mathbf{k}$-module A_n, qed.

[1108] *Proof.* We have

$$\begin{aligned}
\operatorname{Wt}^{-1}(n) &= \{w\in\mathfrak{A}^* \mid \operatorname{Wt}(w)=n\}\\
&= \bigcup_{k\in\mathbb{N}} \left\{ (w_1,w_2,\ldots,w_k)\in\mathfrak{A}^* \mid \underbrace{\operatorname{Wt}((w_1,w_2,\ldots,w_k))}_{\substack{=\operatorname{wt}(w_1)+\operatorname{wt}(w_2)+\cdots+\operatorname{wt}(w_k)\\ \text{(by the definition of } \operatorname{Wt}((w_1,w_2,\ldots,w_k)))}} = n \right\}\\
&\qquad \left(\text{since every word } w\in\mathfrak{A}^* \text{ has the form } (w_1,w_2,\ldots,w_k) \text{ for some } k\in\mathbb{N}\right)\\
&= \bigcup_{k\in\mathbb{N}} \underbrace{\{(w_1,w_2,\ldots,w_k)\in\mathfrak{A}^* \mid \operatorname{wt}(w_1)+\operatorname{wt}(w_2)+\cdots+\operatorname{wt}(w_k)=n\}}_{=\bigcup_{\substack{(i_1,i_2,\ldots,i_k)\in\{1,2,3,\ldots\}^k;\\ i_1+i_2+\cdots+i_k=n}} \{(w_1,w_2,\ldots,w_k)\in\mathfrak{A}^* \mid \operatorname{wt}(w_1)=i_1,\ \operatorname{wt}(w_2)=i_2,\ \ldots,\ \operatorname{wt}(w_k)=i_k\}}\\
&= \underbrace{\bigcup_{k\in\mathbb{N}} \bigcup_{\substack{(i_1,i_2,\ldots,i_k)\in\{1,2,3,\ldots\}^k;\\ i_1+i_2+\cdots+i_k=n}}}_{=\bigcup_{(i_1,i_2,\ldots,i_k)\in\operatorname{Comp}_n}} \underbrace{\{(w_1,w_2,\ldots,w_k)\in\mathfrak{A}^* \mid \operatorname{wt}(w_1)=i_1,\ \operatorname{wt}(w_2)=i_2,\ \ldots,\ \operatorname{wt}(w_k)=i_k\}}_{=\operatorname{wt}^{-1}(i_1)\times\operatorname{wt}^{-1}(i_2)\times\cdots\times\operatorname{wt}^{-1}(i_k)}\\
&= \bigcup_{(i_1,i_2,\ldots,i_k)\in\operatorname{Comp}_n} \underbrace{\operatorname{wt}^{-1}(i_1)\times\operatorname{wt}^{-1}(i_2)\times\cdots\times\operatorname{wt}^{-1}(i_k)}_{\substack{\text{a finite set}\\ \text{(since for every } N\in\{1,2,3,\ldots\}\text{, the set } \operatorname{wt}^{-1}(N) \text{ is finite)}}} .
\end{aligned}$$

Hence, $\operatorname{Wt}^{-1}(n)$ is a union of finitely many finite sets, and thus itself finite, qed.

We have $k \in \left\{1, 2, \ldots, \min\left\{\underbrace{\ell\left(v'\right)}_{=m}, \underbrace{\ell\left(v\right)}_{=m}\right\}\right\} = \left\{1, 2, \ldots, \underbrace{\min\left\{m, m\right\}}_{=m}\right\} = \{1, 2, \ldots, m\}$.

Now, let $(w_1, w_2, \ldots, w_{n+m})$ denote the concatenation $u \cdot v = (u_1, u_2, \ldots, u_n, v_1, v_2, \ldots, v_m)$, and let $\left(w'_1, w'_2, \ldots, w'_{n+m}\right)$ denote the concatenation $u \cdot v' = \left(u_1, u_2, \ldots, u_n, (v')_1, (v')_2, \ldots, (v')_m\right)$. We have $u \underset{\sigma}{\sqcup\!\sqcup} v = \left(w_{\sigma(1)}, w_{\sigma(2)}, \ldots, w_{\sigma(n+m)}\right)$ (by the definition of $u \underset{\sigma}{\sqcup\!\sqcup} v$) and $u \underset{\sigma}{\sqcup\!\sqcup} v' = \left(w'_{\sigma(1)}, w'_{\sigma(2)}, \ldots, w'_{\sigma(n+m)}\right)$ (by the definition of $u \underset{\sigma}{\sqcup\!\sqcup} v'$). Hence, we have $u \underset{\sigma}{\sqcup\!\sqcup} v' \neq u \underset{\sigma}{\sqcup\!\sqcup} v$ [1109]. But our goal is to prove that $u \underset{\sigma}{\sqcup\!\sqcup} v' < u \underset{\sigma}{\sqcup\!\sqcup} v$. Hence, it remains to prove that $u \underset{\sigma}{\sqcup\!\sqcup} v' \leq u \underset{\sigma}{\sqcup\!\sqcup} v$ (since we have shown that $u \underset{\sigma}{\sqcup\!\sqcup} v' \neq u \underset{\sigma}{\sqcup\!\sqcup} v$). Due to the definition of the relation $\leq$, this amounts to proving that

> **either** there exists an $i \in \left\{1, 2, \ldots, \min\left\{\ell\left(u \underset{\sigma}{\sqcup\!\sqcup} v'\right), \ell\left(u \underset{\sigma}{\sqcup\!\sqcup} v\right)\right\}\right\}$ such that
> $$\left(\left(u \underset{\sigma}{\sqcup\!\sqcup} v'\right)_i < \left(u \underset{\sigma}{\sqcup\!\sqcup} v\right)_i, \text{ and every } j \in \{1, 2, \ldots, i-1\} \text{ satisfies } \left(u \underset{\sigma}{\sqcup\!\sqcup} v'\right)_j = \left(u \underset{\sigma}{\sqcup\!\sqcup} v\right)_j\right),$$
> **or** the word $u \underset{\sigma}{\sqcup\!\sqcup} v'$ is a prefix of $u \underset{\sigma}{\sqcup\!\sqcup} v$.

We shall prove that the first of these two alternatives holds, i.e., that there exists an

$$i \in \left\{1, 2, \ldots, \min\left\{\ell\left(u \underset{\sigma}{\sqcup\!\sqcup} v'\right), \ell\left(u \underset{\sigma}{\sqcup\!\sqcup} v\right)\right\}\right\}$$

such that

$$\left(\left(u \underset{\sigma}{\sqcup\!\sqcup} v'\right)_i < \left(u \underset{\sigma}{\sqcup\!\sqcup} v\right)_i, \text{ and every } j \in \{1, 2, \ldots, i-1\} \text{ satisfies } \left(u \underset{\sigma}{\sqcup\!\sqcup} v'\right)_j = \left(u \underset{\sigma}{\sqcup\!\sqcup} v\right)_j\right).$$

Indeed, we claim that $\sigma^{-1}(n+k)$ is such an i. In order to conclude the proof, we then need to show that

$$\sigma^{-1}(n+k) \in \left\{1, 2, \ldots, \min\left\{\ell\left(u \underset{\sigma}{\sqcup\!\sqcup} v'\right), \ell\left(u \underset{\sigma}{\sqcup\!\sqcup} v\right)\right\}\right\}, \tag{13.164.2}$$

that we have

$$\left(u \underset{\sigma}{\sqcup\!\sqcup} v'\right)_{\sigma^{-1}(n+k)} < \left(u \underset{\sigma}{\sqcup\!\sqcup} v\right)_{\sigma^{-1}(n+k)}, \tag{13.164.3}$$

and that

$$\text{every } j \in \left\{1, 2, \ldots, \sigma^{-1}(n+k) - 1\right\} \text{ satisfies } \left(u \underset{\sigma}{\sqcup\!\sqcup} v'\right)_j = \left(u \underset{\sigma}{\sqcup\!\sqcup} v\right)_j. \tag{13.164.4}$$

Proof of (13.164.2): We have $\ell\left(u \underset{\sigma}{\sqcup\!\sqcup} v\right) = \underbrace{\ell(u)}_{=n} + \underbrace{\ell(v)}_{=m} = n + m$ and similarly $\ell\left(u \underset{\sigma}{\sqcup\!\sqcup} v'\right) = n + m$. Thus, $\min\left\{\underbrace{\ell\left(u \underset{\sigma}{\sqcup\!\sqcup} v'\right)}_{=n+m}, \underbrace{\ell\left(u \underset{\sigma}{\sqcup\!\sqcup} v\right)}_{=n+m}\right\} = \min\{n+m, n+m\} = n + m$. But since $\sigma \in \operatorname{Sh}_{n,m} \subset \mathfrak{S}_{n+m}$, we have $\sigma^{-1}(n+k) \in \{1, 2, \ldots, n+m\}$. In other words, $\sigma^{-1}(n+k) \in \left\{1, 2, \ldots, \min\left\{\ell\left(u \underset{\sigma}{\sqcup\!\sqcup} v'\right), \ell\left(u \underset{\sigma}{\sqcup\!\sqcup} v\right)\right\}\right\}$ (since $\min\left\{\ell\left(u \underset{\sigma}{\sqcup\!\sqcup} v'\right), \ell\left(u \underset{\sigma}{\sqcup\!\sqcup} v\right)\right\} = n + m$). This proves (13.164.2).

Proof of (13.164.3): Since $u \underset{\sigma}{\sqcup\!\sqcup} v = \left(w_{\sigma(1)}, w_{\sigma(2)}, \ldots, w_{\sigma(n+m)}\right)$, we have $\left(u \underset{\sigma}{\sqcup\!\sqcup} v\right)_j = w_{\sigma(j)}$ for every $j \in \{1, 2, \ldots, n+m\}$. Applying this to $j = \sigma^{-1}(n+k)$, we obtain $\left(u \underset{\sigma}{\sqcup\!\sqcup} v\right)_{\sigma^{-1}(n+k)} = w_{\sigma(\sigma^{-1}(n+k))} = w_{n+k} = v_k$ (since $k \in \{1, 2, \ldots, m\}$ and $(w_1, w_2, \ldots, w_{n+m}) = (u_1, u_2, \ldots, u_n, v_1, v_2, \ldots, v_m)$). Similarly,

[1109] *Proof.* Assume the contrary. Then, $u \underset{\sigma}{\sqcup\!\sqcup} v' = u \underset{\sigma}{\sqcup\!\sqcup} v$, thus $\left(w'_{\sigma(1)}, w'_{\sigma(2)}, \ldots, w'_{\sigma(n+m)}\right) = u \underset{\sigma}{\sqcup\!\sqcup} v' = u \underset{\sigma}{\sqcup\!\sqcup} v = \left(w_{\sigma(1)}, w_{\sigma(2)}, \ldots, w_{\sigma(n+m)}\right)$. Hence, $\left(w'_1, w'_2, \ldots, w'_{n+m}\right) = (w_1, w_2, \ldots, w_{n+m})$ (since σ is a permutation), and thus $u \cdot v' = \left(w'_1, w'_2, \ldots, w'_{n+m}\right) = (w_1, w_2, \ldots, w_{n+m}) = u \cdot v$, so that $v' = v$ (since u can be cancelled from the equality $u \cdot v' = u \cdot v$). But this contradicts $v' < v$. This contradiction shows that our assumption was wrong, qed.

$\left(u \underset{\sigma}{\sqcup\!\sqcup} v'\right)_{\sigma^{-1}(n+k)} = (v')_k$. Recalling that $(v')_k < v_k$, we now see that $\left(u \underset{\sigma}{\sqcup\!\sqcup} v'\right)_{\sigma^{-1}(n+k)} = (v')_k < v_k = \left(u \underset{\sigma}{\sqcup\!\sqcup} v\right)_{\sigma^{-1}(n+k)}$. This proves (13.164.3).

Proof of (13.164.4): Let $j \in \left\{1, 2, \ldots, \sigma^{-1}(n+k) - 1\right\}$. We need to prove that $\left(u \underset{\sigma}{\sqcup\!\sqcup} v'\right)_j = \left(u \underset{\sigma}{\sqcup\!\sqcup} v\right)_j$.

Assume the contrary. Then, $\left(u \underset{\sigma}{\sqcup\!\sqcup} v'\right)_j \neq \left(u \underset{\sigma}{\sqcup\!\sqcup} v\right)_j$.

We have $j \in \{1, 2, \ldots, n+m\}$ and $j < \sigma^{-1}(n+k)$ (since $j \in \left\{1, 2, \ldots, \sigma^{-1}(n+k) - 1\right\}$).

Since $u \underset{\sigma}{\sqcup\!\sqcup} v = \left(w_{\sigma(1)}, w_{\sigma(2)}, \ldots, w_{\sigma(n+m)}\right)$, we have $\left(u \underset{\sigma}{\sqcup\!\sqcup} v\right)_j = w_{\sigma(j)}$. Similarly, $\left(u \underset{\sigma}{\sqcup\!\sqcup} v'\right)_j = w'_{\sigma(j)}$. Hence, $w'_{\sigma(j)} = \left(u \underset{\sigma}{\sqcup\!\sqcup} v'\right)_j \neq \left(u \underset{\sigma}{\sqcup\!\sqcup} v\right)_j = w_{\sigma(j)}$.

If we had $\sigma(j) \leq n$, then we would have $w_{\sigma(j)} = u_{\sigma(j)}$ (since $(w_1, w_2, \ldots, w_{n+m}) = (u_1, u_2, \ldots, u_n, v_1, v_2, \ldots, v_m)$) and $w'_{\sigma(j)} = u_{\sigma(j)}$ (for similar reasons), which would yield that $w'_{\sigma(j)} = u_{\sigma(j)} = w_{\sigma(j)}$, which would contradict $w'_{\sigma(j)} \neq w_{\sigma(j)}$. Hence, we cannot have $\sigma(j) \leq n$. We thus must have $\sigma(j) > n$, whence $\sigma(j) \in \{n+1, n+2, \ldots, n+m\}$. Consequently, $w_{\sigma(j)} = v_{\sigma(j)-n}$ (because $(w_1, w_2, \ldots, w_{n+m}) = (u_1, u_2, \ldots, u_n, v_1, v_2, \ldots, v_m)$) and $(w')_{\sigma(j)} = (v')_{\sigma(j)-n}$ (similarly).

If we had $\sigma(j) - n \in \{1, 2, \ldots, k-1\}$, then we would have $(v')_{\sigma(j)-n} = v_{\sigma(j)-n}$ (by (13.164.1), applied to $\sigma(j) - n$ instead of j). Hence, if we had $\sigma(j) - n \in \{1, 2, \ldots, k-1\}$, then we would have $(w')_{\sigma(j)} = (v')_{\sigma(j)-n} = v_{\sigma(j)-n} = w_{\sigma(j)}$, which would contradict $w'_{\sigma(j)} \neq w_{\sigma(j)}$. Hence, we cannot have $\sigma(j) - n \in \{1, 2, \ldots, k-1\}$. In other words, we have $\sigma(j) - n \notin \{1, 2, \ldots, k-1\}$, so that $\sigma(j) - n \geq k$. Hence, $\sigma(j) \geq n+k$.

We have $\sigma \in \mathrm{Sh}_{n,m}$, so that $\sigma^{-1}(1) < \sigma^{-1}(2) < \cdots < \sigma^{-1}(n)$ and $\sigma^{-1}(n+1) < \sigma^{-1}(n+2) < \cdots < \sigma^{-1}(n+m)$. In particular, the restriction of the map σ^{-1} to the set $\{n+1, n+2, \ldots, n+m\}$ is strictly increasing (since $\sigma^{-1}(n+1) < \sigma^{-1}(n+2) < \cdots < \sigma^{-1}(n+m)$). Since $\sigma(j)$ and $n+k$ both lie in this set $\{n+1, n+2, \ldots, n+m\}$, we thus have $\sigma^{-1}(\sigma(j)) \geq \sigma^{-1}(n+k)$ (because $\sigma(j) \geq n+k$). Hence, $j = \sigma^{-1}(\sigma(j)) \geq \sigma^{-1}(n+k)$, which contradicts $j < \sigma^{-1}(n+k)$. This contradiction shows that our assumption was wrong. Hence, (13.164.4) is proven.

Now that (13.164.2), (13.164.3) and (13.164.4) are all proven, we conclude that there exists an

$$i \in \left\{1, 2, \ldots, \min\left\{\ell\left(u \underset{\sigma}{\sqcup\!\sqcup} v'\right), \ell\left(u \underset{\sigma}{\sqcup\!\sqcup} v\right)\right\}\right\}$$

such that

$$\left(\left(u \underset{\sigma}{\sqcup\!\sqcup} v'\right)_i < \left(u \underset{\sigma}{\sqcup\!\sqcup} v\right)_i, \text{ and every } j \in \{1, 2, \ldots, i-1\} \text{ satisfies } \left(u \underset{\sigma}{\sqcup\!\sqcup} v'\right)_j = \left(u \underset{\sigma}{\sqcup\!\sqcup} v\right)_j\right)$$

(namely, $i = \sigma^{-1}(n+k)$). This concludes the proof of Lemma 6.3.10(a).

(b) The proof of Lemma 6.3.10(b) is similar to the proof of Lemma 6.3.10(a) above. (One of the changes necessary is to replace $\sigma^{-1}(n+k)$ by $\sigma^{-1}(k)$.)

(Alternatively, it is not hard to derive Lemma 6.3.10(b) from Lemma 6.3.10(a), because if $\tau \in \mathfrak{S}_{n+m}$ denotes the permutation which is written $(m+1, m+2, \ldots, n+m, 1, 2, \ldots, m)$ in one-line notation, then the permutation $\tau \circ \sigma$ belongs to $\mathrm{Sh}_{m,n}$ and satisfies $u \underset{\sigma}{\sqcup\!\sqcup} v = v \underset{\tau\circ\sigma}{\sqcup\!\sqcup} u$ and $u' \underset{\sigma}{\sqcup\!\sqcup} v = v \underset{\tau\circ\sigma}{\sqcup\!\sqcup} u'$, and therefore we can obtain Lemma 6.3.10(b) by applying Lemma 6.3.10(a) to m, n, v, u, u' and $\tau \circ \sigma$ instead of n, m, u, v, v' and σ.)

(c) Let u, v and v' be three words satisfying $\ell(u) = n$, $\ell(v) = m$, $\ell(v') = m$ and $v' \leq v$. We must show that $u \underset{\sigma}{\sqcup\!\sqcup} v' \leq u \underset{\sigma}{\sqcup\!\sqcup} v$. This is obvious if $v' = v$ (in fact, if $v' = v$, then $u \underset{\sigma}{\sqcup\!\sqcup} v' = u \underset{\sigma}{\sqcup\!\sqcup} v \leq u \underset{\sigma}{\sqcup\!\sqcup} v$). Hence, we can WLOG assume that $v' \neq v$. Assuming this, we immediately obtain $v' < v$ (since $v' \neq v$ and $v' \leq v$), so that $u \underset{\sigma}{\sqcup\!\sqcup} v' < u \underset{\sigma}{\sqcup\!\sqcup} v$ (by Lemma 6.3.10(a)). Thus, of course, $u \underset{\sigma}{\sqcup\!\sqcup} v' \leq u \underset{\sigma}{\sqcup\!\sqcup} v$, and Lemma 6.3.10(c) is proven. □

13.165. **Solution to Exercise 6.3.12.** *Solution to Exercise 6.3.12.*

Proof of Proposition 6.3.9. We shall prove Proposition 6.3.9 by strong induction over ℓ. So we fix some $L \in \mathbb{N}$, and we assume that Proposition 6.3.9 holds whenever $\ell < L$. We now need to prove that Proposition

6.3.9 holds for $\ell = L$. In other words, we need to prove that for every word $x \in \mathfrak{A}^L$, there is a family $(\eta_{x,y})_{y\in\mathfrak{A}^L} \in \mathbb{N}^{\mathfrak{A}^L}$ of elements of $\mathbb{N}$ satisfying

$$\mathbf{b}_x = \sum_{\substack{y\in\mathfrak{A}^L;\\ y\leq x}} \eta_{x,y} b_y \tag{13.165.1}$$

and $\eta_{x,x} \neq 0$ (in $\mathbb{N}$).

Let $x \in \mathfrak{A}^L$ be a word. We need to prove that there is a family $(\eta_{x,y})_{y\in\mathfrak{A}^L} \in \mathbb{N}^{\mathfrak{A}^L}$ of elements of $\mathbb{N}$ satisfying (13.165.1) and $\eta_{x,x} \neq 0$.

Let $(a_1, a_2, \ldots, a_p)$ be the CFL factorization of x. Then, a_1, a_2, $\ldots$, a_p are Lyndon words satisfying $x = a_1 a_2 \cdots a_p$ and $a_1 \geq a_2 \geq \cdots \geq a_p$.

If x is the empty word, then our claim is trivial (in fact, we can just set $\eta_{x,x} = 1$ in this case, and (13.165.1) holds obviously). Hence, for the rest of this proof, we WLOG assume that x is not the empty word. Thus, $p \neq 0$ (because otherwise, $x = a_1 a_2 \cdots a_p$ would be an empty product and thus the empty word, contradicting the assumption that x is not the empty word). Hence, we can define two words $u \in \mathfrak{A}^*$ and $v \in \mathfrak{A}^*$ by $u = a_1$ and $v = a_2 a_3 \cdots a_p$. The word u is Lyndon (since $u = a_1$ and since a_1 is Lyndon), and thus has (u) as its CFL factorization. The CFL factorization of the word v is $(a_2, a_3, \ldots, a_p)$ (since the words a_2, a_3, $\ldots$, a_p are Lyndon and satisfy $v = a_2 a_3 \cdots a_p$ and $a_2 \geq a_3 \geq \cdots \geq a_p$ (because $a_1 \geq a_2 \geq \cdots \geq a_p$)). Also, $u = a_1 \geq a_{j+1}$ for every $i \in \{1, 2, \ldots, 1\}$ and $j \in \{1, 2, \ldots, p-1\}$ (because $a_1 \geq a_2 \geq \cdots \geq a_p$). Hence, we can apply Theorem 6.2.2(c) to 1, (u), $p-1$ and $(a_2, a_3, \ldots, a_p)$ instead of p, $(a_1, a_2, \ldots, a_p)$, q and $(b_1, b_2, \ldots, b_q)$. As a result, we conclude that the lexicographically highest element of the multiset $u \sqcup\!\sqcup v$ is $\underbrace{u}_{=a_1}\underbrace{v}_{=a_2a_3\cdots a_p} = a_1 (a_2 a_3 \cdots a_p) = a_1 a_2 \cdots a_p = x$.

But $\mathbf{b}_u = b_u$ (by the definition of $\mathbf{b}_u$, since u has CFL factorization (u)), and $\mathbf{b}_x = \mathbf{b}_u \sqcup\!\sqcup \mathbf{b}_v$ [1110].

The word u is Lyndon and thus nonempty, so that $\ell(u) > 0$. Now, $\ell\left(\underbrace{x}_{=uv}\right) = \ell(uv) = \underbrace{\ell(u)}_{>0} + \ell(v) > \ell(v)$, so that $\ell(v) < \ell(x) = L$ (since $x \in \mathfrak{A}^L$). Hence, the induction hypothesis tells us that we can apply Proposition 6.3.9 to $\ell(v)$ and v instead of ℓ and x (since $v \in \mathfrak{A}^{\ell(v)}$). As a result, we see that there is a family $(\eta_{v,y})_{y\in\mathfrak{A}^{\ell(v)}} \in \mathbb{N}^{\mathfrak{A}^{\ell(v)}}$ of elements of $\mathbb{N}$ satisfying

$$\mathbf{b}_v = \sum_{\substack{y\in\mathfrak{A}^{\ell(v)};\\ y\leq v}} \eta_{v,y} b_y$$

and $\eta_{v,v} \neq 0$ (in $\mathbb{N}$). Consider this family $(\eta_{v,y})_{y\in\mathfrak{A}^{\ell(v)}}$.

Now,

$$\begin{aligned}
\mathbf{b}_x &= \underbrace{\mathbf{b}_u}_{=b_u} \sqcup\!\sqcup \underbrace{\mathbf{b}_v}_{=\sum_{\substack{y\in\mathfrak{A}^{\ell(v)};\\ y\leq v}} \eta_{v,y} b_y} = b_u \sqcup\!\sqcup \left(\sum_{\substack{y\in\mathfrak{A}^{\ell(v)};\\ y\leq v}} \eta_{v,y} b_y \right) = \sum_{\substack{y\in\mathfrak{A}^{\ell(v)};\\ y\leq v}} \eta_{v,y} \underbrace{b_u \sqcup\!\sqcup b_y}_{\substack{=\sum_{\sigma\in\mathrm{Sh}_{\ell(u),\ell(v)}} b_{u \underset{\sigma}{\sqcup\!\sqcup} y} \\ \text{(by Remark 6.3.2, applied to} \\ y,\ \ell(u) \text{ and } \ell(v) \\ \text{instead of } v,\ n \text{ and } m)}} \\
&= \sum_{\substack{y\in\mathfrak{A}^{\ell(v)};\\ y\leq v}} \eta_{v,y} \sum_{\sigma\in\mathrm{Sh}_{\ell(u),\ell(v)}} b_{u \underset{\sigma}{\sqcup\!\sqcup} y} = \sum_{\substack{z\in\mathfrak{A}^{\ell(v)};\\ z\leq v}} \eta_{v,z} \sum_{\sigma\in\mathrm{Sh}_{\ell(u),\ell(v)}} b_{u \underset{\sigma}{\sqcup\!\sqcup} z} \qquad (13.165.2)
\end{aligned}$$

[1110] *Proof.* Since the CFL factorization of v is $(a_2, a_3, \ldots, a_p)$, we have $\mathbf{b}_v = b_{a_2} \sqcup\!\sqcup b_{a_3} \sqcup\!\sqcup \cdots \sqcup\!\sqcup b_{a_p}$ (by the definition of $\mathbf{b}_v$). But since the CFL factorization of x is $(a_1, a_2, \ldots, a_p)$, we have

$$\begin{aligned}
\mathbf{b}_x &= b_{a_1} \sqcup\!\sqcup b_{a_2} \sqcup\!\sqcup \cdots \sqcup\!\sqcup b_{a_p} \qquad \text{(by the definition of } \mathbf{b}_x\text{)} \\
&= \underbrace{b_{a_1}}_{\substack{=b_u \\ \text{(since } a_1=u)}} \sqcup\!\sqcup \underbrace{\left(b_{a_2} \sqcup\!\sqcup b_{a_3} \sqcup\!\sqcup \cdots \sqcup\!\sqcup b_{a_p}\right)}_{=\mathbf{b}_v} = \underbrace{b_u}_{\substack{=\mathbf{b}_u \\ \text{(since } \mathbf{b}_u=b_u)}} \sqcup\!\sqcup \mathbf{b}_v = \mathbf{b}_u \sqcup\!\sqcup \mathbf{b}_v,
\end{aligned}$$

qed.

(here, we renamed the summation index y as z).

Now, let $z \in \mathfrak{A}^{\ell(v)}$ and $\sigma \in \operatorname{Sh}_{\ell(u),\ell(v)}$ be such that $z \leq v$. We will prove that $u \underset{\sigma}{\sqcup\!\sqcup} z \in \mathfrak{A}^L$ and $u \underset{\sigma}{\sqcup\!\sqcup} z \leq x$. First of all, it is clear that $\ell\left(u \underset{\sigma}{\sqcup\!\sqcup} z\right) = \ell(u) + \underbrace{\ell(z)}_{\substack{=\ell(v) \\ (\text{since } z \in \mathfrak{A}^{\ell(v)})}} = \ell(u) + \ell(v) = \ell\left(\underbrace{uv}_{=x}\right) = \ell(x) = L$, so that $u \underset{\sigma}{\sqcup\!\sqcup} z \in \mathfrak{A}^L$. Applying Lemma 6.3.10(c) to $\ell(u)$, $\ell(v)$ and z instead of n, m and v', we obtain $u \underset{\sigma}{\sqcup\!\sqcup} z \leq u \underset{\sigma}{\sqcup\!\sqcup} v$. But $u \underset{\sigma}{\sqcup\!\sqcup} v$ is an element of the multiset $u \sqcup\!\sqcup v$, and thus is $\leq x$ (since the lexicographically highest element of the multiset $u \sqcup\!\sqcup v$ is x). Thus, $u \underset{\sigma}{\sqcup\!\sqcup} v \leq x$, so that $u \underset{\sigma}{\sqcup\!\sqcup} z \leq u \underset{\sigma}{\sqcup\!\sqcup} v \leq x$.

Now, let us forget that we fixed z and σ. We thus have shown that

$$\text{any } z \in \mathfrak{A}^{\ell(v)} \text{ and } \sigma \in \operatorname{Sh}_{\ell(u),\ell(v)} \text{ satisfying } z \leq v \text{ satisfy } u \underset{\sigma}{\sqcup\!\sqcup} z \in \mathfrak{A}^L \text{ and } u \underset{\sigma}{\sqcup\!\sqcup} z \leq x. \tag{13.165.3}$$

A similar argument (but with some of the $\leq$ signs replaced by $<$ signs, and with a reference to Lemma 6.3.10(a) instead of a reference to Lemma 6.3.10(c)) shows that

$$\text{any } z \in \mathfrak{A}^{\ell(v)} \text{ and } \sigma \in \operatorname{Sh}_{\ell(u),\ell(v)} \text{ satisfying } z < v \text{ satisfy } u \underset{\sigma}{\sqcup\!\sqcup} z \in \mathfrak{A}^L \text{ and } u \underset{\sigma}{\sqcup\!\sqcup} z < x. \tag{13.165.4}$$

Now, (13.165.2) becomes

$$\begin{aligned}
\mathbf{b}_x &= \sum_{\substack{z \in \mathfrak{A}^{\ell(v)};\\ z \leq v}} \eta_{v,z} \underbrace{\sum_{\sigma \in \operatorname{Sh}_{\ell(u),\ell(v)}} b_{u \underset{\sigma}{\sqcup\!\sqcup} z}}_{\substack{= \sum\limits_{\substack{y \in \mathfrak{A}^L;\\ y \leq x}} \sum\limits_{\substack{\sigma \in \operatorname{Sh}_{\ell(u),\ell(v)};\\ u \underset{\sigma}{\sqcup\!\sqcup} z = y}} b_{u \underset{\sigma}{\sqcup\!\sqcup} z} \\ (\text{due to (13.165.3)})}} = \sum_{\substack{z \in \mathfrak{A}^{\ell(v)};\\ z \leq v}} \eta_{v,z} \sum_{\substack{y \in \mathfrak{A}^L;\\ y \leq x}} \sum_{\substack{\sigma \in \operatorname{Sh}_{\ell(u),\ell(v)};\\ u \underset{\sigma}{\sqcup\!\sqcup} z = y}} b_{u \underset{\sigma}{\sqcup\!\sqcup} z} \\
&= \sum_{\substack{y \in \mathfrak{A}^L;\\ y \leq x}} \sum_{\substack{z \in \mathfrak{A}^{\ell(v)};\\ z \leq v}} \eta_{v,z} \sum_{\substack{\sigma \in \operatorname{Sh}_{\ell(u),\ell(v)};\\ u \underset{\sigma}{\sqcup\!\sqcup} z = y}} \underbrace{b_{u \underset{\sigma}{\sqcup\!\sqcup} z}}_{\substack{=b_y \\ (\text{since } u \underset{\sigma}{\sqcup\!\sqcup} z = y)}} = \sum_{\substack{y \in \mathfrak{A}^L;\\ y \leq x}} \sum_{\substack{z \in \mathfrak{A}^{\ell(v)};\\ z \leq v}} \eta_{v,z} \underbrace{\sum_{\substack{\sigma \in \operatorname{Sh}_{\ell(u),\ell(v)};\\ u \underset{\sigma}{\sqcup\!\sqcup} z = y}} b_y}_{= \left| \left\{ \sigma \in \operatorname{Sh}_{\ell(u),\ell(v)} \ \mid \ u \underset{\sigma}{\sqcup\!\sqcup} z = y \right\} \right| \cdot b_y} \\
&= \sum_{\substack{y \in \mathfrak{A}^L;\\ y \leq x}} \sum_{\substack{z \in \mathfrak{A}^{\ell(v)};\\ z \leq v}} \eta_{v,z} \left| \left\{ \sigma \in \operatorname{Sh}_{\ell(u),\ell(v)} \ \mid \ u \underset{\sigma}{\sqcup\!\sqcup} z = y \right\} \right| \cdot b_y. \qquad (13.165.5)
\end{aligned}$$

Recall that we must prove that there is a family $(\eta_{x,y})_{y \in \mathfrak{A}^L} \in \mathbb{N}^{\mathfrak{A}^L}$ of elements of $\mathbb{N}$ satisfying (13.165.1) and $\eta_{x,x} \neq 0$ (in $\mathbb{N}$). In order to prove this, we define such a family $(\eta_{x,y})_{y \in \mathfrak{A}^L} \in \mathbb{N}^{\mathfrak{A}^L}$ by setting

$$\left(\eta_{x,y} = \sum_{\substack{z \in \mathfrak{A}^{\ell(v)};\\ z \leq v}} \eta_{v,z} \left| \left\{ \sigma \in \operatorname{Sh}_{\ell(u),\ell(v)} \ \mid \ u \underset{\sigma}{\sqcup\!\sqcup} z = y \right\} \right| \qquad \text{for every } y \in \mathfrak{A}^L \right).$$

Then, (13.165.5) shows that this family satisfies (13.165.1). All that remains to be proven is now to show that $\eta_{x,x} \neq 0$ (in $\mathbb{N}$).

Indeed, the definition of $\eta_{x,x}$ yields

$$\begin{aligned}
\eta_{x,x} &= \sum_{\substack{z\in\mathfrak{A}^{\ell(v)};\\ z\leq v}} \eta_{v,z} \left|\left\{\sigma\in \operatorname{Sh}_{\ell(u),\ell(v)} \ \mid\ u \underset{\sigma}{\sqcup\!\sqcup} z = x\right\}\right| \\
&= \sum_{\substack{z\in\mathfrak{A}^{\ell(v)};\\ z< v}} \eta_{v,z} \left| \underbrace{\left\{\sigma\in \operatorname{Sh}_{\ell(u),\ell(v)} \ \mid\ u \underset{\sigma}{\sqcup\!\sqcup} z = x\right\}}_{\substack{=\varnothing \\ \text{(since every } \sigma\in\operatorname{Sh}_{\ell(u),\ell(v)} \text{ satisfies } u\underset{\sigma}{\sqcup\!\sqcup} z<x \\ \text{(by (13.165.4)), so that it cannot satisfy } u\underset{\sigma}{\sqcup\!\sqcup} z=x)}} \right| + \eta_{v,v} \left|\left\{\sigma\in \operatorname{Sh}_{\ell(u),\ell(v)} \ \mid\ u \underset{\sigma}{\sqcup\!\sqcup} v = x\right\}\right| \\
&\qquad \text{(here, we have split off the addend for } z=v \text{ from the sum)} \\
&= \sum_{\substack{z\in\mathfrak{A}^{\ell(v)};\\ z< v}} \eta_{v,z} \underbrace{|\varnothing|}_{=0} + \eta_{v,v} \left|\left\{\sigma\in \operatorname{Sh}_{\ell(u),\ell(v)} \ \mid\ u \underset{\sigma}{\sqcup\!\sqcup} v = x\right\}\right| \\
&= \underbrace{\sum_{\substack{z\in\mathfrak{A}^{\ell(v)};\\ z< v}} \eta_{v,z} 0}_{=0} + \eta_{v,v} \left|\left\{\sigma\in \operatorname{Sh}_{\ell(u),\ell(v)} \ \mid\ u \underset{\sigma}{\sqcup\!\sqcup} v = x\right\}\right| \\
&= \underbrace{\eta_{v,v}}_{\neq 0 \text{ (in } \mathbb{N})} \underbrace{\left|\left\{\sigma\in \operatorname{Sh}_{\ell(u),\ell(v)} \ \mid\ u \underset{\sigma}{\sqcup\!\sqcup} v = x\right\}\right|}_{\substack{\neq 0 \text{ (in } \mathbb{N}) \\ \text{(since there exists some } \sigma\in\operatorname{Sh}_{\ell(u),\ell(v)} \text{ satisfying } u\underset{\sigma}{\sqcup\!\sqcup} v=x \\ \text{(because } x=uv\in u\sqcup\!\sqcup v))}} \\
&\neq 0 \text{ (in } \mathbb{N}).
\end{aligned}$$

This completes the proof that there is a family $\left(\eta_{x,y}\right)_{y\in\mathfrak{A}^L} \in \mathbb{N}^{\mathfrak{A}^L}$ of elements of $\mathbb{N}$ satisfying (13.165.1) and $\eta_{x,x} \neq 0$ (in $\mathbb{N}$). The induction step is thus complete, and Proposition 6.3.9 is proven by induction. □

13.166. **Solution to Exercise 6.3.13.** *Solution to Exercise 6.3.13.*

Proof of Theorem 6.3.4. It is clearly enough to prove that $\left(b_w\right)_{w\in\mathfrak{L}}$ is an algebraically independent generating set of the $\mathbf{k}$-algebra $\operatorname{Sh}(V)$.

For every word $u\in\mathfrak{A}^*$, define an element $\mathbf{b}_u$ by $\mathbf{b}_u = b_{a_1} \sqcup\!\sqcup\, b_{a_2} \sqcup\!\sqcup \cdots \sqcup\!\sqcup\, b_{a_p}$, where $(a_1, a_2, \ldots, a_p)$ is the CFL factorization of u. According to Lemma 6.3.7(a) (applied to $A = \operatorname{Sh}(V)$) [1111], the family $\left(b_w\right)_{w\in\mathfrak{L}}$ is an algebraically independent generating set of the $\mathbf{k}$-algebra $\operatorname{Sh}(V)$ if and only if the family $\left(\mathbf{b}_u\right)_{u\in\mathfrak{A}^*}$ is a basis of the $\mathbf{k}$-module $\operatorname{Sh}(V)$. In order to prove the former statement (which is our goal), it is therefore enough to prove the latter statement.

So we must prove the family $\left(\mathbf{b}_u\right)_{u\in\mathfrak{A}^*}$ is a basis of the $\mathbf{k}$-module $\operatorname{Sh}(V)$. We shall first prove a particular case of this statement:

Assertion A: If the set $\mathfrak{A}$ is finite, then the family $\left(\mathbf{b}_u\right)_{u\in\mathfrak{A}^*}$ is a basis of the $\mathbf{k}$-module $\operatorname{Sh}(V)$.

Proof of Assertion A: Assume that the set $\mathfrak{A}$ is finite. Fix $\ell\in\mathbb{N}$. Regard $V^{\otimes \ell}$ as a $\mathbf{k}$-submodule of $T(V)$. Then, $\left(b_{u_1}\otimes b_{u_2}\otimes\cdots\otimes b_{u_\ell}\right)_{u\in\mathfrak{A}^\ell}$ is a basis of the $\mathbf{k}$-module $V^{\otimes\ell}$ (since $\left(b_a\right)_{a\in\mathfrak{A}}$ is a basis of the $\mathbf{k}$-module V). In other words,

$$\left(b_u\right)_{u\in\mathfrak{A}^\ell} \text{ is a basis of the } \mathbf{k}\text{-module } V^{\otimes\ell} \tag{13.166.1}$$

(since $b_u = b_{u_1}b_{u_2}\cdots b_{u_\ell} = b_{u_1}\otimes b_{u_2}\otimes\cdots\otimes b_{u_\ell}$ for every $u\in\mathfrak{A}^\ell$).

[1111] Don't be confused by the fact that the multiplication of the $\mathbf{k}$-algebra $\operatorname{Sh}(V)$ is denoted by $\sqcup\!\sqcup$, but the multiplication of the $\mathbf{k}$-algebra A is denoted by $\cdot$ in Lemma 6.3.7(a).

We know that $\mathbb{Q}$ is a subring of $\mathbf{k}$. Hence, every nonzero element of $\mathbb{N}$ is an invertible element of $\mathbf{k}$.

Now, consider the set $\mathfrak{A}^\ell$ as a poset, whose smaller relation is $\leq$. (This poset is actually totally ordered, though we will not need this.) The notion of an invertibly-triangular $\mathfrak{A}^\ell \times \mathfrak{A}^\ell$-matrix is thus defined (according to Definition 11.1.7(c)).

The set $\mathfrak{A}^\ell$ is finite (since $\mathfrak{A}$ is finite); thus, it is a finite poset.

According to Proposition 6.3.9, the family $(\mathbf{b}_u)_{u\in\mathfrak{A}^\ell}$ expands invertibly triangularly in the basis $(b_u)_{u\in\mathfrak{A}^\ell}$ [1112]. Hence, Corollary 11.1.19(e) (applied to $V^{\otimes \ell}$, $\mathfrak{A}^\ell$, $(\mathbf{b}_u)_{u\in\mathfrak{A}^\ell}$ and $(b_u)_{u\in\mathfrak{A}^\ell}$ instead of M, S, $(e_s)_{s\in S}$ and $(f_s)_{s\in S}$) yields that the family $(\mathbf{b}_u)_{u\in\mathfrak{A}^\ell}$ is a basis of the $\mathbf{k}$-module $V^{\otimes \ell}$ if and only if the family $(b_u)_{u\in\mathfrak{A}^\ell}$ is a basis of the $\mathbf{k}$-module $V^{\otimes \ell}$ [1113]. Hence, the family $(\mathbf{b}_u)_{u\in\mathfrak{A}^\ell}$ is a basis of the $\mathbf{k}$-module $V^{\otimes \ell}$ (since the family $(b_u)_{u\in\mathfrak{A}^\ell}$ is a basis of the $\mathbf{k}$-module $V^{\otimes \ell}$).

Now, let us forget that we fixed ℓ. We thus have shown that, for every $\ell \in \mathbb{N}$, the family $(\mathbf{b}_u)_{u\in\mathfrak{A}^\ell}$ is a basis of the $\mathbf{k}$-module $V^{\otimes \ell}$. Hence, the disjoint union of the families $(\mathbf{b}_u)_{u\in\mathfrak{A}^\ell}$ over all $\ell \in \mathbb{N}$ is a basis of the direct sum $\bigoplus_{\ell\in\mathbb{N}} V^{\otimes \ell}$. Since the former disjoint union is the family $(\mathbf{b}_u)_{u\in\mathfrak{A}^*}$, while the latter direct sum is the $\mathbf{k}$-module $\bigoplus_{\ell\in\mathbb{N}} V^{\otimes \ell} = T(V) = \operatorname{Sh}(V)$, this rewrites as follows: The family $(\mathbf{b}_u)_{u\in\mathfrak{A}^*}$ is a basis of the $\mathbf{k}$-module $\operatorname{Sh}(V)$. This proves Assertion A.

Now, we need to prove that $(\mathbf{b}_u)_{u\in\mathfrak{A}^*}$ is a basis of the $\mathbf{k}$-module $\operatorname{Sh}(V)$, without the assumption that $\mathfrak{A}$ be finite. We will reduce this to Assertion A. First, we need some preparations:

Let $\mathfrak{B}$ be any subset of $\mathfrak{A}$. Then, $\mathfrak{B}$ canonically becomes a totally ordered set (since $\mathfrak{A}$ is totally ordered), so that the notion of a Lyndon word over $\mathfrak{B}$ is well-defined. We view $\mathfrak{B}^*$ as a subset of $\mathfrak{A}^*$, and so the Lyndon words over $\mathfrak{B}$ are precisely the Lyndon words over $\mathfrak{A}$ which lie in $\mathfrak{B}^*$.

We define a $\mathbf{k}$-submodule $V_{\mathfrak{B}}$ of V as the $\mathbf{k}$-linear span of the family $(b_a)_{a\in\mathfrak{B}}$. Notice that the family $(b_a)_{a\in\mathfrak{B}}$ is $\mathbf{k}$-linearly independent (being a subfamily of the basis $(b_a)_{a\in\mathfrak{A}}$ of V), and thus is a basis of the $\mathbf{k}$-submodule $V_{\mathfrak{B}}$.

The inclusion $V_{\mathfrak{B}} \to V$ gives rise to an injective $\mathbf{k}$-algebra homomorphism $\operatorname{Sh}(V_{\mathfrak{B}}) \to \operatorname{Sh}(V)$, which sends every b_w and every $\mathbf{b}_u$ (for $w \in \mathfrak{B}^*$ and $u \in \mathfrak{B}^*$, respectively) to the corresponding elements b_w and $\mathbf{b}_u$ of $\operatorname{Sh}(V)$, respectively. We regard this homomorphism as an inclusion, so that $\operatorname{Sh}(V_{\mathfrak{B}})$ is a $\mathbf{k}$-subalgebra of $\operatorname{Sh}(V)$.

[1112] *Proof.* Proposition 6.3.9 shows that, for every $x \in \mathfrak{A}^\ell$, there exists a family $(\eta_{x,y})_{y\in\mathfrak{A}^\ell} \in \mathbb{N}^{\mathfrak{A}^\ell}$ of elements of $\mathbb{N}$ such that

$$\mathbf{b}_x = \sum_{\substack{y\in\mathfrak{A}^\ell;\\ y\leq x}} \eta_{x,y} b_y \tag{13.166.2}$$

and

$$\eta_{x,x} \neq 0 \qquad \text{(in } \mathbb{N}\text{)}. \tag{13.166.3}$$

Consider such a family $(\eta_{x,y})_{y\in\mathfrak{A}^\ell} \in \mathbb{N}^{\mathfrak{A}^\ell}$ for each $x \in \mathfrak{A}^\ell$. Thus, an integer $\eta_{x,y} \in \mathbb{N} \subset \mathbb{Q} \subset \mathbf{k}$ is defined for each $(x,y) \in \mathfrak{A}^\ell \times \mathfrak{A}^\ell$.

We observe that the only elements $\eta_{s,t}$ (with $(s,t) \in \mathfrak{A}^\ell \times \mathfrak{A}^\ell$) appearing in the statements (13.166.2) and (13.166.3) are those which satisfy $t \leq s$. Hence, if some $(s,t) \in \mathfrak{A}^\ell \times \mathfrak{A}^\ell$ does not satisfy $t \leq s$, then the corresponding element $\eta_{s,t}$ does not appear in any of the statements (13.166.2) and (13.166.3); as a consequence, we can arbitrarily change the value of this $\eta_{s,t}$ without running the risk of invalidating (13.166.2) and (13.166.3). Hence, we can WLOG assume that

$$\text{every } (s,t) \in \mathfrak{A}^\ell \times \mathfrak{A}^\ell \text{ which does not satisfy } t \leq s \text{ must satisfy } \eta_{s,t} = 0 \tag{13.166.4}$$

(otherwise, we can just set all such $\eta_{s,t}$ to 0). Assume this. Thus, the matrix $(\eta_{x,y})_{(x,y)\in\mathfrak{A}^\ell\times\mathfrak{A}^\ell}$ is triangular. The diagonal entries $\eta_{x,x}$ of this matrix are nonzero elements of $\mathbb{N}$ (because of (13.166.3)) and therefore invertible elements of $\mathbf{k}$ (since every nonzero element of $\mathbb{N}$ is an invertible element of $\mathbf{k}$). Thus, the matrix $(\eta_{x,y})_{(x,y)\in\mathfrak{A}^\ell\times\mathfrak{A}^\ell}$ (regarded as a matrix in $\mathbf{k}^{\mathfrak{A}^\ell\times\mathfrak{A}^\ell}$) is invertibly triangular.

Now, every $x \in \mathfrak{A}^\ell$ satisfies

$$\sum_{y\in\mathfrak{A}^\ell} \eta_{x,y} b_y = \underbrace{\sum_{\substack{y\in\mathfrak{A}^\ell;\\ y\leq x}} \eta_{x,y} b_y}_{\substack{=\mathbf{b}_x\\ \text{(by (13.166.2))}}} + \sum_{\substack{y\in\mathfrak{A}^\ell;\\ \text{not } y\leq x}} \underbrace{\eta_{x,y}}_{\substack{=0\\ \text{(by (13.166.4), applied to } (s,t)=(x,y))}} b_y = \mathbf{b}_x + \underbrace{\sum_{\substack{y\in\mathfrak{A}^\ell;\\ \text{not } y\leq x}} 0 b_y}_{=0} = \mathbf{b}_x .$$

In other words, every $x \in \mathfrak{A}^\ell$ satisfies $\mathbf{b}_x = \sum_{y\in\mathfrak{A}^\ell} \eta_{x,y} b_y$. In other words, the family $(\mathbf{b}_u)_{u\in\mathfrak{A}^\ell}$ expands in the family $(b_u)_{u\in\mathfrak{A}^\ell}$ through the matrix $(\eta_{x,y})_{(x,y)\in\mathfrak{A}^\ell\times\mathfrak{A}^\ell}$. Since the latter matrix is invertibly triangular, we thus conclude that the family $(\mathbf{b}_u)_{u\in\mathfrak{A}^\ell}$ expands invertibly triangularly in the family $(b_u)_{u\in\mathfrak{A}^\ell}$.

[1113] Here, we have used the fact that the set $\mathfrak{A}^\ell$ is finite.

If $\mathfrak{B}$ is finite, then Assertion A (applied to $\mathfrak{B}$ instead of $\mathfrak{A}$) yields that the family $(\mathbf{b}_u)_{u\in\mathfrak{B}^*}$ is a basis of the $\mathbf{k}$-module $\operatorname{Sh}(V_{\mathfrak{B}})$.

Now, let us forget that we fixed $\mathfrak{B}$. We thus have shown that for every finite subset $\mathfrak{B}$ of $\mathfrak{A}$,

$$\text{the family } (\mathbf{b}_u)_{u\in\mathfrak{B}^*} \text{ is a basis of the } \mathbf{k}\text{-module } \operatorname{Sh}(V_{\mathfrak{B}}). \tag{13.166.5}$$

Let us introduce one more notation: A family of elements of a $\mathbf{k}$-module is said to be *finitely supported* if all but finitely many elements of this family are 0.

Now, let us show that the family $(\mathbf{b}_u)_{u\in\mathfrak{A}^*}$ is $\mathbf{k}$-linearly independent and spans the $\mathbf{k}$-module $\operatorname{Sh}(V)$.

Proof that the family $(\mathbf{b}_u)_{u\in\mathfrak{A}^*}$ is $\mathbf{k}$-linearly independent: Let $(\lambda_u)_{u\in\mathfrak{A}^*} \in \mathbf{k}^{\mathfrak{A}^*}$ be a finitely supported family of elements of $\mathbf{k}$ satisfying $\sum_{u\in\mathfrak{A}^*}\lambda_u\mathbf{b}_u = 0$. We are going to prove that all $u \in \mathfrak{A}^*$ satisfy $\lambda_u = 0$.

Indeed, the family $(\lambda_u)_{u\in\mathfrak{A}^*}$ is finitely supported, so that there exists a finite subset Z of $\mathfrak{A}^*$ such that all $u \in \mathfrak{A}^* \setminus Z$ satisfy $\lambda_u = 0$. Consider this Z. Since Z is finite, there exists a finite subset $\mathfrak{B}$ of $\mathfrak{A}$ satisfying $Z \subset \mathfrak{B}^*$ (in fact, we can take $\mathfrak{B}$ to be the set of all letters occurring in the words lying in Z). Consider this $\mathfrak{B}$. The family $(\mathbf{b}_u)_{u\in\mathfrak{B}^*}$ is a basis of the $\mathbf{k}$-module $\operatorname{Sh}(V_{\mathfrak{B}})$ (by (13.166.5)), and thus $\mathbf{k}$-linearly independent. We have $\mathfrak{A}^* \setminus \mathfrak{B}^* \subset \mathfrak{A}^* \setminus Z$ (since $Z \subset \mathfrak{B}^*$), and therefore all $u \in \mathfrak{A}^* \setminus \mathfrak{B}^*$ satisfy $\lambda_u = 0$ (since all $u \in \mathfrak{A}^* \setminus Z$ satisfy $\lambda_u = 0$). Hence, $\sum_{u\in\mathfrak{A}^*\setminus\mathfrak{B}^*}\underbrace{\lambda_u}_{=0}\mathbf{b}_u = \sum_{u\in\mathfrak{A}^*\setminus\mathfrak{B}^*}0\mathbf{b}_u = 0$. Now,

$$\begin{aligned} 0 &= \sum_{u\in\mathfrak{A}^*}\lambda_u\mathbf{b}_u = \sum_{u\in\mathfrak{B}^*}\lambda_u\mathbf{b}_u + \underbrace{\sum_{u\in\mathfrak{A}^*\setminus\mathfrak{B}^*}\lambda_u\mathbf{b}_u}_{=0} \qquad (\text{since } \mathfrak{B}^* \text{ is a subset of } \mathfrak{A}^*) \\ &= \sum_{u\in\mathfrak{B}^*}\lambda_u\mathbf{b}_u. \end{aligned}$$

So we have $\sum_{u\in\mathfrak{B}^*}\lambda_u\mathbf{b}_u = 0$. Thus, all $u \in \mathfrak{B}^*$ satisfy $\lambda_u = 0$ (since the family $(\mathbf{b}_u)_{u\in\mathfrak{B}^*}$ is $\mathbf{k}$-linearly independent). Combining this with the fact that all $u \in \mathfrak{A}^* \setminus \mathfrak{B}^*$ satisfy $\lambda_u = 0$, we conclude that all $u \in \mathfrak{A}^*$ satisfy $\lambda_u = 0$.

Now forget that we fixed $(\lambda_u)_{u\in\mathfrak{A}^*}$. We thus have shown that if $(\lambda_u)_{u\in\mathfrak{A}^*} \in \mathbf{k}^{\mathfrak{A}^*}$ is a finitely supported family of elements of $\mathbf{k}$ satisfying $\sum_{u\in\mathfrak{A}^*}\lambda_u\mathbf{b}_u = 0$, then all $u \in \mathfrak{A}^*$ satisfy $\lambda_u = 0$. In other words, the family $(\mathbf{b}_u)_{u\in\mathfrak{A}^*}$ is $\mathbf{k}$-linearly independent.

Proof that the family $(\mathbf{b}_u)_{u\in\mathfrak{A}^*}$ spans the $\mathbf{k}$-module $\operatorname{Sh}(V)$: We are going to show that the family $(\mathbf{b}_u)_{u\in\mathfrak{A}^*}$ spans the $\mathbf{k}$-module $\operatorname{Sh}(V)$. In order to prove this, it is enough to show that b_w lies in the $\mathbf{k}$-linear span of the family $(\mathbf{b}_u)_{u\in\mathfrak{A}^*}$ for every $w \in \mathfrak{A}^*$ (because the family $(b_w)_{w\in\mathfrak{A}^*}$ is a basis of the $\mathbf{k}$-module $\operatorname{Sh}(V)$ [1114]). So let us show this now.

Let $w \in \mathfrak{A}^*$. Then, there exists a finite subset $\mathfrak{B}$ of $\mathfrak{A}$ such that $w \in \mathfrak{B}^*$ (namely, we can take $\mathfrak{B}$ to be the set of all letters of w). Consider this $\mathfrak{B}$. The family $(\mathbf{b}_u)_{u\in\mathfrak{B}^*}$ is a basis of the $\mathbf{k}$-module $\operatorname{Sh}(V_{\mathfrak{B}})$ (by (13.166.5)), and thus spans this $\mathbf{k}$-module. But $b_w \in \operatorname{Sh}(V_{\mathfrak{B}})$ (since $w \in \mathfrak{B}^*$), and thus b_w lies in the $\mathbf{k}$-linear span of the family $(\mathbf{b}_u)_{u\in\mathfrak{B}^*}$ (since this family is a basis of the $\mathbf{k}$-module $\operatorname{Sh}(V_{\mathfrak{B}})$). Hence, b_w lies in the $\mathbf{k}$-linear span of the family $(\mathbf{b}_u)_{u\in\mathfrak{A}^*}$ as well (since this family $(\mathbf{b}_u)_{u\in\mathfrak{A}^*}$ includes the family $(\mathbf{b}_u)_{u\in\mathfrak{B}^*}$ as a subfamily). Thus, we have proven that b_w lies in the $\mathbf{k}$-linear span of the family $(\mathbf{b}_u)_{u\in\mathfrak{A}^*}$ for every $w \in \mathfrak{A}^*$. This completes the proof that the family $(\mathbf{b}_u)_{u\in\mathfrak{A}^*}$ spans the $\mathbf{k}$-module $\operatorname{Sh}(V)$.

Altogether, we now know that the family $(\mathbf{b}_u)_{u\in\mathfrak{A}^*}$ is $\mathbf{k}$-linearly independent and spans the $\mathbf{k}$-module $\operatorname{Sh}(V)$. In other words, $(\mathbf{b}_u)_{u\in\mathfrak{A}^*}$ is a basis of the $\mathbf{k}$-module $\operatorname{Sh}(V)$. This completes our proof of Theorem 6.3.4. □

13.167. **Solution to Exercise 6.4.2.** *Solution to Exercise 6.4.2.* We shall give two solutions to this exercise; but they both rest on the following lemmas:

[1114] *Proof.* For every $\ell \in \mathbb{N}$, the family $(b_u)_{u\in\mathfrak{A}^\ell}$ is a basis of the $\mathbf{k}$-module $V^{\otimes\ell}$ (by (13.166.1)). Hence, the disjoint union of the families $(b_u)_{u\in\mathfrak{A}^\ell}$ over all $\ell \in \mathbb{N}$ is a basis of the direct sum $\bigoplus_{\ell\in\mathbb{N}} V^{\otimes\ell}$. Since the former disjoint union is the family $(b_u)_{u\in\mathfrak{A}^*} = (b_w)_{w\in\mathfrak{A}^*}$, whereas the latter direct sum is the $\mathbf{k}$-module $\bigoplus_{\ell\in\mathbb{N}} V^{\otimes\ell} = T(V) = \operatorname{Sh}(V)$, this rewrites as follows: The family $(b_w)_{w\in\mathfrak{A}^*}$ is a basis of the $\mathbf{k}$-module $\operatorname{Sh}(V)$, qed.

Lemma 13.167.1. *For every positive integer N, we have*

$$\sum_{d\mid N}\mu\left(d\right)=\delta_{N,1}. \tag{13.167.1}$$

Lemma 13.167.1 is one of the most fundamental properties of the number-theoretic Möbius function; it will not be proven here.[1115]

Lemma 13.167.2. *Every positive integer n satisfies*

$$\frac{1}{n}\sum_{d\mid n}\mu\left(d\right)\left(2^{n/d}-1\right)=\frac{1}{n}\sum_{d\mid n}\mu\left(d\right)2^{n/d}-\delta_{n,1}.$$

Proof of Lemma 13.167.2. Let n be a positive integer. Then,

$$\begin{aligned}
&\underbrace{\frac{1}{n}\sum_{d\mid n}\mu\left(d\right)\left(2^{n/d}-1\right)}_{=\sum_{d\mid n}\mu(d)2^{n/d}-\sum_{d\mid n}\mu(d)1}\\
&=\frac{1}{n}\left(\sum_{d\mid n}\mu\left(d\right)2^{n/d}-\sum_{d\mid n}\mu\left(d\right)1\right)=\frac{1}{n}\left(\sum_{d\mid n}\mu\left(d\right)2^{n/d}-\sum_{d\mid n}\mu\left(d\right)\right)\\
&=\frac{1}{n}\sum_{d\mid n}\mu\left(d\right)2^{n/d}-\frac{1}{n}\underbrace{\sum_{d\mid n}\mu\left(d\right)}_{\substack{=\delta_{n,1}\\ \text{(by (13.167.1),}\\ \text{applied to } N=n)}}=\frac{1}{n}\sum_{d\mid n}\mu\left(d\right)2^{n/d}-\underbrace{\frac{1}{n}\delta_{n,1}}_{=\delta_{n,1}}=\frac{1}{n}\sum_{d\mid n}\mu\left(d\right)2^{n/d}-\delta_{n,1}.
\end{aligned}$$

This proves Lemma 13.167.2. □

First solution to Exercise 6.4.2. [The following solution is similar to the solution of Exercise 6.1.29.]

For every positive integer n, let $\operatorname{lync} n$ denote the number of Lyndon compositions of size n. We need to prove that

$$\operatorname{lync} n=\frac{1}{n}\sum_{d\mid n}\mu\left(d\right)\left(2^{n/d}-1\right)=\frac{1}{n}\sum_{d\mid n}\mu\left(d\right)2^{n/d}-\delta_{n,1} \tag{13.167.2}$$

for every positive integer n.

We recall that $\mathfrak{A}^*=\operatorname{Comp}$; that is, the elements of $\mathfrak{A}^*$ are the compositions. Hence, the notation $|w|$ makes sense for any $w\in\mathfrak{A}^*$; it denotes the size of the composition w.

For every $n\in\mathbb{N}$, we have

$$\left|\operatorname{Comp}_n\right|=\begin{cases}2^{n-1}, & \text{if } n\geq 1;\\ 1, & \text{if } n=0\end{cases}$$

[1115] For a proof of Lemma 13.167.1, see the solution of Exercise 2.9.6. (More precisely, Lemma 13.167.1 is obtained from (13.84.3) by renaming n as N.)

[1116]. Thus, for every $n \in \mathbb{N}$, we have

$$\left|\left\{w \in \underbrace{\mathfrak{A}^*}_{=\operatorname{Comp}} \ \mid \ |w| = n\right\}\right| = \left|\underbrace{\{w \in \operatorname{Comp} \ \mid \ |w| = n\}}_{=\operatorname{Comp}_n}\right| = |\operatorname{Comp}_n| \\ = \begin{cases} 2^{n-1}, & \text{if } n \geq 1; \\ 1, & \text{if } n = 0 \end{cases}. \tag{13.167.3}$$

Let $\mathfrak{M}$ denote the set of all finite multisets of Lyndon compositions. Define two maps $\mathbf{m} : \mathfrak{M} \to \mathfrak{A}^*$ and $\mathbf{n} : \mathfrak{A}^* \to \mathfrak{M}$ as in Proposition 13.143.1. Then, Proposition 13.143.1 shows that these two maps $\mathbf{m}$ and $\mathbf{n}$ are mutually inverse bijections (since a Lyndon composition is the same thing as a Lyndon word over the alphabet $\mathfrak{A}$). Hence, the map $\mathbf{m}$ is a bijection.

On the other hand, let $\mathfrak{L}$ be the set of all Lyndon compositions. Thus, the definition of $\mathfrak{M}$ says that $\mathfrak{M}$ is the set of all finite multisets of elements of $\mathfrak{L}$. Also, the definition of $\operatorname{lync} n$ now rewrites as

$$\operatorname{lync} n = |\{w \in \mathfrak{L} \ \mid \ |w| = n\}| \qquad \text{for every positive integer } n. \tag{13.167.4}$$

Let $\mathfrak{N}$ be the set of all families $(k_w)_{w \in \mathfrak{L}} \in \mathbb{N}^{\mathfrak{L}}$ of nonnegative integers (indexed by the Lyndon compositions) such that all but finitely many $w \in \mathfrak{L}$ satisfy $k_w = 0$. Proposition 13.143.2 (applied to $S = \mathfrak{L}$) shows that the map $\operatorname{mult} : \mathfrak{M} \to \mathfrak{N}$ that sends each multiset $M \in \mathfrak{M}$ to the family

$$((\text{multiplicity of } w \text{ in the multiset } M))_{w \in S} \in \mathfrak{N}$$

is well-defined and is a bijection. Consider this map mult.

The composition $\mathbf{m} \circ \operatorname{mult}^{-1} : \mathfrak{N} \to \mathfrak{A}^*$ of the bijections $\mathbf{m}$ and mult^{-1} is clearly a bijection. It can easily be seen to satisfy

$$\left|\left(\mathbf{m} \circ \operatorname{mult}^{-1}\right)\left((k_w)_{w \in \mathfrak{L}}\right)\right| = \sum_{w \in \mathfrak{L}} k_w \cdot |w| \qquad \text{for every } (k_w)_{w \in \mathfrak{L}} \in \mathfrak{N}. \tag{13.167.5}$$

[1116] *Proof.* Let $n \in \mathbb{N}$. Recall that we have a bijection $\operatorname{Comp}_n \to 2^{[n-1]}$, where $[n-1] = \{1, 2, \ldots, n-1\}$. Thus,

$$|\operatorname{Comp}_n| = \left|2^{[n-1]}\right| = 2^{|[n-1]|} = \begin{cases} 2^{n-1}, & \text{if } n \geq 1; \\ 2^0, & \text{if } n = 0 \end{cases} \qquad \left(\text{since } |[n-1]| = \begin{cases} n-1, & \text{if } n \geq 1; \\ 0, & \text{if } n = 0 \end{cases}\right) \\ = \begin{cases} 2^{n-1}, & \text{if } n \geq 1; \\ 1, & \text{if } n = 0 \end{cases},$$

qed.

[1117] Now, in the ring $\mathbb{Q}\left[\left[t\right]\right]$ of formal power series, we have

$$\begin{aligned}\sum_{w\in\mathfrak{A}^*} t^{|w|} &= \sum_{n\in\mathbb{N}} \underbrace{\left|\left\{w\in\mathfrak{A}^* \mid |w|=n\right\}\right|}_{\begin{array}{c}=\begin{cases}2^{n-1}, & \text{if } n\geq 1;\\ 1, & \text{if } n=0\end{cases}\\ \text{(by (13.167.3))}\end{array}} t^n = \sum_{n\in\mathbb{N}} \begin{cases}2^{n-1}, & \text{if } n\geq 1;\\ 1, & \text{if } n=0\end{cases} \cdot t^n \\ &= \underbrace{1t^0}_{=1} + \underbrace{\sum_{\substack{n\in\mathbb{N};\\ n\geq 1}} 2^{n-1}t^n}_{\begin{array}{c}=\sum_{n\in\mathbb{N}} 2^n t^{n+1}\\ \text{(here, we have substituted}\\ n+1 \text{ for } n \text{ in the sum)}\end{array}} = 1 + \sum_{n\in\mathbb{N}} \underbrace{2^n t^{n+1}}_{=(2t)^n t} = 1 + \underbrace{\sum_{n\in\mathbb{N}} \left(2t\right)^n}_{=\dfrac{1}{1-2t}} t = 1 + \frac{1}{1-2t} t = \frac{1-t}{1-2t}.\end{aligned}$$

[1117] *Proof of (13.167.5):* Let $\left(k_w\right)_{w\in\mathfrak{L}} \in \mathfrak{N}$. Let $M = \operatorname{mult}^{-1}\left(\left(k_w\right)_{w\in\mathfrak{L}}\right)$. Then, M is a multiset of elements of $\mathfrak{L}$ and satisfies $\left(k_w\right)_{w\in\mathfrak{L}} = \operatorname{mult} M = \left(\left(\text{multiplicity of } w \text{ in the multiset } M\right)\right)_{w\in\mathfrak{L}}$. In other words, every $w\in\mathfrak{L}$ satisfies

$$k_w = \left(\text{multiplicity of } w \text{ in the multiset } M\right). \tag{13.167.6}$$

Let $a_1, a_2, \ldots, a_k$ denote the elements of this multiset M listed in decreasing order. Then, the definition of $\mathbf{m}$ yields $\mathbf{m}\left(M\right) = a_1a_2\cdots a_k$, so that

$$\begin{aligned}\left|\mathbf{m}\left(M\right)\right| &= \left|a_1a_2\cdots a_k\right| = \sum_{i\in\{1,2,\ldots,k\}} \left|a_i\right| = \sum_{w\in\mathfrak{L}} \sum_{\substack{i\in\{1,2,\ldots,k\};\\ a_i=w}} \left|\underbrace{a_i}_{\substack{=w\\ \text{(since } a_i=w)}}\right| \qquad \text{(since every } a_i \text{ belongs to } \mathfrak{L}\text{)} \\ &= \sum_{w\in\mathfrak{L}} \underbrace{\sum_{\substack{i\in\{1,2,\ldots,k\};\\ a_i=w}} |w|}_{=(\text{number of } i\in\{1,2,\ldots,k\} \text{ satisfying } a_i=w)\cdot|w|} = \sum_{w\in\mathfrak{L}} \underbrace{\left(\text{number of } i\in\{1,2,\ldots,k\} \text{ satisfying } a_i=w\right)}_{\substack{=(\text{multiplicity of } w \text{ in the multiset } M)=k_w\\ \text{(by (13.167.6))}}} \cdot |w| \\ &= \sum_{w\in\mathfrak{L}} k_w \cdot |w|.\end{aligned}$$

Now,

$$\left|\left(\mathbf{m}\circ\operatorname{mult}^{-1}\right)\left(\left(k_w\right)_{w\in\mathfrak{L}}\right)\right| = \left|\mathbf{m}\left(\underbrace{\operatorname{mult}^{-1}\left(\left(k_w\right)_{w\in\mathfrak{L}}\right)}_{=M}\right)\right| = \left|\mathbf{m}\left(M\right)\right| = \sum_{w\in\mathfrak{L}} k_w\cdot|w|,$$

which proves (13.167.5).

Hence,

$$\begin{aligned}
\frac{1-t}{1-2t} &= \sum_{w\in\mathfrak{A}^*} t^{|w|} = \sum_{(k_w)_{w\in\mathfrak{L}}\in\mathfrak{N}} t^{\left|\left(\mathbf{m}\circ\operatorname{mult}^{-1}\right)\left((k_w)_{w\in\mathfrak{L}}\right)\right|} \\
&\qquad \left(\begin{array}{c}\text{here, we substituted } \left(\mathbf{m}\circ\operatorname{mult}^{-1}\right)\left((k_w)_{w\in\mathfrak{L}}\right) \text{ for } w \text{ in the sum,}\\ \text{since the map } \mathbf{m}\circ\operatorname{mult}^{-1} : \mathfrak{N}\to\mathfrak{A}^* \text{ is a bijection}\end{array}\right) \\
&= \sum_{(k_w)_{w\in\mathfrak{L}}\in\mathfrak{N}} t^{\sum_{w\in\mathfrak{L}} k_w\cdot|w|} \qquad \text{(by (13.167.5))} \\
&= \sum_{(k_w)_{w\in\mathfrak{L}}\in\mathfrak{N}} \prod_{w\in\mathfrak{L}} t^{k_w\cdot|w|} = \prod_{w\in\mathfrak{L}} \underbrace{\sum_{k\in\mathbb{N}} t^{k\cdot|w|}}_{=\frac{1}{1-t^{|w|}}} \qquad \text{(by the product rule)} \\
&= \prod_{w\in\mathfrak{L}} \frac{1}{1-t^{|w|}} = \prod_{n\geq 1} \prod_{\substack{w\in\mathfrak{L};\\ |w|=n}} \underbrace{\frac{1}{1-t^{|w|}}}_{\substack{=\frac{1}{1-t^n}\\ \text{(since } |w|=n)}} \qquad \text{(since } |w|\geq 1 \text{ for every } w\in\mathfrak{L}) \\
&= \prod_{n\geq 1} \underbrace{\prod_{\substack{w\in\mathfrak{L};\\ |w|=n}} \frac{1}{1-t^n}}_{\substack{=\left(\frac{1}{1-t^n}\right)^{\operatorname{lync} n}\\ \text{(by (13.167.4))}}} = \prod_{n\geq 1} \left(\frac{1}{1-t^n}\right)^{\operatorname{lync} n}.
\end{aligned}$$

Taking the logarithm of both sides of this identity, we obtain

$$\begin{aligned}
\log\frac{1-t}{1-2t} &= \log\left(\prod_{n\geq 1}\left(\frac{1}{1-t^n}\right)^{\operatorname{lync} n}\right) = \sum_{n\geq 1} (\operatorname{lync} n)\cdot \underbrace{\log\left(\frac{1}{1-t^n}\right)}_{\substack{=-\log(1-t^n)=\sum_{u\geq 1}\frac{1}{u}(t^n)^u\\ \text{(by the Mercator series for the logarithm)}}} \\
&= \sum_{n\geq 1} (\operatorname{lync} n)\cdot \sum_{u\geq 1}\frac{1}{u}(t^n)^u = \sum_{n\geq 1}\sum_{u\geq 1} (\operatorname{lync} n)\frac{1}{u}\underbrace{(t^n)^u}_{=t^{nu}} = \sum_{n\geq 1}\sum_{u\geq 1} (\operatorname{lync} n)\frac{1}{u}t^{nu} \\
&= \underbrace{\sum_{n\geq 1}\sum_{\substack{v\geq 1;\\ n\mid v}}}_{=\sum_{v\geq 1}\sum_{n\mid v}} (\operatorname{lync} n)\underbrace{\frac{1}{v/n}}_{=\frac{n}{v}} t^v \qquad \text{(here, we substituted } v/n \text{ for } u \text{ in the second sum)} \\
&= \sum_{v\geq 1}\sum_{n\mid v} (\operatorname{lync} n)\frac{n}{v}t^v = \sum_{n\geq 1}\sum_{d\mid n} (\operatorname{lync} d)\frac{d}{n}t^n
\end{aligned}$$

(here, we renamed the summation indices v and n as n and d). Since

$$\begin{aligned}\log\frac{1-t}{1-2t} &= \log(1-t) - \log(1-2t) \\ &= \underbrace{(-\log(1-2t))}_{\substack{=\sum_{n\geq 1}\frac{1}{n}(2t)^n \\ \text{(by the Mercator series for the logarithm)}}} - \underbrace{(-\log(1-t))}_{\substack{=\sum_{n\geq 1}\frac{1}{n}t^n \\ \text{(by the Mercator series for the logarithm)}}} \\ &= \sum_{n\geq 1}\frac{1}{n}(2t)^n - \sum_{n\geq 1}\frac{1}{n}t^n = \sum_{n\geq 1}\frac{1}{n}\underbrace{\left((2t)^n - t^n\right)}_{=(2^n-1)t^n} = \sum_{n\geq 1}\frac{1}{n}(2^n-1)t^n,\end{aligned}$$

this rewrites as

$$\sum_{n\geq 1}\frac{1}{n}(2^n-1)t^n = \sum_{n\geq 1}\sum_{d\mid n}(\operatorname{lync} d)\frac{d}{n}t^n.$$

Comparing coefficients, we conclude that every positive integer n satisfies

$$\frac{1}{n}(2^n-1) = \sum_{d\mid n}(\operatorname{lync} d)\frac{d}{n}.$$

Multiplying this with n, we obtain

$$2^n - 1 = \sum_{d\mid n}(\operatorname{lync} d)\, d. \tag{13.167.7}$$

Now, every positive integer n satisfies

$$\begin{aligned}\sum_{d\mid n}\mu(d)\left(2^{n/d}-1\right) &= \sum_{e\mid n}\mu(e)\underbrace{\left(2^{n/e}-1\right)}_{\substack{=\sum_{d\mid n/e}(\operatorname{lync} d)d \\ \text{(by (13.167.7), applied} \\ \text{to } n/e \text{ instead of } n)}} = \sum_{e\mid n}\mu(e)\sum_{d\mid n/e}(\operatorname{lync} d)\, d \\ &= \underbrace{\sum_{e\mid n}\sum_{d\mid n/e}}_{=\sum_{d\mid n}\sum_{e\mid n/d}}\mu(e)(\operatorname{lync} d)\, d = \sum_{d\mid n}\underbrace{\sum_{e\mid n/d}\mu(e)}_{\substack{=\delta_{n/d,1} \\ \text{(by (13.167.1), applied} \\ \text{to } N=n/d)}}(\operatorname{lync} d)\, d \\ &= \sum_{d\mid n}\underbrace{\delta_{n/d,1}}_{=\delta_{n,d}}(\operatorname{lync} d)\, d = \sum_{d\mid n}\delta_{n,d}(\operatorname{lync} d)\, d = (\operatorname{lync} n)\, n.\end{aligned}$$

Dividing this by n, we obtain $\frac{1}{n}\sum_{d\mid n}\mu(d)\left(2^{n/d}-1\right) = \operatorname{lync} n$. Hence,

$$\operatorname{lync} n = \frac{1}{n}\sum_{d\mid n}\mu(d)\left(2^{n/d}-1\right) = \frac{1}{n}\sum_{d\mid n}\mu(d)\, 2^{n/d} - \delta_{n,1}$$

(by Lemma 13.167.2). This proves (13.167.2). Thus, Exercise 6.4.2 is solved.

Second solution to Exercise 6.4.2. Let $\mathfrak{B}$ denote the two-element set $\{\mathbf{0},\mathbf{1}\}$, where $\mathbf{0}$ and $\mathbf{1}$ are two new objects. We make $\mathfrak{B}$ into a totally ordered set by setting $\mathbf{0}<\mathbf{1}$. In the following, we will study not only words over the alphabet $\mathfrak{A}=\{1,2,3,\ldots\}$, but also words over the alphabet $\mathfrak{B}$. The latter words form the set $\mathfrak{B}^*$. Let $\mathfrak{L}_{\mathfrak{B}}$ denote the set of all Lyndon words over the alphabet $\mathfrak{B}$.

For every $k \in \mathfrak{A}$, define an element $\widetilde{k}$ of $\mathfrak{B}^*$ by $\widetilde{k} = \left(\mathbf{0}, \underbrace{\mathbf{1}, \mathbf{1}, \ldots, \mathbf{1}}_{k-1 \text{ times}}\right)$. (This is well-defined, since $\mathfrak{A} = \{1, 2, 3, \ldots\}$.) We can rewrite the definition of $\widetilde{k}$ as follows: We have[1118]

$$\widetilde{k} = \mathbf{01}^{k-1} \qquad \text{for every positive integer } k \tag{13.167.8}$$

(where "$\mathbf{0}$" and "$\mathbf{1}$" are regarded as one-letter words). Thus,

$$\ell\left(\widetilde{k}\right) = k \qquad \text{for every positive integer } k \tag{13.167.9}$$

[1119]. Also,

$$\widetilde{k+1} = \widetilde{k}\mathbf{1} \qquad \text{for every positive integer } k \tag{13.167.10}$$

(where "$\widetilde{k}\mathbf{1}$" means the concatenation of $\widetilde{k}$ with the one-letter word $\mathbf{1}$). [1120] As a consequence, $\widetilde{k}$ is a prefix of $\widetilde{k+1}$ for every positive integer k. Thus, $\widetilde{k} < \widetilde{k+1}$ (in the lexicographic order on $\mathfrak{B}^*$) for every positive integer k [1121]. In other words,

$$\widetilde{1} < \widetilde{2} < \widetilde{3} < \cdots \qquad \text{in the lexicographic order on } \mathfrak{B}^*.$$

We notice a slightly stronger property: For any $a \in \mathfrak{A}$ and $b \in \mathfrak{A}$ satisfying $a < b$, we have

$$\widetilde{a}\left(\widetilde{c_1}\widetilde{c_2}\cdots\widetilde{c_s}\right) < \widetilde{b}w \tag{13.167.11}$$

for any $s \in \mathbb{N}$, any $c_1, c_2, \ldots, c_s \in \mathfrak{A}$ and any $w \in \mathfrak{B}^*$. [1122]

[1118] Here and in the following, expressions like $\mathbf{01}^n$ (for $n \in \mathbb{N}$) have to be understood as $\mathbf{0}\left(\mathbf{1}^n\right)$ rather than as $\left(\mathbf{01}\right)^n$. (The objects $\mathbf{0}$ and $\mathbf{1}$ are not actually numbers; they don't form digital expansions.)

[1119] *Proof.* Let k be a positive integer. Then, (13.167.8) yields $\widetilde{k} = \mathbf{01}^{k-1}$. Now,

$$\ell\left(\underbrace{\widetilde{k}}_{=\mathbf{01}^{k-1}}\right) = \ell\left(\mathbf{01}^{k-1}\right) = \underbrace{\ell\left(\mathbf{0}\right)}_{=1} + \underbrace{\ell\left(\mathbf{1}^{k-1}\right)}_{=(k-1)\ell(\mathbf{1})} = 1 + (k-1)\underbrace{\ell\left(\mathbf{1}\right)}_{=1} = 1 + (k-1) = k,$$

qed.

[1120] *Proof.* Let k be a positive integer. Then, $k \geq 1$, so that $k - 1 \geq 0$. But (13.167.8) (applied to $k+1$ instead of k) yields

$$\widetilde{k+1} = \mathbf{0} \underbrace{\mathbf{1}^{(k+1)-1}}_{\substack{=\mathbf{1}^k = \mathbf{1}^{(k-1)+1} = \mathbf{1}^{k-1}\mathbf{1} \\ (\text{since } k-1 \geq 0)}} = \underbrace{\mathbf{01}^{k-1}}_{\substack{=\widetilde{k} \\ (\text{by } (13.167.8))}} \mathbf{1} = \widetilde{k}\mathbf{1},$$

qed.

[1121] *Proof.* Let k be a positive integer. Then, $\widetilde{k} \leq \widetilde{k+1}$ (since $\widetilde{k}$ is a prefix of $\widetilde{k+1}$). Also, (13.167.9) (applied to $k+1$ instead of k) yields $\ell\left(\widetilde{k+1}\right) = k+1$. Now, (13.167.9) yields $\ell\left(\widetilde{k}\right) = k < k+1 = \ell\left(\widetilde{k+1}\right)$, so that $\ell\left(\widetilde{k}\right) \neq \ell\left(\widetilde{k+1}\right)$ and thus $\widetilde{k} \neq \widetilde{k+1}$. Combined with $\widetilde{k} \leq \widetilde{k+1}$, this yields $\widetilde{k} < \widetilde{k+1}$, qed.

[1122] *Proof of (13.167.11):* Let $a \in \mathfrak{A}$ and $b \in \mathfrak{A}$ be such that $a < b$. Let $s \in \mathbb{N}$, let $c_1, c_2, \ldots, c_s \in \mathfrak{A}$ and let $w \in \mathfrak{B}^*$. We need to prove that (13.167.11) holds.

Assume the contrary. Thus, $\widetilde{a}\left(\widetilde{c_1}\widetilde{c_2}\cdots\widetilde{c_s}\right) < \widetilde{b}w$ does not hold. We thus have $\widetilde{a}\left(\widetilde{c_1}\widetilde{c_2}\cdots\widetilde{c_s}\right) \geq \widetilde{b}w$. Hence, $\widetilde{a}\left(\widetilde{c_1}\widetilde{c_2}\cdots\widetilde{c_s}\right) \geq \widetilde{b}w \geq \widetilde{b}$.

We notice that $b - a > 0$ (since $a < b$), so that the word $\mathbf{1}^{b-a}$ is nonempty. Also, $b - a > 0$, so that $b - a \geq 1$ (since $b - a$ is an integer) and thus $(b-a) - 1 \geq 0$. Hence, the word $\mathbf{1}^{(b-a)-1}$ is well-defined, and we have $\mathbf{1}^{b-a} = \mathbf{11}^{(b-a)-1}$.

Applying (13.167.8) to $k = a$, we obtain $\widetilde{a} = \mathbf{01}^{a-1}$. But $b - 1 = (a-1) + (b-a)$, so that $\mathbf{1}^{b-1} = \mathbf{1}^{(a-1)+(b-a)} = \mathbf{1}^{a-1}\mathbf{1}^{b-a}$ (since $b - a > 0$). Now, (13.167.8) (applied to $k = b$) yields

$$\widetilde{b} = \mathbf{0} \underbrace{\mathbf{1}^{b-1}}_{=\mathbf{1}^{a-1}\mathbf{1}^{b-a}} = \underbrace{\mathbf{01}^{a-1}}_{=\widetilde{a}} \mathbf{1}^{b-a} = \widetilde{a}\mathbf{1}^{b-a} > \widetilde{a}$$

(since $\mathbf{1}^{b-a}$ is a nonempty word). Hence, $\widetilde{a} < \widetilde{b}$. Now, if we had $s = 0$, then we would have $\widetilde{a} \underbrace{\left(\widetilde{c_1}\widetilde{c_2}\cdots\widetilde{c_s}\right)}_{=(\text{empty product})=\varnothing} = \widetilde{a} < \widetilde{b}$, which would contradict $\widetilde{a}\left(\widetilde{c_1}\widetilde{c_2}\cdots\widetilde{c_s}\right) \geq \widetilde{b}$. Hence, we cannot have $s = 0$. We thus have $s \geq 1$. Consequently, c_1 is well-defined. From (13.167.8) (applied to $k = c_1$), we have $\widetilde{c_1} = \mathbf{01}^{c_1 - 1}$. Thus, $\mathbf{0}$ is a prefix of $\widetilde{c_1}$. Since $\widetilde{c_1}$ is, in turn, a prefix of $\widetilde{c_1}\widetilde{c_2}\cdots\widetilde{c_s}$, this yields that $\mathbf{0}$ is a prefix of $\widetilde{c_1}\widetilde{c_2}\cdots\widetilde{c_s}$. In other words, there exists a $t \in \mathfrak{B}^*$ such that $\widetilde{c_1}\widetilde{c_2}\cdots\widetilde{c_s} = \mathbf{0}t$. Consider this t. Now,

$$\widetilde{a} \underbrace{\mathbf{0}t}_{=\widetilde{c_1}\widetilde{c_2}\cdots\widetilde{c_s}} = \widetilde{a}\left(\widetilde{c_1}\widetilde{c_2}\cdots\widetilde{c_s}\right) \geq \widetilde{b} = \widetilde{a}\mathbf{1}^{b-a}.$$

Also,

$$\widetilde{k} \text{ is a Lyndon word over the alphabet } \mathfrak{B} \text{ for every positive integer } k. \tag{13.167.12}$$

[1123]

We can now define a map $\Phi : \mathfrak{A}^* \to \mathfrak{B}^*$ by setting

$$\left(\Phi\left((w_1, w_2, \ldots, w_k) \right) = \widetilde{w_1}\widetilde{w_2}\cdots\widetilde{w_k} \qquad \text{for every word } (w_1, w_2, \ldots, w_k) \in \mathfrak{A}^* \right).$$

In other words, $\widetilde{a}\mathbf{1}^{b-a} \leq \widetilde{a}\mathbf{0}t$. Thus, Proposition 6.1.2(c) (applied to $\mathfrak{B}$, $\widetilde{a}$, $\mathbf{1}^{b-a}$ and $\mathbf{0}t$ instead of $\mathfrak{A}$, a, c and d) yields $\mathbf{1}^{b-a} \leq \mathbf{0}t$. Now, $\mathbf{1}\mathbf{1}^{(b-a)-1} = \mathbf{1}^{b-a} \leq \mathbf{0}t$. Therefore, Proposition 6.1.2(e) (applied to $\mathfrak{B}$, $\mathbf{1}$, $\mathbf{1}^{(b-a)-1}$, $\mathbf{0}$ and t instead of $\mathfrak{A}$, a, b, c and d) yields that either we have $\mathbf{1} \leq \mathbf{0}$ or the word $\mathbf{0}$ is a prefix of $\mathbf{1}$. Since the word $\mathbf{0}$ is not a prefix of $\mathbf{1}$, this shows that $\mathbf{1} \leq \mathbf{0}$. But this contradicts $\mathbf{0} < \mathbf{1}$. This contradiction proves that our assumption was wrong, qed.

[1123] *Proof.* Let k be a positive integer. From (13.167.8), we obtain $\widetilde{k} = \mathbf{0}\mathbf{1}^{k-1}$.

From (13.167.9), we have $\ell\left(\widetilde{k}\right) = k \geq 1$. Thus, the word $\widetilde{k}$ is nonempty.

Let v be a nonempty proper suffix of $\widetilde{k}$. We will show that $v > \widetilde{k}$.

Indeed, there exists a nonempty $u \in \mathfrak{B}^*$ satisfying $\widetilde{k} = uv$ (since v is a proper suffix of $\widetilde{k}$). Consider this u. Since u is nonempty, the first letter of u is well-defined. We have

$$\begin{aligned}(\text{the first letter of } u) &= \left(\text{the first letter of } \underbrace{\widetilde{k}}_{=\mathbf{0}\mathbf{1}^{k-1}}\right) \qquad \left(\text{since } u \text{ is a prefix of } \widetilde{k} \text{ (since } \widetilde{k} = uv)\right) \\ &= \left(\text{the first letter of } \mathbf{0}\mathbf{1}^{k-1}\right) = \mathbf{0}.\end{aligned}$$

Thus, $\mathbf{0}$ is a prefix of u. In other words, there exists a word $u' \in \mathfrak{B}^*$ satisfying $u = \mathbf{0}u'$. Consider this u'. We have $\underbrace{\mathbf{0}u'}_{=u} v = uv = \widetilde{k} = \mathbf{0}\mathbf{1}^{k-1}$. Cancelling $\mathbf{0}$ from this equality, we obtain $u'v = \mathbf{1}^{k-1}$. Hence, v is a suffix of the word $\mathbf{1}^{k-1}$. Thus, v has the form $\mathbf{1}^p$ for some $p \in \mathbb{N}$ (since every suffix of the word $\mathbf{1}^{k-1}$ has this form). Consider this p. The word v is nonempty; thus, $v \neq \varnothing$. We have $p \neq 0$ (since otherwise, we would have

$$\begin{aligned} v = \mathbf{1}^p &= \mathbf{1}^0 \qquad (\text{since } p = 0) \\ &= \varnothing, \end{aligned}$$

contradicting $v \neq \varnothing$). Hence, $p \geq 1$, so that $p - 1 \geq 0$ and thus $\mathbf{1}^p = \mathbf{1}\mathbf{1}^{p-1}$. Now, $\mathbf{0} < \mathbf{1}$. Hence, Proposition 6.1.2(d) (applied to $\mathfrak{B}$, $\mathbf{0}$, $\mathbf{1}^{k-1}$, $\mathbf{1}$ and $\mathbf{1}^{p-1}$ instead of $\mathfrak{A}$, a, b, c and d) yields that either we have $\mathbf{0}\mathbf{1}^{k-1} \leq \mathbf{1}\mathbf{1}^{p-1}$ or the word $\mathbf{0}$ is a prefix of $\mathbf{1}$. Since the word $\mathbf{0}$ is not a prefix of $\mathbf{1}$, we thus obtain $\mathbf{0}\mathbf{1}^{k-1} \leq \mathbf{1}\mathbf{1}^{p-1}$, so that $\widetilde{k} = \mathbf{0}\mathbf{1}^{k-1} \leq \mathbf{1}\mathbf{1}^{p-1} = \mathbf{1}^p = v$. Hence, $v \geq \widetilde{k}$. Since $v \neq \widetilde{k}$ (because v is a **proper** suffix of $\widetilde{k}$), this yields $v > \widetilde{k}$.

Now, let us forget that we fixed v. We thus have proven that every nonempty proper suffix v of $\widetilde{k}$ satisfies $v > \widetilde{k}$. Since the word $\widetilde{k}$ is nonempty, this shows that the word $\widetilde{k}$ is Lyndon (by the definition of a Lyndon word), qed.

This map Φ is clearly a monoid homomorphism[1124]. Hence, $\Phi(\varnothing) = \varnothing$. Also, the map Φ is strictly order-preserving (with respect to the lexicographical orders on $\mathfrak{A}^*$ and $\mathfrak{B}^*$) [1125]. Consequently, the map Φ is injective[1126].

[1124] Indeed, we could just as well have defined Φ as the unique monoid homomorphism $\mathfrak{A}^* \to \mathfrak{B}^*$ which sends every $k \in \mathfrak{A}$ to $\widetilde{k} \in \mathfrak{B}^*$. This definition makes sense since $\mathfrak{A}^*$ is the free monoid on $\mathfrak{A}$.

[1125] *Proof.* Let u and v be two words in $\mathfrak{A}^*$ satisfying $u < v$. We are going to prove that $\Phi(u) < \Phi(v)$ in $\mathfrak{B}^*$.

We have $u < v$. Thus, $u \leq v$. By the definition of the relation $\leq$, this means that

$$\begin{aligned}
&\textbf{either } \text{there exists an } i \in \{1,2,\ldots,\min\{\ell(u),\ell(v)\}\} \\
&\qquad \text{such that } (u_i < v_i, \text{ and every } j \in \{1,2,\ldots,i-1\} \text{ satisfies } u_j = v_j), \\
&\textbf{or } \text{the word } u \text{ is a prefix of } v.
\end{aligned}$$

We thus must be in one of the following two cases:

Case 1: There exists an $i \in \{1,2,\ldots,\min\{\ell(u),\ell(v)\}\}$ such that ($u_i < v_i$, and every $j \in \{1,2,\ldots,i-1\}$ satisfies $u_j = v_j$).

Case 2: The word u is a prefix of v.

Let us first consider Case 1. In this case, there exists an $i \in \{1,2,\ldots,\min\{\ell(u),\ell(v)\}\}$ such that ($u_i < v_i$, and every $j \in \{1,2,\ldots,i-1\}$ satisfies $u_j = v_j$). Consider this i. We have $\widetilde{a} < \widetilde{b}$ for any two elements a and b of $\mathfrak{A}$ satisfying $a < b$ (because $\widetilde{1} < \widetilde{2} < \widetilde{3} < \cdots$). Applying this to $a = u_i$ and $b = v_i$, we obtain $\widetilde{u_i} < \widetilde{v_i}$.

But let $g = \widetilde{u_1}\widetilde{u_2}\cdots\widetilde{u_{i-1}}$. Every $j \in \{1,2,\ldots,i-1\}$ satisfies $u_j = v_j$. Thus, every $j \in \{1,2,\ldots,i-1\}$ satisfies $\widetilde{u_j} = \widetilde{v_j}$. Taking the product of these equalities over all $j \in \{1,2,\ldots,i-1\}$, we obtain $\widetilde{u_1}\widetilde{u_2}\cdots\widetilde{u_{i-1}} = \widetilde{v_1}\widetilde{v_2}\cdots\widetilde{v_{i-1}}$, so that $g = \widetilde{u_1}\widetilde{u_2}\cdots\widetilde{u_{i-1}} = \widetilde{v_1}\widetilde{v_2}\cdots\widetilde{v_{i-1}}$.

We have $u_i < v_i$. Thus, $\widetilde{u_i}\left(\widetilde{u_{i+1}}\widetilde{u_{i+2}}\cdots\widetilde{u_{\ell(u)}}\right) < \widetilde{v_i}\left(\widetilde{v_{i+1}}\widetilde{v_{i+2}}\cdots\widetilde{v_{\ell(v)}}\right)$ (by (13.167.11), applied to $a = u_i$, $b = v_i$, $s = \ell(u) - i$, $c_k = u_{i+k}$ and $w = \widetilde{v_{i+1}}\widetilde{v_{i+2}}\cdots\widetilde{v_{\ell(v)}}$). Thus,

$$\widetilde{u_i}\widetilde{u_{i+1}}\cdots\widetilde{u_{\ell(u)}} = \widetilde{u_i}\left(\widetilde{u_{i+1}}\widetilde{u_{i+2}}\cdots\widetilde{u_{\ell(u)}}\right) < \widetilde{v_i}\left(\widetilde{v_{i+1}}\widetilde{v_{i+1}}\cdots\widetilde{v_{\ell(v)}}\right) = \widetilde{v_i}\widetilde{v_{i+1}}\cdots\widetilde{v_{\ell(v)}}.$$

Hence, Proposition 6.1.2(b) (applied to $\mathfrak{B}$, g, $\widetilde{u_i}\widetilde{u_{i+1}}\cdots\widetilde{u_{\ell(u)}}$ and $\widetilde{v_i}\widetilde{v_{i+1}}\cdots\widetilde{v_{\ell(v)}}$ instead of $\mathfrak{A}$, a, c and d) yields

$$g\left(\widetilde{u_i}\widetilde{u_{i+1}}\cdots\widetilde{u_{\ell(u)}}\right) \leq g\left(\widetilde{v_i}\widetilde{v_{i+1}}\cdots\widetilde{v_{\ell(v)}}\right).$$

Moreover, $g\left(\widetilde{u_i}\widetilde{u_{i+1}}\cdots\widetilde{u_{\ell(u)}}\right) \neq g\left(\widetilde{v_i}\widetilde{v_{i+1}}\cdots\widetilde{v_{\ell(v)}}\right)$ (because otherwise, we would have $g\left(\widetilde{u_i}\widetilde{u_{i+1}}\cdots\widetilde{u_{\ell(u)}}\right) = g\left(\widetilde{v_i}\widetilde{v_{i+1}}\cdots\widetilde{v_{\ell(v)}}\right)$, and thus (by cancelling g from the equality $g\left(\widetilde{u_i}\widetilde{u_{i+1}}\cdots\widetilde{u_{\ell(u)}}\right) = g\left(\widetilde{v_i}\widetilde{v_{i+1}}\cdots\widetilde{v_{\ell(v)}}\right)$) we would obtain $\widetilde{u_i}\widetilde{u_{i+1}}\cdots\widetilde{u_{\ell(u)}} = \widetilde{v_i}\widetilde{v_{i+1}}\cdots\widetilde{v_{\ell(v)}}$, which would contradict $\widetilde{u_i}\widetilde{u_{i+1}}\cdots\widetilde{u_{\ell(u)}} < \widetilde{v_i}\widetilde{v_{i+1}}\cdots\widetilde{v_{\ell(v)}}$). Combining this with $g\left(\widetilde{u_i}\widetilde{u_{i+1}}\cdots\widetilde{u_{\ell(u)}}\right) \leq g\left(\widetilde{v_i}\widetilde{v_{i+1}}\cdots\widetilde{v_{\ell(v)}}\right)$, we obtain $g\left(\widetilde{u_i}\widetilde{u_{i+1}}\cdots\widetilde{u_{\ell(u)}}\right) < g\left(\widetilde{v_i}\widetilde{v_{i+1}}\cdots\widetilde{v_{\ell(v)}}\right)$.

But $u = \left(u_1, u_2, \ldots, u_{\ell(u)}\right)$, so that

$$\begin{aligned}
\Phi(u) &= \Phi\left(\left(u_1, u_2, \ldots, u_{\ell(u)}\right)\right) \qquad \text{(by the definition of } \Phi(u)) \\
&= \widetilde{u_1}\widetilde{u_2}\cdots\widetilde{u_{\ell(u)}} = \underbrace{\left(\widetilde{u_1}\widetilde{u_2}\cdots\widetilde{u_{i-1}}\right)}_{=g}\left(\widetilde{u_i}\widetilde{u_{i+1}}\cdots\widetilde{u_{\ell(u)}}\right) = g\left(\widetilde{u_i}\widetilde{u_{i+1}}\cdots\widetilde{u_{\ell(u)}}\right) \\
&< \underbrace{g}_{=\widetilde{v_1}\widetilde{v_2}\cdots\widetilde{v_{i-1}}}\left(\widetilde{v_i}\widetilde{v_{i+1}}\cdots\widetilde{v_{\ell(v)}}\right) = \left(\widetilde{v_1}\widetilde{v_2}\cdots\widetilde{v_{i-1}}\right)\left(\widetilde{v_i}\widetilde{v_{i+1}}\cdots\widetilde{v_{\ell(v)}}\right) = \widetilde{v_1}\widetilde{v_2}\cdots\widetilde{v_{\ell(v)}} = \Phi(v)
\end{aligned}$$

(since $\Phi(v) = \widetilde{v_1}\widetilde{v_2}\cdots\widetilde{v_{\ell(v)}}$ (by the definition of $\Phi(v)$)). Thus, $\Phi(u) < \Phi(v)$ is proven in Case 1.

Let us now consider Case 2. In this case, the word u is a prefix of v. That is, there exists a word $r \in \mathfrak{A}^*$ such that $v = ur$. Consider this r. We have $\Phi\left(\underbrace{v}_{=ur}\right) = \Phi(ur) = \Phi(u)\Phi(r)$ (since Φ is a monoid homomorphism). If we had $r = \varnothing$, then we would have $v = u\underbrace{r}_{=\varnothing} = u < v$, which is absurd. Hence, we cannot have $r = \varnothing$. In other words, r is nonempty, so that $\ell(r) \geq 1$.

Now, let us check that $\Phi(r)$ is nonempty. In fact, $r = \left(r_1, r_2, \ldots, r_{\ell(r)}\right)$, so that $\Phi(r) = \widetilde{r_1}\widetilde{r_2}\cdots\widetilde{r_{\ell(r)}}$ (by the definition of $\Phi(r)$). Notice that r_1 is well-defined, since $\ell(r) \geq 1$. From (13.167.8) (applied to $k = r_1$), we obtain $\widetilde{r_1} = \mathbf{0}\mathbf{1}^{r_1 - 1}$, so that $\mathbf{0}$ is a prefix of the word $\widetilde{r_1}$. Since $\widetilde{r_1}$ (in turn) is a prefix of the word $\widetilde{r_1}\widetilde{r_2}\cdots\widetilde{r_{\ell(r)}}$, this yields that $\mathbf{0}$ is a prefix of the word $\widetilde{r_1}\widetilde{r_2}\cdots\widetilde{r_{\ell(r)}}$. In other words, $\mathbf{0}$ is a prefix of the word $\Phi(r)$ (since $\Phi(r) = \widetilde{r_1}\widetilde{r_2}\cdots\widetilde{r_{\ell(r)}}$). Hence, the word $\Phi(r)$ has a nonempty prefix (namely, $\mathbf{0}$). Thus, the word $\Phi(r)$ is nonempty. In other words, $\Phi(r) \neq \varnothing$. That is, $\varnothing \neq \Phi(r)$.

Since $\Phi(v) = \Phi(u)\Phi(r)$, we now see that the word $\Phi(u)$ is a prefix of $\Phi(v)$ (since $\Phi(r)$ is nonempty). Hence, $\Phi(u) \leq \Phi(v)$. On the other hand, if we had $\Phi(u) = \Phi(v)$, then we would have $\Phi(u)\varnothing = \Phi(u) = \Phi(v) = \Phi(u)\Phi(r)$, and therefore we would have $\varnothing = \Phi(r)$ (as a result of cancelling $\Phi(u)$ from the equality $\Phi(u)\varnothing = \Phi(u)\Phi(r)$), which would contradict $\varnothing \neq \Phi(r)$. Hence, we cannot have $\Phi(u) = \Phi(v)$. We therefore must have $\Phi(u) \neq \Phi(v)$. Hence, $\Phi(u) \leq \Phi(v)$ becomes $\Phi(u) < \Phi(v)$. We thus have proven $\Phi(u) < \Phi(v)$ in Case 2.

Now, we have proven that $\Phi(u) < \Phi(v)$ in each of the two Cases 1 and 2. Since these two Cases cover all possibilities, this shows that $\Phi(u) < \Phi(v)$ always holds.

For every $w \in \mathfrak{B}^*$, we denote by $w\mathfrak{B}^*$ the set $\{wu \mid u \in \mathfrak{B}^*\}$ (that is, the set of all words in $\mathfrak{B}^*$ which have w as a prefix). Every nonempty word in $\mathfrak{B}^*$ starts with either the letter $\mathbf{0}$ or the letter $\mathbf{1}$, but not both. In other words, every nonempty word in $\mathfrak{B}^*$ belongs to either $\mathbf{0}\mathfrak{B}^*$ or $\mathbf{1}\mathfrak{B}^*$ (where $\mathbf{0}$ and $\mathbf{1}$ are regarded as one-letter words), but not both. In other words, $\mathfrak{B}^* \setminus \{\varnothing\} = \mathbf{0}\mathfrak{B}^* \cup \mathbf{1}\mathfrak{B}^*$ and $\mathbf{0}\mathfrak{B}^* \cap \mathbf{1}\mathfrak{B}^* = \varnothing$. Thus, $\mathfrak{B}^* \setminus \mathbf{1}\mathfrak{B}^* = \{\varnothing\} \cup \mathbf{0}\mathfrak{B}^*$ [1127].

Now, let us forget that we fixed u and v. We thus have proven that if u and v are two words in $\mathfrak{A}^*$ satisfying $u < v$, then $\Phi(u) < \Phi(v)$ in $\mathfrak{B}^*$. In other words, the map Φ is strictly order-preserving, qed.

[1126] because any strictly order-preserving map from a totally ordered set to a poset must be injective

[1127] *Proof.* We have $\mathfrak{B}^* \setminus \{\varnothing\} = \mathbf{0}\mathfrak{B}^* \cup \mathbf{1}\mathfrak{B}^*$ and $\mathbf{0}\mathfrak{B}^* \cap \mathbf{1}\mathfrak{B}^* = \varnothing$. Thus, the sets $\mathbf{0}\mathfrak{B}^*$ and $\mathbf{1}\mathfrak{B}^*$ are disjoint and have union $\mathfrak{B}^* \setminus \{\varnothing\}$. In other words, the sets $\mathbf{0}\mathfrak{B}^*$ and $\mathbf{1}\mathfrak{B}^*$ are complementary subsets of $\mathfrak{B}^* \setminus \{\varnothing\}$. Hence, $(\mathbf{0}\mathfrak{B}^* \cup \mathbf{1}\mathfrak{B}^*) \setminus \mathbf{1}\mathfrak{B}^* = \mathbf{0}\mathfrak{B}^*$. Now,

$$\begin{aligned}\underbrace{\mathfrak{B}^*}_{=\{\varnothing\}\cup(\mathfrak{B}^*\setminus\{\varnothing\})} \setminus \mathbf{1}\mathfrak{B}^* &= (\{\varnothing\}\cup(\mathfrak{B}^*\setminus\{\varnothing\}))\setminus\mathbf{1}\mathfrak{B}^* = \underbrace{(\{\varnothing\}\setminus\mathbf{1}\mathfrak{B}^*)}_{=\{\varnothing\}} \cup \left(\underbrace{(\mathfrak{B}^*\setminus\{\varnothing\})}_{=(\mathbf{0}\mathfrak{B}^*\cup\mathbf{1}\mathfrak{B}^*)}\setminus\mathbf{1}\mathfrak{B}^*\right) \\ &= \{\varnothing\}\cup\underbrace{((\mathbf{0}\mathfrak{B}^*\cup\mathbf{1}\mathfrak{B}^*)\setminus\mathbf{1}\mathfrak{B}^*)}_{=\mathbf{0}\mathfrak{B}^*} = \{\varnothing\}\cup\mathbf{0}\mathfrak{B}^*,\end{aligned}$$

qed.

Furthermore, $\Phi\left(\mathfrak{A}^*\right) \subset \mathfrak{B}^* \setminus \mathbf{1}\mathfrak{B}^*$ [1128] and $\mathfrak{B}^* \setminus \mathbf{1}\mathfrak{B}^* \subset \Phi\left(\mathfrak{A}^*\right)$ [1129]. Combining these two relations, we obtain

$$\Phi\left(\mathfrak{A}^*\right) = \mathfrak{B}^* \setminus \mathbf{1}\mathfrak{B}^*. \tag{13.167.15}$$

[1128] *Proof.* Let $u \in \Phi\left(\mathfrak{A}^*\right)$. We will prove that $u \in \mathfrak{B}^* \setminus \mathbf{1}\mathfrak{B}^*$.

Indeed, there exists some $w \in \mathfrak{A}^*$ satisfying $u = \Phi(w)$ (since $u \in \Phi\left(\mathfrak{A}^*\right)$). Consider this w. If $w = \varnothing$, then $u = \Phi\left(\underbrace{w}_{=\varnothing}\right) = \Phi(\varnothing) = \varnothing \in \mathfrak{B}^* \setminus \mathbf{1}\mathfrak{B}^*$ (since clearly, $\varnothing \notin \mathbf{1}\mathfrak{B}^*$). Hence, if $w = \varnothing$, then $u \in \mathfrak{B}^* \setminus \mathbf{1}\mathfrak{B}^*$ is proven. Thus, for the rest of this proof, we can WLOG assume that we don't have $w = \varnothing$. Assume this.

The word w is nonempty (since $w \neq \varnothing$), so that $\ell(w) \geq 1$. Hence, the letter w_1 is well-defined. We have $w = \left(w_1, w_2, \ldots, w_{\ell(w)}\right)$ and thus

$$\begin{aligned}\Phi(w) &= \widetilde{w_1}\widetilde{w_2}\cdots\widetilde{w_{\ell(w)}} \qquad \text{(by the definition of } \Phi\text{)} \\ &= \underbrace{\widetilde{w_1}}_{\substack{=\mathbf{0}\mathbf{1}^{w_1-1} \\ \text{(by (13.167.8), applied} \\ \text{to } k=w_1\text{)}}} \left(\widetilde{w_2}\widetilde{w_3}\cdots\widetilde{w_{\ell(w)}}\right) = \mathbf{0}\underbrace{\mathbf{1}^{w_1-1}\left(\widetilde{w_2}\widetilde{w_3}\cdots\widetilde{w_{\ell(w)}}\right)}_{\in\mathfrak{B}^*} \in \mathbf{0}\mathfrak{B}^* \subset \{\varnothing\} \cup \mathbf{0}\mathfrak{B}^* = \mathfrak{B}^* \setminus \mathbf{1}\mathfrak{B}^*.\end{aligned}$$

Hence, $u = \Phi(w) \in \mathfrak{B}^* \setminus \mathbf{1}\mathfrak{B}^*$.

Now, let us forget that we fixed u. We thus have proven that $u \in \mathfrak{B}^* \setminus \mathbf{1}\mathfrak{B}^*$ for every $u \in \Phi\left(\mathfrak{A}^*\right)$. In other words, $\Phi\left(\mathfrak{A}^*\right) \subset \mathfrak{B}^* \setminus \mathbf{1}\mathfrak{B}^*$, qed.

[1129] *Proof.* We are going to show that

$$\text{every } u \in \mathfrak{B}^* \setminus \mathbf{1}\mathfrak{B}^* \text{ satisfies } u \in \Phi\left(\mathfrak{A}^*\right). \tag{13.167.13}$$

Proof of (13.167.13): We will prove (13.167.13) by strong induction over $\ell(u)$:

Induction step: Let $N \in \mathbb{N}$. Assume that (13.167.13) holds whenever $\ell(u) < N$. We now will prove that (13.167.13) holds whenever $\ell(u) = N$.

We know that (13.167.13) holds whenever $\ell(u) < N$. In other words,

$$\text{every } u \in \mathfrak{B}^* \setminus \mathbf{1}\mathfrak{B}^* \text{ satisfying } \ell(u) < N \text{ satisfies } u \in \Phi\left(\mathfrak{A}^*\right). \tag{13.167.14}$$

Now, fix an $u \in \mathfrak{B}^* \setminus \mathbf{1}\mathfrak{B}^*$ satisfying $\ell(u) = N$. We shall show that $u \in \Phi\left(\mathfrak{A}^*\right)$.

If $u = \varnothing$, then $u = \varnothing = \Phi(\varnothing) \in \Phi\left(\mathfrak{A}^*\right)$. Hence, for the rest of our proof of $u \in \Phi\left(\mathfrak{A}^*\right)$, we can WLOG assume that we don't have $u = \varnothing$. Assume this.

We have $u \neq \varnothing$ (since we don't have $u = \varnothing$), thus $u \notin \{\varnothing\}$. We have $u \in \mathfrak{B}^* \setminus \mathbf{1}\mathfrak{B}^* = \{\varnothing\} \cup \mathbf{0}\mathfrak{B}^*$ but $u \notin \{\varnothing\}$. Hence, $u \in \left(\{\varnothing\} \cup \mathbf{0}\mathfrak{B}^*\right) \setminus \{\varnothing\} \subset \mathbf{0}\mathfrak{B}^*$. In other words, there exists a $p \in \mathfrak{B}^*$ such that $u = \mathbf{0}p$. Consider this p. The one-letter word $\mathbf{0}$ is a prefix of u (since $u = \mathbf{0}p$).

Applying (13.167.8) to $k = 1$, we obtain $\widetilde{1} = \mathbf{0}\underbrace{\mathbf{1}^{1-1}}_{=\mathbf{1}^0=\varnothing} = \mathbf{0}$, so that $\widetilde{1}$ is a prefix of u (since $\mathbf{0}$ is a prefix of u). Hence, there exists a $k \in \mathfrak{A}$ such that $\widetilde{k}$ is a prefix of u (namely, $k = 1$).

On the other hand, if $k \in \mathfrak{A}$ is such that $\widetilde{k}$ is a prefix of u, then $\ell\left(\widetilde{k}\right) \leq \ell(u)$. Hence, if $k \in \mathfrak{A}$ is such that $\widetilde{k}$ is a prefix of u, then $\ell(u) \geq \ell\left(\widetilde{k}\right) = k$ (by (13.167.9)). Hence, if $k \in \mathfrak{A}$ is such that $\widetilde{k}$ is a prefix of u, then $k \leq \ell(u)$. Hence, only finitely many $k \in \mathfrak{A}$ have the property that $\widetilde{k}$ is a prefix of u (because only finitely many $k \in \mathfrak{A}$ have the property that $k \leq \ell(u)$).

We now have made the following two observations:

- There exists a $k \in \mathfrak{A}$ such that $\widetilde{k}$ is a prefix of u.
- Only finitely many $k \in \mathfrak{A}$ have the property that $\widetilde{k}$ is a prefix of u.

Combining these two observations, we conclude that there exists a **largest** $k \in \mathfrak{A}$ such that $\widetilde{k}$ is a prefix of u. Consider this largest k. Then, $\widetilde{k}$ is a prefix of u, but $\widetilde{k+1}$ is not a prefix of u.

There exists a $v \in \mathfrak{B}^*$ such that $u = \widetilde{k}v$ (since $\widetilde{k}$ is a prefix of u). Consider this v. We have $\ell\left(\underbrace{u}_{=\widetilde{k}v}\right) = \ell\left(\widetilde{k}v\right) = \underbrace{\ell\left(\widetilde{k}\right)}_{\substack{=k \\ \text{(by (13.167.9))}}} + \ell(v) = \underbrace{k}_{\substack{\geq 1 \\ \text{(since } k\in\mathfrak{A}\text{)}}} + \ell(v) \geq 1 + \ell(v) > \ell(v)$, so that $\ell(v) < \ell(u) = N$.

Let us now assume that $v \in \mathfrak{B}^* \setminus \mathbf{1}\mathfrak{B}^*$. Then, $v \in \Phi\left(\mathfrak{A}^*\right)$ (by (13.167.14), applied to v instead of u). In other words, there exists a $v' \in \mathfrak{A}^*$ such that $v = \Phi(v')$. Consider this v'. The definition of $\Phi(k)$ (where k stands for the one-letter word $(k) \in \mathfrak{A}^*$)

Next, we recall that $\mathfrak{L}_{\mathfrak{B}}$ is the set of all Lyndon words over the alphabet $\mathfrak{B}$. We notice that $\Phi\left(\mathfrak{L}\right) \subset \mathfrak{L}_{\mathfrak{B}} \setminus \{\mathbf{1}\}$ [1130] and $\mathfrak{L}_{\mathfrak{B}} \setminus \{\mathbf{1}\} \subset \Phi\left(\mathfrak{L}\right)$ [1131]. Combining the latter two relations, we obtain

$$\Phi\left(\mathfrak{L}\right) = \mathfrak{L}_{\mathfrak{B}} \setminus \{\mathbf{1}\} . \tag{13.167.16}$$

Hence, the map Φ restricts to a bijection $\mathfrak{L} \to \mathfrak{L}_{\mathfrak{B}} \setminus \{\mathbf{1}\}$.

Next, we notice that

$$\ell\left(\Phi\left(w\right)\right) = \left|w\right| \qquad \text{for every } w \in \mathfrak{A}^* \tag{13.167.17}$$

yields $\Phi\left(k\right) = \widetilde{k}$. Hence,

$$u = \underbrace{\widetilde{k}}_{=\Phi(k)} \underbrace{v}_{=\Phi(v')} = \Phi\left(k\right) \Phi\left(v'\right) = \Phi\left(kv'\right) \qquad (\text{since } \Phi \text{ is a monoid homomorphism})$$
$$\in \Phi\left(\mathfrak{A}^*\right).$$

Now, let us forget that we have assumed that $v \in \mathfrak{B}^* \setminus \mathbf{1}\mathfrak{B}^*$. We have thus proven that $u \in \Phi\left(\mathfrak{A}^*\right)$ under the assumption that $v \in \mathfrak{B}^* \setminus \mathbf{1}\mathfrak{B}^*$. Thus, for the rest of the proof of $u \in \Phi\left(\mathfrak{A}^*\right)$, we can WLOG assume that we don't have $v \in \mathfrak{B}^* \setminus \mathbf{1}\mathfrak{B}^*$. Assume this.

We have $v \in \mathbf{1}\mathfrak{B}^*$ (since we don't have $v \in \mathfrak{B}^* \setminus \mathbf{1}\mathfrak{B}^*$). Thus, there exists a $v' \in \mathfrak{B}^*$ such that $v = \mathbf{1}v'$. Consider this v'. We have

$$u = \widetilde{k} \underbrace{v}_{=\mathbf{1}v'} = \underbrace{\widetilde{k}\mathbf{1}}_{\substack{=\widetilde{k+1} \\ (\text{by (13.167.10)})}} v' = \left(\widetilde{k+1}\right) v'.$$

Hence, $\widetilde{k+1}$ is a prefix of u. This contradicts the fact that $\widetilde{k+1}$ is not a prefix of u. From this contradiction, we conclude that $u \in \Phi\left(\mathfrak{A}^*\right)$ (since ex falso quodlibet). Thus, $u \in \Phi\left(\mathfrak{A}^*\right)$ is proven.

Now, let us forget that we fixed u. We thus have shown that every $u \in \mathfrak{B}^* \setminus \mathbf{1}\mathfrak{B}^*$ satisfying $\ell\left(u\right) = N$ satisfies $u \in \Phi\left(\mathfrak{A}^*\right)$. In other words, (13.167.13) holds whenever $\ell\left(u\right) = N$. This completes the induction step. Thus, (13.167.13) is proven by induction.

But from (13.167.13), we immediately obtain $\mathfrak{B}^* \setminus \mathbf{1}\mathfrak{B}^* \subset \Phi\left(\mathfrak{A}^*\right)$, qed.

[1130] *Proof.* Let $u \in \Phi\left(\mathfrak{L}\right)$. We are going to prove that $u \in \mathfrak{L}_{\mathfrak{B}} \setminus \{\mathbf{1}\}$.

There exists a $w \in \mathfrak{L}$ such that $u = \Phi\left(w\right)$ (since $u \in \Phi\left(\mathfrak{L}\right)$). Consider this w. The word w is Lyndon (since $w \in \mathfrak{L}$), and thus nonempty. Hence, $\ell\left(w\right) \geq 1$.

Let $n = \ell\left(w\right)$. Thus, $n = \ell\left(w\right) \geq 1$. We have $w = \left(w_1, w_2, \ldots, w_{\ell(w)}\right) = \left(w_1, w_2, \ldots, w_n\right)$ (since $\ell\left(w\right) = n$).

We have $w = \left(w_1, w_2, \ldots, w_n\right)$, so that $\Phi\left(w\right) = \widetilde{w_1}\widetilde{w_2}\cdots\widetilde{w_n}$ (by the definition of $\Phi\left(w\right)$). For every $i \in \{1, 2, \ldots, n\}$, the word $\widetilde{w_i}$ is a Lyndon word over the alphabet $\mathfrak{B}$ (by (13.167.12), applied to $k = w_i$). Thus, $\widetilde{w}_1$, $\widetilde{w}_2$, $\ldots$, $\widetilde{w}_n$ are Lyndon words over the alphabet $\mathfrak{B}$.

Let now $i \in \{1, 2, \ldots, n\}$. Then, $\left(w_i, w_{i+1}, \ldots, w_n\right)$ is a nonempty suffix of w (since $w = \left(w_1, w_2, \ldots, w_n\right)$).

Recall that w is Lyndon. Hence, Corollary 6.1.15 (applied to $v = \left(w_i, w_{i+1}, \ldots, w_n\right)$) yields that $\left(w_i, w_{i+1}, \ldots, w_n\right) \geq w$ (since $\left(w_i, w_{i+1}, \ldots, w_n\right)$ is a nonempty suffix of w). Hence, $\Phi\left(\left(w_i, w_{i+1}, \ldots, w_n\right)\right) \geq \Phi\left(w\right)$ (since the map Φ is strictly order-preserving). But the definition of the map Φ yields $\Phi\left(\left(w_i, w_{i+1}, \ldots, w_n\right)\right) = \widetilde{w_i}\widetilde{w_{i+1}}\cdots\widetilde{w_n}$. Thus, $\widetilde{w_i}\widetilde{w_{i+1}}\cdots\widetilde{w_n} = \Phi\left(\left(w_i, w_{i+1}, \ldots, w_n\right)\right) \geq \Phi\left(w\right) = \widetilde{w_1}\widetilde{w_2}\cdots\widetilde{w_n}$.

Now, let us forget that we fixed i. We thus have shown that $\widetilde{w_i}\widetilde{w_{i+1}}\cdots\widetilde{w_n} \geq \widetilde{w_1}\widetilde{w_2}\cdots\widetilde{w_n}$ for every $i \in \{1, 2, \ldots, n\}$. Hence, Exercise 6.1.24 (applied to $\mathfrak{B}$ and $\widetilde{w_i}$ instead of $\mathfrak{A}$ and w_i) yields that $\widetilde{w_1}\widetilde{w_2}\cdots\widetilde{w_n}$ is a Lyndon word. In other words, u is a Lyndon word (since $u = \Phi\left(w\right) = \widetilde{w_1}\widetilde{w_2}\cdots\widetilde{w_n}$). More precisely, u is a Lyndon word over the alphabet $\mathfrak{B}$. In other words, $u \in \mathfrak{L}_{\mathfrak{B}}$ (since $\mathfrak{L}_{\mathfrak{B}}$ is the set of all Lyndon words over the alphabet $\mathfrak{B}$).

On the other hand, $u \in \Phi\left(\underbrace{\mathfrak{L}}_{\subset\mathfrak{A}^*}\right) \subset \Phi\left(\mathfrak{A}^*\right) = \mathfrak{B}^* \setminus \mathbf{1}\mathfrak{B}^*$ (by (13.167.15)), so that $u \notin \mathbf{1}\mathfrak{B}^*$. Hence, we cannot have $u = \mathbf{1}$ (because if we had $u = \mathbf{1}$, we would have $u = \mathbf{1} = \mathbf{1}\underbrace{\varnothing}_{\in\mathfrak{B}^*} \in \mathbf{1}\mathfrak{B}^*$, which would contradict $u \notin \mathbf{1}\mathfrak{B}^*$). Thus, we have $u \neq \mathbf{1}$, so that $u \notin \{\mathbf{1}\}$. Combined with $u \in \mathfrak{L}_{\mathfrak{B}}$, this yields $u \in \mathfrak{L}_{\mathfrak{B}} \setminus \{\mathbf{1}\}$.

Now, let us forget that we fixed u. We thus have proven that $u \in \mathfrak{L}_{\mathfrak{B}} \setminus \{\mathbf{1}\}$ for every $u \in \Phi\left(\mathfrak{L}\right)$. In other words, $\Phi\left(\mathfrak{L}\right) \subset \mathfrak{L}_{\mathfrak{B}} \setminus \{\mathbf{1}\}$, qed.

[1131] *Proof.* Let $u \in \mathfrak{L}_{\mathfrak{B}} \setminus \{\mathbf{1}\}$. We are going to prove that $u \in \Phi\left(\mathfrak{L}\right)$.

We have $u \in \mathfrak{L}_{\mathfrak{B}} \setminus \{\mathbf{1}\} \subset \mathfrak{L}_{\mathfrak{B}}$. Thus, u is a Lyndon word over the alphabet $\mathfrak{B}$ (since $\mathfrak{L}_{\mathfrak{B}}$ is the set of all Lyndon words over the alphabet $\mathfrak{B}$). Also, $u \notin \{\mathbf{1}\}$ (since $u \in \mathfrak{L}_{\mathfrak{B}} \setminus \{\mathbf{1}\}$), so that $u \neq \mathbf{1}$.

Let us now assume (for the sake of contradiction) that $u \in \mathbf{1}\mathfrak{B}^*$. Then, there exists a $u' \in \mathfrak{B}^*$ such that $u = \mathbf{1}u'$. Consider this u'.

The word u is nonempty (since it is Lyndon), and thus the last letter of u is well-defined. Let g be this last letter of u. Then, g is a suffix of u. Clearly, g is nonempty (when regarded as a word). Thus, Corollary 6.1.15 (applied to $\mathfrak{B}$, u and g instead of $\mathfrak{A}$, w and v) yields $g \geq u = \mathbf{1}u'$, so that $\mathbf{1}u' \leq g = g\varnothing$. Also, $\ell\left(g\right) \geq 1$ (since g is nonempty), whence $\ell\left(g\right) \geq 1 = \ell\left(\mathbf{1}\right)$ and thus $\ell\left(\mathbf{1}\right) \leq \ell\left(g\right)$. Now, Proposition 6.1.2(f) (applied to $\mathfrak{B}$, $\mathbf{1}$, u', g and $\varnothing$ instead of $\mathfrak{A}$, a, b, c and d) yields $\mathbf{1} \leq g$. Thus, $g \geq \mathbf{1}$. Since g is a single letter, this yields that $g = \mathbf{1}$ (because the only letter of $\mathfrak{B}$ which is $\geq \mathbf{1}$ is $\mathbf{1}$ itself). Thus, $\mathbf{1}u' \leq g = \mathbf{1} = \mathbf{1}\varnothing$. Now, Proposition 6.1.2(f) (applied to $\mathfrak{B}$, $\mathbf{1}$, u' and $\varnothing$ instead of $\mathfrak{A}$, a, c and d) yields $u' \leq \varnothing$. This yields $u' = \varnothing$ (since the only (where the notation $|w|$ makes sense because every $w \in \mathfrak{A}^*$ is a composition and thus has a size)[1132].

[Incidentally, here is a similar identity which will not use:

$$(13.167.18) \qquad (\text{the number of letters } \mathbf{0} \text{ in } \Phi(w)) = \ell(w) \qquad \text{for every } w \in \mathfrak{A}^*$$

[1133].]

Now, fix a positive integer n. The alphabet $\mathfrak{B}$ is finite and satisfies $2 = |\mathfrak{B}|$. Hence, Exercise 6.1.29 (applied to $\mathfrak{B}$ and 2 instead of $\mathfrak{A}$ and q) yields that the number of Lyndon words of length n over the alphabet $\mathfrak{B}$ equals $\dfrac{1}{n}\sum_{d|n}\mu(d)\,2^{n/d}$. That is,

$$(13.167.20) \qquad (\text{the number of all Lyndon words of length } n \text{ over the alphabet } \mathfrak{B}) = \frac{1}{n}\sum_{d|n}\mu(d)\,2^{n/d}.$$

word which is $\leq \varnothing$ is $\varnothing$ itself). Hence, $u = \mathbf{1}\underbrace{u'}_{=\varnothing} = \mathbf{1}$, which contradicts $u \neq \mathbf{1}$. This contradiction shows that our assumption (that $u \in \mathbf{1}\mathfrak{B}^*$) was wrong. Hence, we have $u \notin \mathbf{1}\mathfrak{B}^*$.

Since $u \in \mathfrak{B}^*$ and $u \notin \mathbf{1}\mathfrak{B}^*$, we must have $u \in \mathfrak{B}^* \setminus \mathbf{1}\mathfrak{B}^* = \Phi(\mathfrak{A}^*)$ (by (13.167.15)). In other words, there exists a $w \in \mathfrak{A}^*$ such that $u = \Phi(w)$. Consider this w.

We have $u \neq \varnothing$ (since u is nonempty) and thus $w \neq \varnothing$ (because otherwise, we would have $w = \varnothing$ and thus $u = \Phi\left(\underbrace{w}_{=\varnothing}\right) = \Phi(\varnothing) = \varnothing$, contradicting $u \neq \varnothing$). Hence, the word w is nonempty.

Let now v be a nonempty proper suffix of w. We assume (for the sake of contradiction) that $v \leq w$. Since $v \neq w$ (because v is a **proper** suffix of w) and $v \leq w$, we have $v < w$. Thus, $\Phi(v) < \Phi(w)$ (since the map Φ is strictly order-preserving). Hence, $\Phi(v) < \Phi(w) = u$.

Also, $v \neq \varnothing$ (since v is nonempty). Thus, $\Phi(v) \neq \Phi(\varnothing)$ (because otherwise, we would have $\Phi(v) = \Phi(\varnothing)$, so that $v = \varnothing$ (since the map Φ is injective), which would contradict $v \neq \varnothing$). Hence, $\Phi(v) \neq \Phi(\varnothing) = \varnothing$, so that the word $\Phi(v)$ is nonempty.

But there exists a $p \in \mathfrak{A}^*$ such that $w = pv$ (since v is a suffix of w). Consider this p. We have $u = \Phi\left(\underbrace{w}_{=pv}\right) = \Phi(pv) = \Phi(p)\,\Phi(v)$ (since Φ is a monoid homomorphism). Thus, $\Phi(v)$ is a suffix of u. Now, Corollary 6.1.15 (applied to $\mathfrak{B}$, u and $\Phi(v)$ instead of $\mathfrak{A}$, w and v) yields $\Phi(v) \geq u$. This contradicts $\Phi(v) < u$. This contradiction shows that our assumption (that $v \leq w$) was wrong. Hence, we cannot have $v \leq w$. We thus have $v > w$.

Now, let us forget that we fixed v. We thus have proven that every nonempty proper suffix v of w satisfies $v > w$. Since the word w is nonempty, this yields that the word w is Lyndon (by the definition of a Lyndon word). In other words, $w \in \mathfrak{L}$ (since $\mathfrak{L}$ is the set of all Lyndon words over the alphabet $\mathfrak{A}$). Now, $u = \Phi\left(\underbrace{w}_{\in\mathfrak{L}}\right) \in \Phi(\mathfrak{L})$.

Now, let us forget that we fixed u. We thus have shown that $u \in \Phi(\mathfrak{L})$ for every $u \in \mathfrak{L}_{\mathfrak{B}} \setminus \{\mathbf{1}\}$. In other words, $\mathfrak{L}_{\mathfrak{B}} \setminus \{\mathbf{1}\} \subset \Phi(\mathfrak{L})$, qed.

[1132] *Proof.* Let $w \in \mathfrak{A}^*$. Then, $w = \left(w_1, w_2, \ldots, w_{\ell(w)}\right)$, so that $\Phi(w) = \widetilde{w_1}\widetilde{w_2}\cdots\widetilde{w_{\ell(w)}}$ (by the definition of $\Phi(w)$). Hence,

$$\ell\left(\underbrace{\Phi(w)}_{=\widetilde{w_1}\widetilde{w_2}\cdots\widetilde{w_{\ell(w)}}}\right) = \ell\left(\widetilde{w_1}\widetilde{w_2}\cdots\widetilde{w_{\ell(w)}}\right) = \ell(\widetilde{w_1}) + \ell(\widetilde{w_2}) + \cdots + \ell\left(\widetilde{w_{\ell(w)}}\right) = \sum_{i=1}^{\ell(w)} \underbrace{\ell(\widetilde{w_i})}_{\substack{=w_i \\ \text{(by (13.167.9), applied to } k=w_i)}}$$
$$= \sum_{i=1}^{\ell(w)} w_i = w_1 + w_2 + \cdots + w_{\ell(w)} = |w|$$

(since $|w| = w_1 + w_2 + \cdots + w_{\ell(w)}$ (since $w = \left(w_1, w_2, \ldots, w_{\ell(w)}\right)$)), qed.

[1133] *Proof.* Let $w \in \mathfrak{A}^*$.

We notice that

$$(13.167.19) \qquad \left(\text{the number of letters } \mathbf{0} \text{ in } \underbrace{\widetilde{k}}_{\substack{=\mathbf{0}\mathbf{1}^{k-1} \\ \text{(by (13.167.8))}}}\right) = \left(\text{the number of letters } \mathbf{0} \text{ in } \mathbf{0}\mathbf{1}^{k-1}\right) = 1$$

for every $k \in \mathfrak{A}$.

Now, let $\operatorname{lync} n$ denote the number of Lyndon compositions of size n. We need to prove that

$$\operatorname{lync} n = \frac{1}{n}\sum_{d\mid n}\mu(d)\left(2^{n/d}-1\right) = \frac{1}{n}\sum_{d\mid n}\mu(d)\,2^{n/d}-\delta_{n,1}. \qquad (13.167.21)$$

By the definition of $\operatorname{lync} n$, we have

$$\begin{aligned}
\operatorname{lync} n &= (\text{the number of Lyndon compositions of size } n)\\
&= (\text{the number of all } w\in\mathfrak{L} \text{ such that } |w|=n)\\
&\qquad (\text{since the set of all Lyndon compositions is } \mathfrak{L})\\
&= \left|\left\{ w\in\mathfrak{L} \ \mid\ \underbrace{|w|}_{\substack{=\ell(\Phi(w))\\ (\text{by (13.167.17)})}} = n\right\}\right|\\
&= \left|\left\{ w\in\mathfrak{L} \ \mid\ \ell(\Phi(w))=n\right\}\right| = \left|\underbrace{\left\{ w\in\mathfrak{L}_{\mathfrak{B}}\setminus\{\mathbf{1}\} \ \mid\ \ell(w)=n\right\}}_{=\{w\in\mathfrak{L}_{\mathfrak{B}}\ \mid\ \ell(w)=n\}\setminus\{w\in\{\mathbf{1}\}\ \mid\ \ell(w)=n\}}\right|\\
&\qquad \left(\begin{array}{c}\text{here, we have substituted } w \text{ for } \Phi(w)\text{, since the map}\\ \Phi \text{ restricts to a bijection } \mathfrak{L}\to\mathfrak{L}_{\mathfrak{B}}\setminus\{\mathbf{1}\}\end{array}\right)\\
&= \left|\left\{ w\in\mathfrak{L}_{\mathfrak{B}} \ \mid\ \ell(w)=n\right\}\setminus\left\{ w\in\{\mathbf{1}\} \ \mid\ \ell(w)=n\right\}\right|\\
&= \underbrace{\left|\left\{ w\in\mathfrak{L}_{\mathfrak{B}} \ \mid\ \ell(w)=n\right\}\right|}_{\substack{=(\text{the number of all } w\in\mathfrak{L}_{\mathfrak{B}} \text{ such that } \ell(w)=n)\\ =(\text{the number of all Lyndon words of length } n \text{ over the alphabet } \mathfrak{B})\\ (\text{since } \mathfrak{L}_B \text{ is the set of all Lyndon words over the alphabet } \mathfrak{B})}} - \underbrace{\left|\left\{ w\in\{\mathbf{1}\} \ \mid\ \ell(w)=n\right\}\right|}_{\substack{=\delta_{n,1}\\ (\text{since } \ell(\mathbf{1})=1)}}\\
&\qquad \left(\text{since } \left\{ w\in\underbrace{\{\mathbf{1}\}}_{\subset\mathfrak{L}_{\mathfrak{B}}} \ \mid\ \ell(w)=n\right\} \subset \left\{ w\in\mathfrak{L}_{\mathfrak{B}} \ \mid\ \ell(w)=n\right\}\right)\\
&= \underbrace{(\text{the number of all Lyndon words of length } n \text{ over the alphabet } \mathfrak{B})}_{\substack{=\frac{1}{n}\sum_{d\mid n}\mu(d)2^{n/d}\\ (\text{by (13.167.20)})}} - \delta_{n,1}\\
&= \frac{1}{n}\sum_{d\mid n}\mu(d)\,2^{n/d}-\delta_{n,1} = \frac{1}{n}\sum_{d\mid n}\mu(d)\left(2^{n/d}-1\right)
\end{aligned}$$

(by Lemma 13.167.2). This solves Exercise 6.4.2 again.

We have $w=\left(w_1,w_2,\ldots,w_{\ell(w)}\right)$, so that $\Phi(w)=\widetilde{w_1}\widetilde{w_2}\cdots\widetilde{w_{\ell(w)}}$ (by the definition of $\Phi(w)$). Hence,

$$\begin{aligned}
&\left(\text{the number of letters } \mathbf{0} \text{ in } \underbrace{\Phi(w)}_{=\widetilde{w_1}\widetilde{w_2}\cdots\widetilde{w_{\ell(w)}}}\right)\\
&= \left(\text{the number of letters } \mathbf{0} \text{ in } \widetilde{w_1}\widetilde{w_2}\cdots\widetilde{w_{\ell(w)}}\right)\\
&= \left(\text{the number of letters } \mathbf{0} \text{ in } \widetilde{w_1}\right) + \left(\text{the number of letters } \mathbf{0} \text{ in } \widetilde{w_2}\right)\\
&\qquad + \cdots + \left(\text{the number of letters } \mathbf{0} \text{ in } \widetilde{w_{\ell(w)}}\right)\\
&= \sum_{i=1}^{\ell(w)} \underbrace{\left(\text{the number of letters } \mathbf{0} \text{ in } \widetilde{w_i}\right)}_{\substack{=1\\ (\text{by (13.167.19), applied to } k=w_i)}} = \sum_{i=1}^{\ell(w)} 1 = \ell(w),
\end{aligned}$$

qed.

13.168. **Solution to Exercise 6.4.6.** *Solution to Exercise 6.4.6.*

Proof of Proposition 6.4.5. Write α and β as $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_\ell)$ and $\beta = (\beta_1, \beta_2, \ldots, \beta_m)$, respectively; then $\ell(\alpha) = \ell$ and $\ell(\beta) = m$.

Fix three disjoint chain posets $(i_1 < i_2 < \cdots < i_\ell)$, $(j_1 < j_2 < \cdots < j_m)$ and $(k_1 < k_2 < k_3 < \cdots)$. For any $p \in \mathbb{N}$, we define a *p-shuffling map* to mean a map f from the disjoint union of two chains to a chain

$$(i_1 < i_2 < \cdots < i_\ell) \sqcup (j_1 < j_2 < \cdots < j_m) \xrightarrow{f} (k_1 < k_2 < \cdots < k_p)$$

which is both surjective and strictly order-preserving (that is, if x and y are two elements in the domain of f satisfying $x < y$, then $f(x) < f(y)$). For every $p \in \mathbb{N}$ and every p-shuffling map f, we define a composition $\operatorname{wt}(f) := (\operatorname{wt}_1(f), \operatorname{wt}_2(f), \ldots, \operatorname{wt}_p(f))$ by $\operatorname{wt}_s(f) := \sum_{i_u \in f^{-1}(k_s)} \alpha_u + \sum_{j_v \in f^{-1}(k_s)} \beta_v$. Then, Proposition 5.1.3 yields

$$M_\alpha M_\beta = \sum_f M_{\operatorname{wt}(f)},$$

where the sum is over all $p \in \mathbb{N}$ and all p-shuffling maps f. In other words,

$$\begin{align*}
M_\alpha M_\beta &= \sum_{\substack{(p,f);\\ p\in\mathbb{N};\\ f \text{ is a } p\text{-shuffling map}}} M_{\operatorname{wt}(f)} = \sum_{p\in\mathbb{N}} \sum_{f \text{ is a } p\text{-shuffling map}} M_{\operatorname{wt}(f)} \\
&= \sum_{\substack{p\in\mathbb{N};\\ p\leq\ell+m}} \sum_{f \text{ is a } p\text{-shuffling map}} M_{\operatorname{wt}(f)} + \sum_{\substack{p\in\mathbb{N};\\ p>\ell+m}} \underbrace{\sum_{f \text{ is a } p\text{-shuffling map}} M_{\operatorname{wt}(f)}}_{\substack{=(\text{empty sum})\\ (\text{because there exist no } p\text{-shuffling maps if } p>\ell+m\\ (\text{since a } p\text{-shuffling map has to be a surjective map}\\ \text{from an } (\ell+m)\text{-element set to a } p\text{-element set,}\\ \text{and such maps don't exist if } p>\ell+m))}} \\
&= \sum_{\substack{p\in\mathbb{N};\\ p\leq\ell+m}} \sum_{f \text{ is a } p\text{-shuffling map}} M_{\operatorname{wt}(f)} + \sum_{\substack{p\in\mathbb{N};\\ p>\ell+m}} \underbrace{(\text{empty sum})}_{=0} \\
&= \sum_{\substack{p\in\mathbb{N};\\ p\leq\ell+m}} \sum_{f \text{ is a } p\text{-shuffling map}} M_{\operatorname{wt}(f)} + \underbrace{\sum_{\substack{p\in\mathbb{N};\\ p>\ell+m}} 0}_{=0} = \sum_{\substack{p\in\mathbb{N};\\ p\leq\ell+m}} \sum_{f \text{ is a } p\text{-shuffling map}} M_{\operatorname{wt}(f)} \tag{13.168.1} \\
&= \sum_{f \text{ is an } (\ell+m)\text{-shuffling map}} M_{\operatorname{wt}(f)} + \sum_{\substack{p\in\mathbb{N};\\ p<\ell+m}} \sum_{f \text{ is a } p\text{-shuffling map}} M_{\operatorname{wt}(f)} \tag{13.168.2}
\end{align*}$$

(here, we have split off the addend for $p = \ell + m$ from the sum).

We now notice that

$$\begin{align*}
&\sum_{\substack{p\in\mathbb{N};\\ p<\ell+m}} \sum_{f \text{ is a } p\text{-shuffling map}} M_{\operatorname{wt}(f)} \\
&= (\text{a sum of terms of the form } M_\delta \text{ with } \delta \in \mathfrak{A}^* \text{ satisfying } \ell(\delta) < \ell(\alpha) + \ell(\beta)) \tag{13.168.3}
\end{align*}$$

[1134]. Our next goal is to prove

$$\sum_{f \text{ is an } (\ell+m)\text{-shuffling map}} M_{\operatorname{wt}(f)} = \sum_{\sigma \in \operatorname{Sh}_{\ell,m}} M_{\alpha \underset{\sigma}{\sqcup\!\sqcup} \beta}. \tag{13.168.4}$$

[1134] *Proof of (13.168.3):* If $p \in \mathbb{N}$ is such that $p < \ell + m$, and if f is a p-shuffling map, then $\operatorname{wt}(f) = (\operatorname{wt}_1(f), \ldots, \operatorname{wt}_p(f))$ is a composition of length $p < \underbrace{\ell}_{=\ell(\alpha)} + \underbrace{m}_{=\ell(\beta)} = \ell(\alpha) + \ell(\beta)$. Hence, if $p \in \mathbb{N}$ is such that $p < \ell + m$, and if f is a p-shuffling map, then $M_{\operatorname{wt}(f)}$ is a term of the form M_δ with $\delta \in \mathfrak{A}^*$ satisfying $\ell(\delta) < \ell(\alpha) + \ell(\beta)$. Therefore, $\sum\limits_{\substack{p\in\mathbb{N};\\ p<\ell+m}} \sum_{f \text{ is a } p\text{-shuffling map}} M_{\operatorname{wt}(f)}$ is a sum of terms of the form M_δ with $\delta \in \mathfrak{A}^*$ satisfying $\ell(\delta) < \ell(\alpha) + \ell(\beta)$. This proves (13.168.3).

Once (13.168.4) is proven, we will immediately conclude that

$$\sum_{f \text{ is an } (\ell+m)\text{-shuffling map}} M_{\operatorname{wt}(f)} = \sum_{\sigma \in \operatorname{Sh}_{\ell,m}} M_{\alpha \underset{\sigma}{\sqcup\!\sqcup} \beta} = \sum_{\gamma \in \alpha \sqcup\!\sqcup \beta} M_\gamma$$

(since the multiset $\alpha \sqcup\!\sqcup \beta$ is defined as $\left\{\alpha \underset{\sigma}{\sqcup\!\sqcup} \beta \ : \ \sigma \in \operatorname{Sh}_{\ell,m}\right\}_{\text{multiset}}$), and therefore we will obtain

$$M_\alpha M_\beta = \underbrace{\sum_{f \text{ is an } (\ell+m)\text{-shuffling map}} M_{\operatorname{wt}(f)}}_{=\sum_{\gamma\in\alpha\sqcup\!\sqcup\beta} M_\gamma} + \underbrace{\sum_{\substack{p\in\mathbb{N};\\ p<\ell+m}} \sum_{f \text{ is a } p\text{-shuffling map}} M_{\operatorname{wt}(f)}}_{\substack{=(\text{a sum of terms of the form } M_\delta \text{ with } \delta\in\mathfrak{A}^* \text{ satisfying } \ell(\delta)<\ell(\alpha)+\ell(\beta))\\ (\text{by } (13.168.3))}}$$

$$= \sum_{\gamma\in\alpha\sqcup\!\sqcup\beta} M_\gamma + (\text{a sum of terms of the form } M_\delta \text{ with } \delta \in \mathfrak{A}^* \text{ satisfying } \ell(\delta) < \ell(\alpha) + \ell(\beta)),$$

which will complete the proof of Proposition 6.4.5. Hence, it only remains to prove (13.168.4).

We let D be the poset $(i_1 < i_2 < \cdots < i_\ell) \sqcup (j_1 < j_2 < \cdots < j_m)$, and we let R be the poset $(k_1 < k_2 < \cdots < k_{\ell+m})$. Note that $|D| = \ell + m = |R|$. Hence, D and R are two finite sets of the same cardinality. Consequently, a given map from D to R is surjective if and only if it is bijective.[1135]

Recall that an $(\ell+m)$-shuffling map means a map f from $(i_1 < i_2 < \cdots < i_\ell) \sqcup (j_1 < j_2 < \cdots < j_m)$ to $(k_1 < k_2 < \cdots < k_{\ell+m})$ which is both surjective and strictly order-preserving (by the definition of an "$(\ell+m)$-shuffling map"). In other words, an $(\ell+m)$-shuffling map means a map f from D to R which is both surjective and strictly order-preserving[1136]. In other words, an $(\ell+m)$-shuffling map means a map f from D to R which is both bijective and strictly order-preserving[1137].

Let us now define a bijection $\mathbf{d} : \{1, 2, \ldots, \ell+m\} \to D$ by

$$\mathbf{d}(u) = \begin{cases} i_u, & \text{if } u \leq \ell; \\ j_{u-\ell}, & \text{if } u > \ell \end{cases} \qquad \text{for every } u \in \{1, 2, \ldots, \ell+m\}.$$

Let us further define a bijection $\mathbf{r} : \{1, 2, \ldots, \ell+m\} \to R$ by

$$\mathbf{r}(u) = k_u \qquad \text{for every } u \in \{1, 2, \ldots, \ell+m\}.$$

Notice that $\mathbf{r}$ is an isomorphism of posets, where we endow the set $\{1, 2, \ldots, \ell+m\}$ with its natural total order (i.e., the order $1 < 2 < \cdots < \ell+m$).

We can then define a bijection

$$\Phi : (\text{the set of all bijective maps from } \{1, 2, \ldots, \ell+m\} \text{ to } \{1, 2, \ldots, \ell+m\})$$
$$\to (\text{the set of all bijective maps from } D \text{ to } R)$$

by setting

$$\Phi(\sigma) = \mathbf{r} \circ \sigma \circ \mathbf{d}^{-1} \qquad \text{for every bijective map } \sigma : \{1, 2, \ldots, \ell+m\} \to \{1, 2, \ldots, \ell+m\}.$$

[1138] Consider this bijection Φ. Then, Φ is a bijection from $\mathfrak{S}_{\ell+m}$ to (the set of all bijective maps from D to R) (because (the set of all bijective maps from $\{1, 2, \ldots, \ell+m\}$ to $\{1, 2, \ldots, \ell+m\}$) $= \mathfrak{S}_{\ell+m}$). It is now easy to see that if $\sigma \in \mathfrak{S}_{\ell+m}$, then we have the following equivalence of assertions:

$$(\sigma \in \operatorname{Sh}_{\ell,m}) \Longleftrightarrow \left(\text{the map } \Phi\left(\sigma^{-1}\right) : D \to R \text{ is strictly order-preserving}\right). \tag{13.168.5}$$

[1139] Moreover, every $\sigma \in \operatorname{Sh}_{\ell,m}$ satisfies

$$\operatorname{wt}\left(\Phi\left(\sigma^{-1}\right)\right) = \alpha \underset{\sigma}{\sqcup\!\sqcup} \beta. \tag{13.168.9}$$

[1135] Of course, the letters D and R have been chosen to remind of "domain" and "range".

[1136] since $D = (i_1 < i_2 < \cdots < i_\ell) \sqcup (j_1 < j_2 < \cdots < j_m)$ and $R = (k_1 < k_2 < \cdots < k_{\ell+m})$

[1137] This is because a given map from D to R is surjective if and only if it is bijective.

[1138] This is a bijection since both $\mathbf{d}$ and $\mathbf{r}$ are bijections.

[1139] *Proof of (13.168.5):* Let $\sigma \in \mathfrak{S}_{\ell+m}$. The poset D is the disjoint union of the totally ordered posets $(i_1 < i_2 < \cdots < i_\ell)$ and $(j_1 < j_2 < \cdots < j_m)$. Hence, if P is any other poset, and $f : D \to P$ is any map, then the map $f : D \to P$ is strictly order-preserving if and only if it satisfies

$$(f(i_1) < f(i_2) < \cdots < f(i_\ell) \text{ and } f(j_1) < f(j_2) < \cdots < f(j_m)).$$

[1140]

Now, recall that an $(\ell+m)$-shuffling map means a map f from D to R which is both bijective and strictly order-preserving. In other words, an $(\ell+m)$-shuffling map means a bijective map from D to R which is

---

Applying this to $P=R$ and $f=\Phi\left(\sigma^{-1}\right)$, we conclude that the map $\Phi\left(\sigma^{-1}\right):D\to R$ is strictly order-preserving if and only if it satisfies

$$\left(\left(\Phi\left(\sigma^{-1}\right)\right)(i_1)<\left(\Phi\left(\sigma^{-1}\right)\right)(i_2)<\cdots<\left(\Phi\left(\sigma^{-1}\right)\right)(i_\ell)\right.$$
$$\left.\text{and }\left(\Phi\left(\sigma^{-1}\right)\right)(j_1)<\left(\Phi\left(\sigma^{-1}\right)\right)(j_2)<\cdots<\left(\Phi\left(\sigma^{-1}\right)\right)(j_m)\right).$$

In other words, we have the following equivalence of assertions:

$$\begin{aligned}&\left(\text{the map }\Phi\left(\sigma^{-1}\right):D\to R\text{ is strictly order-preserving}\right)\\&\iff\left(\left(\Phi\left(\sigma^{-1}\right)\right)(i_1)<\left(\Phi\left(\sigma^{-1}\right)\right)(i_2)<\cdots<\left(\Phi\left(\sigma^{-1}\right)\right)(i_\ell)\right.\\&\qquad\left.\text{and }\left(\Phi\left(\sigma^{-1}\right)\right)(j_1)<\left(\Phi\left(\sigma^{-1}\right)\right)(j_2)<\cdots<\left(\Phi\left(\sigma^{-1}\right)\right)(j_m)\right).\end{aligned}\tag{13.168.6}$$

Now, recall that $\Phi\left(\sigma^{-1}\right)=\mathbf{r}\circ\sigma^{-1}\circ\mathbf{d}^{-1}$ (by the definition of Φ). Hence, every $u\in\{1,2,\ldots,\ell\}$ satisfies

$$\left(\Phi\left(\sigma^{-1}\right)\right)(i_u)=\left(\mathbf{r}\circ\sigma^{-1}\circ\mathbf{d}^{-1}\right)(i_u)=\mathbf{r}\left(\sigma^{-1}\left(\underbrace{\mathbf{d}^{-1}(i_u)}_{\substack{=u\\(\text{since }\mathbf{d}(u)=i_u\\(\text{by the definition of }\mathbf{d}))}}\right)\right)=\mathbf{r}\left(\sigma^{-1}(u)\right).$$

Hence, we have the following equivalence of assertions:

$$\begin{aligned}&\left(\left(\Phi\left(\sigma^{-1}\right)\right)(i_1)<\left(\Phi\left(\sigma^{-1}\right)\right)(i_2)<\cdots<\left(\Phi\left(\sigma^{-1}\right)\right)(i_\ell)\right)\\&\iff\left(\mathbf{r}\left(\sigma^{-1}(1)\right)<\mathbf{r}\left(\sigma^{-1}(2)\right)<\cdots<\mathbf{r}\left(\sigma^{-1}(\ell)\right)\right)\\&\iff\left(\sigma^{-1}(1)<\sigma^{-1}(2)<\cdots<\sigma^{-1}(\ell)\right)\qquad(\text{since }\mathbf{r}\text{ is an isomorphism of posets}).\end{aligned}\tag{13.168.7}$$

Similarly, we have the following equivalence of assertions:

$$\begin{aligned}&\left(\left(\Phi\left(\sigma^{-1}\right)\right)(j_1)<\left(\Phi\left(\sigma^{-1}\right)\right)(j_2)<\cdots<\left(\Phi\left(\sigma^{-1}\right)\right)(j_m)\right)\\&\iff\left(\sigma^{-1}(\ell+1)<\sigma^{-1}(\ell+2)<\cdots<\sigma^{-1}(\ell+m)\right).\end{aligned}\tag{13.168.8}$$

Now, the equivalence (13.168.6) becomes

$$\begin{aligned}&\left(\text{the map }\Phi\left(\sigma^{-1}\right):D\to R\text{ is strictly order-preserving}\right)\\&\iff\left(\underbrace{\left(\Phi\left(\sigma^{-1}\right)\right)(i_1)<\left(\Phi\left(\sigma^{-1}\right)\right)(i_2)<\cdots<\left(\Phi\left(\sigma^{-1}\right)\right)(i_\ell)}_{\substack{\text{this is equivalent to}\\(\sigma^{-1}(1)<\sigma^{-1}(2)<\cdots<\sigma^{-1}(\ell))\\(\text{by (13.168.7)})}}\right.\\&\qquad\left.\text{and }\underbrace{\left(\Phi\left(\sigma^{-1}\right)\right)(j_1)<\left(\Phi\left(\sigma^{-1}\right)\right)(j_2)<\cdots<\left(\Phi\left(\sigma^{-1}\right)\right)(j_m)}_{\substack{\text{this is equivalent to}\\(\sigma^{-1}(\ell+1)<\sigma^{-1}(\ell+2)<\cdots<\sigma^{-1}(\ell+m))\\(\text{by (13.168.8)})}}\right)\\&\iff\left(\sigma^{-1}(1)<\sigma^{-1}(2)<\cdots<\sigma^{-1}(\ell)\text{ and }\sigma^{-1}(\ell+1)<\sigma^{-1}(\ell+2)<\cdots<\sigma^{-1}(\ell+m)\right)\\&\iff\left(\sigma\in\operatorname{Sh}_{\ell,m}\right)\qquad\left(\text{by the definition of }\operatorname{Sh}_{\ell,m}\right).\end{aligned}$$

This proves (13.168.5).

[1140] *Proof of (13.168.9):* Let $\sigma\in\operatorname{Sh}_{\ell,m}$. Set $f=\Phi\left(\sigma^{-1}\right)$. Then, f is a bijective map from D to R (because the domain of Φ is (the set of all bijective maps from D to R)).

By the definition of $\operatorname{wt}(f)$, we have $\operatorname{wt}(f)=(\operatorname{wt}_1(f),\operatorname{wt}_2(f),\ldots,\operatorname{wt}_{\ell+m}(f))$, where we set $\operatorname{wt}_s(f):=\sum_{i_u\in f^{-1}(k_s)}\alpha_u+\sum_{j_v\in f^{-1}(k_s)}\beta_v$ for every $s\in\{1,2,\ldots,\ell+m\}$.

On the other hand, let $(\gamma_1,\gamma_2,\ldots,\gamma_{\ell+m})$ be the concatenation $\alpha\cdot\beta=(\alpha_1,\alpha_2,\ldots,\alpha_\ell,\beta_1,\beta_2,\ldots,\beta_m)$. Then, $\alpha\underset{\sigma}{\sqcup\!\sqcup}\beta=\left(\gamma_{\sigma(1)},\gamma_{\sigma(2)},\ldots,\gamma_{\sigma(\ell+m)}\right)$.

Let $s\in\{1,2,\ldots,\ell+m\}$. We are going to prove that $\gamma_{\sigma(s)}=\operatorname{wt}_s(f)$.

We must be in one of the following two cases:

strictly order-preserving. Hence,

$$\begin{aligned}
&\sum_{f \text{ is an } (\ell+m)\text{-shuffling map}} M_{\operatorname{wt}(f)} \\
&= \sum_{\substack{f \text{ is a bijective map from } D \text{ to } R;\\ f \text{ is strictly order-preserving}}} M_{\operatorname{wt}(f)} = \sum_{\substack{\sigma\in\mathfrak{S}_{\ell+m};\\ \Phi(\sigma) \text{ is strictly order-preserving}}} M_{\operatorname{wt}(\Phi(\sigma))} \\
&\qquad \left(\begin{array}{c}\text{here, we have substituted } \Phi(\sigma) \text{ for } f \text{ in the sum,}\\ \text{since } \Phi : \mathfrak{S}_{\ell+m} \to (\text{the set of all bijective maps from } D \text{ to } R)\\ \text{is a bijection}\end{array}\right) \\
&= \underbrace{\sum_{\substack{\sigma\in\mathfrak{S}_{\ell+m};\\ \Phi\left(\sigma^{-1}\right) \text{ is strictly order-preserving}}}}_{\substack{=\sum_{\substack{\sigma\in\mathfrak{S}_{\ell+m};\\ \sigma\in\operatorname{Sh}_{\ell,m}}} \\ \text{(because } \Phi\left(\sigma^{-1}\right) \text{ is strictly order-preserving if and only if } \sigma\in\operatorname{Sh}_{\ell,m} \\ \text{(according to (13.168.5)))}}} M_{\operatorname{wt}(\Phi(\sigma^{-1}))} \\
&\qquad \left(\begin{array}{c}\text{here, we have substituted } \sigma^{-1} \text{ for } \sigma \text{ in the sum,}\\ \text{since the map } \mathfrak{S}_{\ell+m} \to \mathfrak{S}_{\ell+m},\ \sigma \mapsto \sigma^{-1} \text{ is a bijection}\end{array}\right) \\
&= \underbrace{\sum_{\substack{\sigma\in\mathfrak{S}_{\ell+m};\\ \sigma\in\operatorname{Sh}_{\ell,m}}}}_{=\sum_{\sigma\in\operatorname{Sh}_{\ell,m}}} \underbrace{M_{\operatorname{wt}(\Phi(\sigma^{-1}))}}_{\substack{=M_{\alpha \underset{\sigma}{\sqcup\!\sqcup} \beta} \\ \text{(by (13.168.9))}}} = \sum_{\sigma\in\operatorname{Sh}_{\ell,m}} M_{\alpha \underset{\sigma}{\sqcup\!\sqcup} \beta}.
\end{aligned}$$

This proves (13.168.4). The proof of Proposition 6.4.5 is thus complete. □

Case 1: We have $\sigma(s) \leq \ell$.

Case 2: We have $\sigma(s) > \ell$.

Let us consider Case 1 first. In this case, we have $\sigma(s) \leq \ell$. Thus, $\sigma(s) \in \{1, 2, \ldots, \ell\}$, so that $i_{\sigma(s)}$ is well-defined. We have $\mathbf{d}(\sigma(s)) = i_{\sigma(s)}$ (by the definition of $\mathbf{d}$, since $\sigma(s) \leq \ell$), so that $\mathbf{d}^{-1}\left(i_{\sigma(s)}\right) = \sigma(s)$. Now, $f = \Phi\left(\sigma^{-1}\right) = \mathbf{r} \circ \sigma^{-1} \circ \mathbf{d}^{-1}$ (by the definition of Φ), so that

$$f\left(i_{\sigma(s)}\right) = \left(\mathbf{r} \circ \sigma^{-1} \circ \mathbf{d}^{-1}\right)\left(i_{\sigma(s)}\right) = \mathbf{r}\left(\sigma^{-1}\left(\underbrace{\mathbf{d}^{-1}\left(i_{\sigma(s)}\right)}_{=\sigma(s)}\right)\right) = \mathbf{r}\left(\underbrace{\sigma^{-1}(\sigma(s))}_{=s}\right) = \mathbf{r}(s) = k_s$$

(by the definition of $\mathbf{r}$). Since the map f is a bijection, this yields that the **set** $f^{-1}(k_s)$ equals $\left\{i_{\sigma(s)}\right\}$. Hence, the sum $\sum_{i_u \in f^{-1}(k_s)} \alpha_u$ contains precisely one addend, namely the one for $u = \sigma(s)$; as a consequence, this sum simplifies to $\sum_{i_u \in f^{-1}(k_s)} \alpha_u = \alpha_{\sigma(s)}$. On the other hand, the sum $\sum_{j_v \in f^{-1}(k_s)} \beta_v$ is empty (since the set $f^{-1}(k_s)$ equals $\left\{i_{\sigma(s)}\right\}$, and thus contains no elements of the form j_v), and thus vanishes, i.e., we have $\sum_{j_v \in f^{-1}(k_s)} \beta_v = 0$. Now,

$$\operatorname{wt}_s(f) = \underbrace{\sum_{i_u \in f^{-1}(k_s)} \alpha_u}_{=\alpha_{\sigma(s)}} + \underbrace{\sum_{j_v \in f^{-1}(k_s)} \beta_v}_{=0} = \alpha_{\sigma(s)} = \gamma_{\sigma(s)}$$

(because $\gamma_{\sigma(s)} = \alpha_{\sigma(s)}$ (since $(\gamma_1, \gamma_2, \ldots, \gamma_{\ell+m}) = (\alpha_1, \alpha_2, \ldots, \alpha_\ell, \beta_1, \beta_2, \ldots, \beta_m)$ and $\sigma(s) \leq \ell$)). In other words, $\gamma_{\sigma(s)} = \operatorname{wt}_s(f)$.

We have thus shown that $\gamma_{\sigma(s)} = \operatorname{wt}_s(f)$ holds in Case 1. A similar argument (but relying on $\mathbf{d}(\sigma(s)) = j_{\sigma(s)-\ell}$ instead of $\mathbf{d}(\sigma(s)) = i_{\sigma(s)}$) shows that $\gamma_{\sigma(s)} = \operatorname{wt}_s(f)$ holds in Case 2.

Thus, $\gamma_{\sigma(s)} = \operatorname{wt}_s(f)$ holds in both Cases 1 and 2. Since these two Cases cover all possibilities, this yields that $\gamma_{\sigma(s)} = \operatorname{wt}_s(f)$ always holds.

Hence, $\gamma_{\sigma(s)} = \operatorname{wt}_s(f)$ for every $s \in \{1, 2, \ldots, \ell+m\}$. Thus, $\left(\gamma_{\sigma(1)}, \gamma_{\sigma(2)}, \ldots, \gamma_{\sigma(\ell+m)}\right) = (\operatorname{wt}_1(f), \operatorname{wt}_2(f), \ldots, \operatorname{wt}_{\ell+m}(f))$. Hence,

$$\alpha \underset{\sigma}{\sqcup\!\sqcup} \beta = \left(\gamma_{\sigma(1)}, \gamma_{\sigma(2)}, \ldots, \gamma_{\sigma(\ell+m)}\right) = (\operatorname{wt}_1(f), \operatorname{wt}_2(f), \ldots, \operatorname{wt}_{\ell+m}(f)) = \operatorname{wt}\left(\underbrace{f}_{=\Phi\left(\sigma^{-1}\right)}\right) = \operatorname{wt}\left(\Phi\left(\sigma^{-1}\right)\right).$$

This proves (13.168.9).

13.169. **Solution to Exercise 6.4.8.** *Solution to Exercise 6.4.8.*

Proof of Corollary 6.4.7. Write α and β as $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_\ell)$ and $\beta = (\beta_1, \beta_2, \ldots, \beta_m)$, respectively; then $\ell(\alpha) = \ell$ and $\ell(\beta) = m$.

Fix three disjoint chain posets $(i_1 < i_2 < \cdots < i_\ell)$, $(j_1 < j_2 < \cdots < j_m)$ and $(k_1 < k_2 < k_3 < \cdots)$. We have

$$\begin{aligned}&\sum_{\substack{p\in\mathbb{N};\\ p\leq \ell+m}} \sum_{f \text{ is a } p\text{-shuffling map}} M_{\operatorname{wt}(f)}\\ &= (\text{a sum of terms of the form } M_\delta \text{ with } \delta \in \mathfrak{A}^* \text{ satisfying } \ell(\delta) \leq \ell(\alpha) + \ell(\beta)) \qquad (13.169.1)\end{aligned}$$

[1141]. Now, (13.168.1) becomes

$$\begin{aligned}M_\alpha M_\beta &= \sum_{\substack{p\in\mathbb{N};\\ p\leq \ell+m}} \sum_{f \text{ is a } p\text{-shuffling map}} M_{\operatorname{wt}(f)}\\ &= (\text{a sum of terms of the form } M_\delta \text{ with } \delta \in \mathfrak{A}^* \text{ satisfying } \ell(\delta) \leq \ell(\alpha) + \ell(\beta))\end{aligned}$$

(by (13.169.1)). This proves Corollary 6.4.7. □

13.170. **Solution to Exercise 6.4.12.** *Solution to Exercise 6.4.12.*

Proof of Lemma 6.4.11. (a) Let $\gamma \in u ⧢ v$ be arbitrary. Then, the multiset of all letters of γ is the disjoint union of the multiset of all letters of u with the multiset of all letters of v. Hence, the sum of all letters of γ equals the sum of all letters of u plus the sum of all letters of v. In other words, $|\gamma| = |u| + |v|$ (because $|\gamma|$ is the sum of all letters of γ, because $|u|$ is the sum of all letters of u, and because $|v|$ is the sum of all letters of v). Since $u \in \operatorname{Comp}_n$, we have $|u| = n$. Similarly, $|v| = m$. Thus, $|\gamma| = \underbrace{|u|}_{=n} + \underbrace{|v|}_{=m} = n + m$, so that $\gamma \in \operatorname{Comp}_{n+m}$.

Now, let us forget that we fixed γ. We thus have proven that

$$\gamma \in \operatorname{Comp}_{n+m} \text{ for every } \gamma \in u ⧢ v. \qquad (13.170.1)$$

Applying (13.170.1) to $\gamma = z$, we obtain $z \in \operatorname{Comp}_{n+m}$. This proves Lemma 6.4.11(a).

(b) Since $z \in u ⧢ v$, we have $\ell(z) = \ell(u) + \ell(v)$.

We have $u \in \operatorname{Comp}_n \subset \operatorname{Comp} = \mathfrak{A}^*$ and similarly $v \in \mathfrak{A}^*$. Thus, Proposition 6.4.5 (applied to $\alpha = u$ and $\beta = v$) yields

$$\begin{aligned}&M_u M_v\\ &= \sum_{\gamma \in u ⧢ v} M_\gamma + (\text{a sum of terms of the form } M_\delta \text{ with } \delta \in \mathfrak{A}^* \text{ satisfying } \ell(\delta) < \ell(u) + \ell(v)),\end{aligned}$$

so that

$$\begin{aligned}&M_u M_v - \sum_{\gamma \in u ⧢ v} M_\gamma\\ &= \left(\text{a sum of terms of the form } M_\delta \text{ with } \delta \in \underbrace{\mathfrak{A}^*}_{=\operatorname{Comp}} \text{ satisfying } \ell(\delta) < \underbrace{\ell(u) + \ell(v)}_{=\ell(z)}\right)\\ &= (\text{a sum of terms of the form } M_\delta \text{ with } \delta \in \operatorname{Comp} \text{ satisfying } \ell(\delta) < \ell(z)). \qquad (13.170.2)\end{aligned}$$

[1141] *Proof of (13.169.1):* The proof of this is analogous to the proof of (13.168.3), with the only difference that some $<$ signs are replaced by $\leq$ signs.

Now, let π denote the projection from the direct sum $\operatorname{QSym} = \bigoplus_{k\in\mathbb{N}} \operatorname{QSym}_k$ onto its $(n+m)$-th homogeneous component $\operatorname{QSym}_{n+m}$. Notice that $\underbrace{M_u}_{\substack{\in \operatorname{QSym}_n \\ (\text{since } u \in \operatorname{Comp}_n)}} \underbrace{M_v}_{\substack{\in \operatorname{QSym}_m \\ (\text{since } v \in \operatorname{Comp}_m)}} \in \operatorname{QSym}_n \cdot \operatorname{QSym}_m \subset \operatorname{QSym}_{n+m}$, so that $\pi\left(M_u M_v\right) = M_u M_v$. Also, for every $\gamma \in u \sqcup v$, we have $M_\gamma \in \operatorname{QSym}_{n+m}$ (because (13.170.1) shows that $\gamma \in \operatorname{Comp}_{n+m}$) and therefore $\pi\left(M_\gamma\right) = M_\gamma$. Hence, the $\mathbf{k}$-linearity of π yields

$$\begin{aligned}
\pi\left(M_u M_v - \sum_{\gamma \in u \sqcup v} M_\gamma\right) &= \underbrace{\pi\left(M_u M_v\right)}_{= M_u M_v} - \sum_{\gamma \in u \sqcup v} \underbrace{\pi\left(M_\gamma\right)}_{= M_\gamma} \\
&= M_u M_v - \sum_{\gamma \in u \sqcup v} M_\gamma . \qquad (13.170.3)
\end{aligned}$$

On the other hand,

$$\pi\left(M_\delta\right) = 0 \qquad \text{for every } \delta \in \operatorname{Comp} \setminus \operatorname{Comp}_{n+m} \qquad (13.170.4)$$

[1142].

But applying the projection π to the equality (13.170.2), we obtain

$$\begin{aligned}
&\pi\left(M_u M_v - \sum_{\gamma \in u \sqcup v} M_\gamma\right) \\
&= \pi\left(\text{a sum of terms of the form } M_\delta \text{ with } \delta \in \operatorname{Comp} \text{ satisfying } \ell\left(\delta\right) < \ell\left(z\right)\right) \\
&= \left(\text{a sum of terms of the form } \pi\left(M_\delta\right) \text{ with } \delta \in \operatorname{Comp} \text{ satisfying } \ell\left(\delta\right) < \ell\left(z\right)\right) \\
&= \left(\text{a sum of terms of the form } \underbrace{\pi\left(M_\delta\right)}_{\substack{= M_\delta \\ (\text{since } M_\delta \in \operatorname{QSym}_{n+m} \\ (\text{since } \delta \in \operatorname{Comp}_{n+m}))}} \text{ with } \delta \in \operatorname{Comp}_{n+m} \text{ satisfying } \ell\left(\delta\right) < \ell\left(z\right)\right) \\
&\qquad + \left(\text{a sum of terms of the form } \underbrace{\pi\left(M_\delta\right)}_{\substack{= 0 \\ (\text{by } (13.170.4))}} \text{ with } \delta \in \operatorname{Comp} \setminus \operatorname{Comp}_{n+m} \text{ satisfying } \ell\left(\delta\right) < \ell\left(z\right)\right) \\
&\qquad \left(\text{since every } \delta \in \operatorname{Comp} \text{ satisfies either } \delta \in \operatorname{Comp}_{n+m} \text{ or } \delta \in \operatorname{Comp} \setminus \operatorname{Comp}_{n+m}\right) \\
&= \left(\text{a sum of terms of the form } M_\delta \text{ with } \delta \in \operatorname{Comp}_{n+m} \text{ satisfying } \ell\left(\delta\right) < \ell\left(z\right)\right) \\
&\qquad + \underbrace{\left(\text{a sum of terms of the form } 0 \text{ with } \delta \in \operatorname{Comp} \setminus \operatorname{Comp}_{n+m} \text{ satisfying } \ell\left(\delta\right) < \ell\left(z\right)\right)}_{=0} \\
&= \left(\text{a sum of terms of the form } M_\delta \text{ with } \delta \in \operatorname{Comp}_{n+m} \text{ satisfying } \ell\left(\delta\right) < \ell\left(z\right)\right) \\
&= \left(\text{a sum of terms of the form } M_w \text{ with } w \in \operatorname{Comp}_{n+m} \text{ satisfying } \ell\left(w\right) < \ell\left(z\right)\right) \\
&\qquad \left(\text{here, we renamed the index } \delta \text{ as } w\right) \\
&= \left(\text{a sum of terms of the form } M_w \text{ with } w \in \operatorname{Comp}_{n+m} \text{ satisfying } w \underset{\text{wll}}{<} z\right)
\end{aligned}$$

[1142] *Proof of (13.170.4):* Let $\delta \in \operatorname{Comp} \setminus \operatorname{Comp}_{n+m}$. Then, $\delta \in \operatorname{Comp}$ but $\delta \notin \operatorname{Comp}_{n+m}$. In other words, δ is a composition with size $\left|\delta\right| \neq n+m$. As a consequence, M_δ is a homogeneous element of QSym of degree $\left|\delta\right| \neq n+m$. Therefore, $\pi\left(M_\delta\right) = 0$ (since π is the projection from the direct sum $\operatorname{QSym} = \bigoplus_{k\in\mathbb{N}} \operatorname{QSym}_k$ onto its $(n+m)$-th homogeneous component $\operatorname{QSym}_{n+m}$). This proves (13.170.4).

(since every $w \in \operatorname{Comp}_{n+m}$ satisfying $\ell(w) < \ell(z)$ must satisfy $w \underset{\text{wll}}{<} z$). Compared with (13.170.3), this yields

$$M_u M_v - \sum_{\gamma \in u \sqcup v} M_\gamma$$
$$= \left(\text{a sum of terms of the form } M_w \text{ with } w \in \operatorname{Comp}_{n+m} \text{ satisfying } w \underset{\text{wll}}{<} z\right).$$

Adding $\sum_{\gamma \in u \sqcup v} M_\gamma$ to both sides of this equality, we obtain

$$M_u M_v = \sum_{\gamma \in u \sqcup v} M_\gamma + \left(\text{a sum of terms of the form } M_w \text{ with } w \in \operatorname{Comp}_{n+m} \text{ satisfying } w \underset{\text{wll}}{<} z\right). \tag{13.170.5}$$

But every $\gamma \in u \sqcup v$ satisfying $\gamma \neq z$ must satisfy $\gamma \in \operatorname{Comp}_{n+m}$ (by (13.170.1)) and $\gamma \underset{\text{wll}}{<} z$ [1143]. Hence,

$$\sum_{\substack{\gamma \in u \sqcup v;\\ \gamma \neq z}} M_\gamma = \left(\text{a sum of terms of the form } M_\gamma \text{ with } \gamma \in \operatorname{Comp}_{n+m} \text{ satisfying } \gamma \underset{\text{wll}}{<} z\right). \tag{13.170.6}$$

But let h be the multiplicity with which the word z appears in the multiset $u \sqcup v$. Then, h is a positive integer (since z is an element of the multiset $u \sqcup v$), and satisfies $\sum_{\substack{\gamma \in u \sqcup v;\\ \gamma = z}} M_z = h M_z$. Now,

$$\sum_{\gamma \in u \sqcup v} M_\gamma = \sum_{\substack{\gamma \in u \sqcup v;\\ \gamma = z}} \underbrace{M_\gamma}_{\substack{= M_z \\ (\text{since } \gamma = z)}} + \sum_{\substack{\gamma \in u \sqcup v;\\ \gamma \neq z}} M_\gamma$$
$$= \underbrace{\sum_{\substack{\gamma \in u \sqcup v;\\ \gamma = z}} M_z}_{= h M_z} + \underbrace{\sum_{\substack{\gamma \in u \sqcup v;\\ \gamma \neq z}} M_\gamma}_{\substack{= \left(\text{a sum of terms of the form } M_\gamma \text{ with } \gamma \in \operatorname{Comp}_{n+m} \text{ satisfying } \gamma \underset{\text{wll}}{<} z\right) \\ (\text{by (13.170.6)})}}$$
$$= h M_z + \underbrace{\left(\text{a sum of terms of the form } M_\gamma \text{ with } \gamma \in \operatorname{Comp}_{n+m} \text{ satisfying } \gamma \underset{\text{wll}}{<} z\right)}_{\substack{= \left(\text{a sum of terms of the form } M_w \text{ with } w \in \operatorname{Comp}_{n+m} \text{ satisfying } w \underset{\text{wll}}{<} z\right) \\ (\text{here, we have renamed the index } \gamma \text{ as } w)}}$$
$$= h M_z + \left(\text{a sum of terms of the form } M_w \text{ with } w \in \operatorname{Comp}_{n+m} \text{ satisfying } w \underset{\text{wll}}{<} z\right).$$

[1143] *Proof.* Let $\gamma \in u \sqcup v$ be such that $\gamma \neq z$. Since $\gamma \in u \sqcup v$, we must have $\ell(\gamma) = \ell(u) + \ell(v) = \ell(z)$. Also, γ and z both belong to $\operatorname{Comp}_{n+m}$. Now, γ is an element of the multiset $u \sqcup v$, whereas z is the lexicographically highest element of this multiset. Hence, $\gamma \leq z$ with respect to the lexicographic order. Since the elements γ and z of $\operatorname{Comp}_{n+m}$ satisfy $\ell(\gamma) = \ell(z)$ and $\gamma \leq z$ with respect to the lexicographic order, we must have $\gamma \underset{\text{wll}}{\leq} z$, and thus $\gamma \underset{\text{wll}}{<} z$ (since $\gamma \neq z$), qed.

Hence, (13.170.5) becomes

$$\begin{aligned}
&M_u M_v \\
&= \underbrace{\sum_{\gamma \in u \shuffle v} M_\gamma}_{= hM_z + \left(\text{a sum of terms of the form } M_w \text{ with } w \in \operatorname{Comp}_{n+m} \text{ satisfying } w \underset{\text{wll}}{<} z\right)} \\
&\qquad + \left(\text{a sum of terms of the form } M_w \text{ with } w \in \operatorname{Comp}_{n+m} \text{ satisfying } w \underset{\text{wll}}{<} z\right) \\
&= hM_z + \left(\text{a sum of terms of the form } M_w \text{ with } w \in \operatorname{Comp}_{n+m} \text{ satisfying } w \underset{\text{wll}}{<} z\right) \\
&\qquad + \left(\text{a sum of terms of the form } M_w \text{ with } w \in \operatorname{Comp}_{n+m} \text{ satisfying } w \underset{\text{wll}}{<} z\right) \\
&= hM_z + \left(\text{a sum of terms of the form } M_w \text{ with } w \in \operatorname{Comp}_{n+m} \text{ satisfying } w \underset{\text{wll}}{<} z\right).
\end{aligned}$$

This proves Lemma 6.4.11(b).

(c) Since $z \in u \shuffle v$, we have $\ell(z) = \ell(u) + \ell(v)$. Since $v' \underset{\text{wll}}{<} v$, we have $v' \neq v$.

Let z' be the lexicographically highest element of the multiset $u \shuffle v'$. Lemma 6.4.11(a) (applied to v' and z' instead of v and z) yields $z' \in \operatorname{Comp}_{n+m}$. Since z' is an element of the multiset $u \shuffle v'$, we have $\ell(z') = \ell(u) + \ell(v')$.

Now, it is easy to see (using Lemma 6.3.10) that $z' \underset{\text{wll}}{<} z$ [1144]. Hence, every $w \in \operatorname{Comp}_{n+m}$ satisfying $w \underset{\text{wll}}{<} z'$ also satisfies $w \underset{\text{wll}}{<} z$ (because it satisfies $w \underset{\text{wll}}{<} z' \underset{\text{wll}}{<} z$). Applying Lemma 6.4.11(b) to v' and z' instead of v and z, we conclude that there exists a positive integer h such that

$$M_u M_{v'} = hM_{z'} + \left(\text{a sum of terms of the form } M_w \text{ with } w \in \operatorname{Comp}_{n+m} \text{ satisfying } w \underset{\text{wll}}{<} z'\right).$$

[1144] *Proof.* We have $v' \underset{\text{wll}}{<} v$, thus $v' \underset{\text{wll}}{\leq} v$. According to the definition of the relation $\underset{\text{wll}}{\leq}$, this means that we are in one of the following two cases:

Case 1: We have $\ell(v') < \ell(v)$.

Case 2: We have $\ell(v') = \ell(v)$ and $v' \leq v$ in lexicographic order.

Let us first consider Case 1. In this case, we have $\ell(v') < \ell(v)$. Now, $\ell(z') = \ell(u) + \underbrace{\ell(v')}_{<\ell(v)} < \ell(u) + \ell(v) = \ell(z)$. But any two elements α and β of $\operatorname{Comp}_{n+m}$ satisfying $\ell(\alpha) < \ell(\beta)$ must satisfy $\alpha \underset{\text{wll}}{\leq} \beta$ (by the definition of $\underset{\text{wll}}{\leq}$). Applying this to $\alpha = z'$ and $\beta = z$, we obtain $z' \underset{\text{wll}}{\leq} z$. Combined with $z' \neq z$ (since $\ell(z') < \ell(z)$), this yields $z' \underset{\text{wll}}{<} z$. Thus, $z' \underset{\text{wll}}{<} z$ is proven in Case 1.

Let us now consider Case 2. In this case, we have $\ell(v') = \ell(v)$ and $v' \leq v$ in lexicographic order. Now, $\ell(z') = \ell(u) + \underbrace{\ell(v')}_{=\ell(v)} = \ell(u) + \ell(v) = \ell(z)$.

We know that z' is an element of the multiset $u \shuffle v'$; in other words, z' can be written in the form $z' = u \underset{\sigma}{\shuffle} v'$ for some $\sigma \in \operatorname{Sh}_{\ell(u),\ell(v')}$ (because every element of the multiset $u \shuffle v'$ can be written in this form). Consider this σ. We have $\sigma \in \operatorname{Sh}_{\ell(u),\ell(v')} = \operatorname{Sh}_{\ell(u),\ell(v)}$ (since $\ell(v') = \ell(v)$), and thus $u \underset{\sigma}{\shuffle} v$ is a well-defined element of the multiset $u \shuffle v$. Therefore, $u \underset{\sigma}{\shuffle} v \leq z$ (because z is the lexicographically highest element of this multiset $u \shuffle v$).

But we have $v' \leq v$ with respect to the relation $\leq$ on $\mathfrak{A}^*$ defined in Definition 6.1.1 (since $v' \leq v$ in lexicographic order). Thus, $v' < v$ with respect to this relation (since $v' \neq v$). Hence, Lemma 6.3.10(a) (applied to $\ell(u)$ and $\ell(v)$ instead of n and m) yields $u \underset{\sigma}{\shuffle} v' < u \underset{\sigma}{\shuffle} v \leq z$. Thus, $z' = u \underset{\sigma}{\shuffle} v' < z$ with respect to the relation $\leq$ on $\mathfrak{A}^*$. In other words, $z' < z$ in lexicographic order (since $\ell(z') = \ell(z)$), thus $z' \leq z$.

Now, the two elements z' and z of $\operatorname{Comp}_{n+m}$ satisfy $\ell(z') = \ell(z)$ and $z' \leq z$ in lexicographic order. But any two elements α and β of $\operatorname{Comp}_{n+m}$ satisfying ($\ell(\alpha) = \ell(\beta)$ and $\alpha \leq \beta$ in lexicographic order) must satisfy $\alpha \underset{\text{wll}}{\leq} \beta$ (by the definition of $\underset{\text{wll}}{\leq}$). Applying this to $\alpha = z'$ and $\beta = z$, we obtain $z' \underset{\text{wll}}{\leq} z$. Since $z' \neq z$, this yields $z' \underset{\text{wll}}{<} z$. Hence, $z' \underset{\text{wll}}{<} z$ is proven in Case 2.

Now, we have proved $z' \underset{\text{wll}}{<} z$ in each of the two Cases 1 and 2. Since these two Cases cover all possibilities, this yields that $z' \underset{\text{wll}}{<} z$ always holds, qed.

Consider this h. We have

$$\begin{aligned}
M_u M_{v'} &= \underbrace{h M_{z'}}_{\substack{= \sum_{i=1}^{h} M_{z'} = \left(\text{a sum of terms of the form } M_w \text{ with } w \in \operatorname{Comp}_{n+m} \text{ satisfying } w \underset{\text{wll}}{<} z\right) \\ \left(\text{since } z' \in \operatorname{Comp}_{n+m} \text{ satisfies } z' \underset{\text{wll}}{<} z\right)}} \\
&\qquad + \underbrace{\left(\text{a sum of terms of the form } M_w \text{ with } w \in \operatorname{Comp}_{n+m} \text{ satisfying } w \underset{\text{wll}}{<} z'\right)}_{\substack{= \left(\text{a sum of terms of the form } M_w \text{ with } w \in \operatorname{Comp}_{n+m} \text{ satisfying } w \underset{\text{wll}}{<} z\right) \\ \left(\text{since every } w \in \operatorname{Comp}_{n+m} \text{ satisfying } w \underset{\text{wll}}{<} z' \text{ also satisfies } w \underset{\text{wll}}{<} z\right)}} \\
&= \left(\text{a sum of terms of the form } M_w \text{ with } w \in \operatorname{Comp}_{n+m} \text{ satisfying } w \underset{\text{wll}}{<} z\right) \\
&\qquad + \left(\text{a sum of terms of the form } M_w \text{ with } w \in \operatorname{Comp}_{n+m} \text{ satisfying } w \underset{\text{wll}}{<} z\right) \\
&= \left(\text{a sum of terms of the form } M_w \text{ with } w \in \operatorname{Comp}_{n+m} \text{ satisfying } w \underset{\text{wll}}{<} z\right).
\end{aligned}$$

This proves Lemma 6.4.11(c). □

13.171. **Solution to Exercise 6.4.13.** *Solution to Exercise 6.4.13.*

Proof of Proposition 6.4.10. We shall prove Proposition 6.4.10 by strong induction over n. So we fix some $N \in \mathbb{N}$, and we assume that Proposition 6.4.10 holds whenever $n < N$. We now need to prove that Proposition 6.4.10 holds for $n = N$. In other words, we need to prove that for every $x \in \operatorname{Comp}_N$, there is a family $(\eta_{x,y})_{y \in \operatorname{Comp}_N} \in \mathbb{N}^{\operatorname{Comp}_N}$ of elements of $\mathbb{N}$ satisfying

$$\mathbf{M}_x = \sum_{\substack{y \in \operatorname{Comp}_N; \\ y \underset{\text{wll}}{\leq} x}} \eta_{x,y} M_y \tag{13.171.1}$$

and $\eta_{x,x} \neq 0$ (in $\mathbb{N}$).

Let $x \in \operatorname{Comp}_N$ be arbitrary. We need to prove that there is a family $(\eta_{x,y})_{y \in \operatorname{Comp}_N} \in \mathbb{N}^{\operatorname{Comp}_N}$ of elements of $\mathbb{N}$ satisfying (13.171.1) and $\eta_{x,x} \neq 0$.

Let $(a_1, a_2, \ldots, a_p)$ be the CFL factorization of x. Then, a_1, a_2, $\ldots$, a_p are Lyndon words satisfying $x = a_1 a_2 \cdots a_p$ and $a_1 \geq a_2 \geq \cdots \geq a_p$.

If x is the empty word, then our claim is trivial (in fact, we can just set $\eta_{x,x} = 1$ in this case, and (13.171.1) holds obviously). Hence, for the rest of this proof, we WLOG assume that x is not the empty word. Thus, $p \neq 0$ (because otherwise, $x = a_1 a_2 \cdots a_p$ would be an empty product and thus the empty word, contradicting the assumption that x is not the empty word). Hence, we can define two words $u \in \mathfrak{A}^*$ and $v \in \mathfrak{A}^*$ by $u = a_1$ and $v = a_2 a_3 \cdots a_p$. The word u is Lyndon (since $u = a_1$ and since a_1 is Lyndon), and thus has (u) as its CFL factorization. The CFL factorization of the word v is $(a_2, a_3, \ldots, a_p)$ (since the words a_2, a_3, $\ldots$, a_p are Lyndon and satisfy $v = a_2 a_3 \cdots a_p$ and $a_2 \geq a_3 \geq \cdots \geq a_p$ (because $a_1 \geq a_2 \geq \cdots \geq a_p$)). Also, $u = a_1 \geq a_{j+1}$ for every $i \in \{1, 2, \ldots, 1\}$ and $j \in \{1, 2, \ldots, p-1\}$ (because $a_1 \geq a_2 \geq \cdots \geq a_p$). Hence, we can apply Theorem 6.2.2(c) to 1, (u), $p-1$ and $(a_2, a_3, \ldots, a_p)$ instead of p, $(a_1, a_2, \ldots, a_p)$, q and $(b_1, b_2, \ldots, b_q)$. As a result, we conclude that the lexicographically highest element of the multiset $u \sqcup v$ is $\underbrace{u}_{=a_1} \underbrace{v}_{=a_2 a_3 \cdots a_p} = a_1 (a_2 a_3 \cdots a_p) = a_1 a_2 \cdots a_p = x$.

But $\mathbf{M}_u = M_u$ (by the definition of $\mathbf{M}_u$, since u has CFL factorization (u)), and $\mathbf{M}_x = \mathbf{M}_u \mathbf{M}_v$ [1145].

The word u is Lyndon and thus nonempty, so that $|u| > 0$. Now, $\left| \underbrace{x}_{=uv} \right| = |uv| = |u| + |v|$, so that $|u| + |v| = |x| = N$ (since $x \in \operatorname{Comp}_N$). Thus, $N = \underbrace{|u|}_{>0} + |v| > |v|$, so that $|v| < N$. Hence, the induction hypothesis tells us that we can apply Proposition 6.4.10 to $|v|$ and v instead of n and x (since $v \in \operatorname{Comp}_{|v|}$). As a result, we see that there is a family $(\eta_{v,y})_{y \in \operatorname{Comp}_{|v|}} \in \mathbb{N}^{\operatorname{Comp}_{|v|}}$ of elements of $\mathbb{N}$ satisfying

$$\mathbf{M}_v = \sum_{\substack{y \in \operatorname{Comp}_{|v|}; \\ y \underset{\mathrm{wll}}{\leq} v}} \eta_{v,y} M_y$$

and $\eta_{v,v} \neq 0$ (in $\mathbb{N}$). Consider this family $(\eta_{v,y})_{y \in \operatorname{Comp}_{|v|}}$.

We have $u \in \operatorname{Comp}_{|u|}$ and $v \in \operatorname{Comp}_{|v|}$. Hence, Lemma 6.4.11(b) (applied to x, $|u|$ and $|v|$ instead of z, u and v) yields that there exists a positive integer h such that

$$\begin{aligned} M_u M_v &= h M_x + \left(\text{a sum of terms of the form } M_w \text{ with } w \in \underbrace{\operatorname{Comp}_{|u|+|v|}}_{\substack{=\operatorname{Comp}_N \\ (\text{since } |u|+|v|=N)}} \text{ satisfying } w \underset{\mathrm{wll}}{<} x \right) \\ &= h M_x + \left(\text{a sum of terms of the form } M_w \text{ with } w \in \operatorname{Comp}_N \text{ satisfying } w \underset{\mathrm{wll}}{<} x \right). \qquad (13.171.2) \end{aligned}$$

Consider this h. Then, $h \in \mathbb{N}$ and $h \neq 0$ (since h is a positive integer).

[1145] *Proof.* Since the CFL factorization of v is $(a_2, a_3, \ldots, a_p)$, we have $\mathbf{M}_v = M_{a_2} M_{a_3} \cdots M_{a_p}$ (by the definition of $\mathbf{M}_v$). But since the CFL factorization of x is $(a_1, a_2, \ldots, a_p)$, we have

$$\mathbf{M}_x = M_{a_1} M_{a_2} \cdots M_{a_p} = \underbrace{M_{a_1}}_{\substack{=M_u \\ (\text{since } a_1 = u)}} \underbrace{\left(M_{a_2} M_{a_3} \cdots M_{a_p} \right)}_{=\mathbf{M}_v} = \underbrace{M_u}_{\substack{=\mathbf{M}_u \\ (\text{since } \mathbf{M}_u = M_u)}} \mathbf{M}_v = \mathbf{M}_u \mathbf{M}_v,$$

qed.

Now,

$$\begin{aligned}
\mathbf{M}_x &= \underbrace{\mathbf{M}_u}_{=M_u} \underbrace{\mathbf{M}_v}_{=\sum\limits_{\substack{y\in\operatorname{Comp}_{|v|};\\ y \underset{\mathrm{wll}}{\leq} v}} \eta_{v,y} M_y} = M_u \left(\sum_{\substack{y\in\operatorname{Comp}_{|v|};\\ y \underset{\mathrm{wll}}{\leq} v}} \eta_{v,y} M_y \right) = \sum_{\substack{y\in\operatorname{Comp}_{|v|};\\ y \underset{\mathrm{wll}}{\leq} v}} \eta_{v,y} M_u M_y \\
&= \eta_{v,v} \underbrace{M_u M_v}_{\substack{=hM_x + \left(\text{a sum of terms of the form } M_w \text{ with } w\in\operatorname{Comp}_N \text{ satisfying } w \underset{\mathrm{wll}}{<} x\right) \\ \text{(by (13.171.2))}}} \\
&\qquad + \sum_{\substack{y\in\operatorname{Comp}_{|v|};\\ y \underset{\mathrm{wll}}{<} v}} \eta_{v,y} \underbrace{M_u M_y}_{\substack{=\left(\text{a sum of terms of the form } M_w \text{ with } w\in\operatorname{Comp}_{|u|+|v|} \text{ satisfying } w \underset{\mathrm{wll}}{<} x\right) \\ \text{(by Lemma 6.4.11(c), applied to } n=|u|,\ m=|v|,\ v'=y \text{ and } z=x\text{)}}} \\
&\qquad\qquad \text{(here, we have split off the addend for } y = v \text{ from the sum)} \\
&= \underbrace{\eta_{v,v} \left(hM_x + \left(\text{a sum of terms of the form } M_w \text{ with } w \in \operatorname{Comp}_N \text{ satisfying } w \underset{\mathrm{wll}}{<} x \right) \right)}_{=\eta_{v,v} hM_x + \eta_{v,v} \left(\text{a sum of terms of the form } M_w \text{ with } w\in\operatorname{Comp}_N \text{ satisfying } w \underset{\mathrm{wll}}{<} x\right)} \\
&\qquad + \sum_{\substack{y\in\operatorname{Comp}_{|v|};\\ y \underset{\mathrm{wll}}{<} v}} \eta_{v,y} \left(\text{a sum of terms of the form } M_w \text{ with } w \in \underbrace{\operatorname{Comp}_{|u|+|v|}}_{\substack{=\operatorname{Comp}_N \\ \text{(since } |u|+|v|=N\text{)}}} \text{ satisfying } w \underset{\mathrm{wll}}{<} x \right) \\
&= \eta_{v,v} hM_x + \eta_{v,v} \left(\text{a sum of terms of the form } M_w \text{ with } w \in \operatorname{Comp}_N \text{ satisfying } w \underset{\mathrm{wll}}{<} x \right) \\
&\qquad \underbrace{+ \sum_{\substack{y\in\operatorname{Comp}_{|v|};\\ y \underset{\mathrm{wll}}{<} v}} \eta_{v,y} \left(\text{a sum of terms of the form } M_w \text{ with } w \in \operatorname{Comp}_N \text{ satisfying } w \underset{\mathrm{wll}}{<} x \right)}_{\substack{=\left(\text{an } \mathbb{N}\text{-linear combination of terms of the form } M_w \text{ with } w\in\operatorname{Comp}_N \text{ satisfying } w \underset{\mathrm{wll}}{<} x\right) \\ \text{(since } \eta_{v,y}\in\mathbb{N} \text{ for every } y\in\operatorname{Comp}_{|v|} \text{ satisfying } y \underset{\mathrm{wll}}{<} v\text{)}}} \\
&= \eta_{v,v} hM_x + \underbrace{\eta_{v,v} \left(\text{a sum of terms of the form } M_w \text{ with } w \in \operatorname{Comp}_N \text{ satisfying } w \underset{\mathrm{wll}}{<} x \right)}_{\substack{=\left(\text{an } \mathbb{N}\text{-linear combination of terms of the form } M_w \text{ with } w\in\operatorname{Comp}_N \text{ satisfying } w \underset{\mathrm{wll}}{<} x\right) \\ \text{(since } \eta_{v,v}\in\mathbb{N}\text{)}}} \\
&\qquad + \left(\text{an } \mathbb{N}\text{-linear combination of terms of the form } M_w \text{ with } w \in \operatorname{Comp}_N \text{ satisfying } w \underset{\mathrm{wll}}{<} x \right) \\
&= \eta_{v,v} hM_x + \left(\text{an } \mathbb{N}\text{-linear combination of terms of the form } M_w \text{ with } w \in \operatorname{Comp}_N \text{ satisfying } w \underset{\mathrm{wll}}{<} x \right) \\
&\qquad + \left(\text{an } \mathbb{N}\text{-linear combination of terms of the form } M_w \text{ with } w \in \operatorname{Comp}_N \text{ satisfying } w \underset{\mathrm{wll}}{<} x \right) \\
&= \eta_{v,v} hM_x + \left(\text{an } \mathbb{N}\text{-linear combination of terms of the form } M_w \text{ with } w \in \operatorname{Comp}_N \text{ satisfying } w \underset{\mathrm{wll}}{<} x \right).
\end{aligned}$$

In other words, we can write $\mathbf{M}_x$ in the form

$$\mathbf{M}_x = \eta_{v,v} hM_x + \mathbf{c}, \tag{13.171.3}$$

where $\mathbf{c}$ is an $\mathbb{N}$-linear combination of terms of the form M_w with $w \in \operatorname{Comp}_N$ satisfying $w \underset{\text{wll}}{<} x$. Consider this $\mathbf{c}$.

Write $\mathbf{c}$ in the form

$$\mathbf{c} = \sum_{\substack{y \in \operatorname{Comp}_N; \\ y \underset{\text{wll}}{<} x}} \eta_{x,y} M_y, \tag{13.171.4}$$

where $\eta_{x,y}$ is an element of $\mathbb{N}$ for every $y \in \operatorname{Comp}_N$ satisfying $y \underset{\text{wll}}{<} x$. (This is possible since $\mathbf{c}$ is an $\mathbb{N}$-linear combination of terms of the form M_w with $w \in \operatorname{Comp}_N$ satisfying $w \underset{\text{wll}}{<} x$.) Thus, we have defined a family $(\eta_{x,y})_{y \in \operatorname{Comp}_N;\ y \underset{\text{wll}}{<} x}$ of elements of $\mathbb{N}$. Extend this family to a family $(\eta_{x,y})_{y \in \operatorname{Comp}_N} \in \mathbb{N}^{\operatorname{Comp}_N}$ of elements of $\mathbb{N}$ by defining

$$\eta_{x,y} = \delta_{x,y} \eta_{v,v} h \qquad \text{for every } y \in \operatorname{Comp}_N \text{ which does not satisfy } y \underset{\text{wll}}{<} x.$$

Notice that the definition of $\eta_{x,x}$ yields $\eta_{x,x} = \delta_{x,x} \eta_{v,v} h$ (since x does not satisfy $x \underset{\text{wll}}{<} x$), and thus $\eta_{x,x} = \underbrace{\delta_{x,x}}_{=1} \eta_{v,v} h = \underbrace{\eta_{v,v}}_{\neq 0} \underbrace{h}_{\neq 0} \neq 0$.

Now,

$$\begin{aligned} \sum_{\substack{y \in \operatorname{Comp}_N; \\ y \underset{\text{wll}}{\leq} x}} \eta_{x,y} M_y &= \underbrace{\eta_{x,x}}_{= \eta_{v,v} h} M_x + \underbrace{\sum_{\substack{y \in \operatorname{Comp}_N; \\ y \underset{\text{wll}}{<} x}} \eta_{x,y} M_y}_{\substack{= \mathbf{c} \\ \text{(by (13.171.4))}}} \\ &\qquad \text{(here, we have split off the addend for } y = x \text{ from the sum)} \\ &= \eta_{v,v} h M_x + \mathbf{c} = \mathbf{M}_x \qquad \text{(by (13.171.3))}, \end{aligned}$$

and thus $\mathbf{M}_x = \sum_{\substack{y \in \operatorname{Comp}_N; \\ y \underset{\text{wll}}{\leq} x}} \eta_{x,y} M_y$. In other words, (13.171.1) is satisfied.

Hence, we have constructed a family $(\eta_{x,y})_{y \in \operatorname{Comp}_N} \in \mathbb{N}^{\operatorname{Comp}_N}$ of elements of $\mathbb{N}$ satisfying (13.171.1) and $\eta_{x,x} \neq 0$. Thus, we have shown the existence of such a family. The induction step is thus complete, and Proposition 6.4.10 is proven by induction. □

13.172. **Solution to Exercise 6.4.15.** *Solution to Exercise 6.4.15.*

Proof of Proposition 6.4.14. We know that $\mathbb{Q}$ is a subring of $\mathbf{k}$. Hence, every nonzero element of $\mathbb{N}$ is an invertible element of $\mathbf{k}$.

For every composition $u \in \operatorname{Comp} = \mathfrak{A}^*$, define an element $\mathbf{M}_u \in \operatorname{QSym}$ by $\mathbf{M}_u = M_{a_1} M_{a_2} \cdots M_{a_p}$, where $(a_1, a_2, \ldots, a_p)$ is the CFL factorization of the word u.

Fix $n \in \mathbb{N}$. Consider the set Comp_n as a poset whose smaller relation is the relation $\underset{\text{wll}}{\leq}$. We shall use the notations introduced in Section 11.1.

It is easy to see that $(M_u)_{u \in \operatorname{Comp}_n}$ is a basis of the $\mathbf{k}$-module QSym_n. According to Proposition 6.4.10, the family $(\mathbf{M}_u)_{u \in \operatorname{Comp}_n}$ expands invertibly triangularly in the basis $(M_u)_{u \in \operatorname{Comp}_n}$ [1146]. Hence, Corollary 11.1.19(e) (applied to QSym_n, Comp_n, $(\mathbf{M}_u)_{u \in \operatorname{Comp}_n}$ and $(M_u)_{u \in \operatorname{Comp}_n}$ instead of M, S, $(e_s)_{s \in S}$ and

[1146] *Proof.* Proposition 6.4.10 shows that, for every $x \in \operatorname{Comp}_n$, there exists a family $(\eta_{x,y})_{y \in \operatorname{Comp}_n} \in \mathbb{N}^{\operatorname{Comp}_n}$ of elements of $\mathbb{N}$ such that

$$\mathbf{M}_x = \sum_{\substack{y \in \operatorname{Comp}_n; \\ y \underset{\text{wll}}{\leq} x}} \eta_{x,y} M_y \tag{13.172.1}$$

and

$$\eta_{x,x} \neq 0 \qquad (\text{in } \mathbb{N}). \tag{13.172.2}$$

$(f_s)_{s\in S}$) yields that the family $(\mathbf{M}_u)_{u\in\operatorname{Comp}_n}$ is a basis of the $\mathbf{k}$-module QSym_n if and only if the family $(M_u)_{u\in\operatorname{Comp}_n}$ is a basis of the $\mathbf{k}$-module QSym_n [1147]. Hence, the family $(\mathbf{M}_u)_{u\in\operatorname{Comp}_n}$ is a basis of the $\mathbf{k}$-module QSym_n (since the family $(M_u)_{u\in\operatorname{Comp}_n}$ is a basis of the $\mathbf{k}$-module QSym_n).

Now, let us forget that we fixed n. We thus have proven that the family $(\mathbf{M}_u)_{u\in\operatorname{Comp}_n}$ is a basis of QSym_n for every $n\in\mathbb{N}$. Hence, the disjoint union of the families $(\mathbf{M}_u)_{u\in\operatorname{Comp}_n}$ over all $n\in\mathbb{N}$ is a basis of the direct sum $\bigoplus_{n\in\mathbb{N}}\operatorname{QSym}_n$. Since the former disjoint union is the family $(\mathbf{M}_u)_{u\in\operatorname{Comp}}$, while the latter direct sum is the $\mathbf{k}$-module $\bigoplus_{n\in\mathbb{N}}\operatorname{QSym}_n=\operatorname{QSym}$, this rewrites as follows: The family $(\mathbf{M}_u)_{u\in\operatorname{Comp}}$ is a basis of the $\mathbf{k}$-module QSym. In other words, the family $(\mathbf{M}_u)_{u\in\mathfrak{A}^*}$ is a basis of the $\mathbf{k}$-module QSym (since $\mathfrak{A}^*=\operatorname{Comp}$).

But Lemma 6.3.7(a) (applied to $A=\operatorname{QSym}$, $b_w=M_w$ and $\mathbf{b}_u=\mathbf{M}_u$) yields that the family $(M_w)_{w\in\mathfrak{L}}$ is an algebraically independent generating set of the $\mathbf{k}$-algebra QSym if and only if the family $(\mathbf{M}_u)_{u\in\mathfrak{A}^*}$ is a basis of the $\mathbf{k}$-module QSym. Hence, the family $(M_w)_{w\in\mathfrak{L}}$ is an algebraically independent generating set of the $\mathbf{k}$-algebra QSym (since we know that the family $(\mathbf{M}_u)_{u\in\mathfrak{A}^*}$ is a basis of the $\mathbf{k}$-module QSym). This proves Proposition 6.4.14. □

13.173. **Solution to Exercise 6.5.4.** *Solution to Exercise 6.5.4.* We first notice that every $s\in\{1,2,3,\ldots\}$ and $\mathbf{i}\in\operatorname{SIS}(\ell)$ satisfy

$$(\mathbf{x}_{\mathbf{i}}^{\alpha})^s=\mathbf{x}_{\mathbf{i}}^{\alpha\{s\}} \tag{13.173.1}$$

[1148].

Consider such a family $(\eta_{x,y})_{y\in\operatorname{Comp}_n}\in\mathbb{N}^{\operatorname{Comp}_n}$ for each $x\in\operatorname{Comp}_n$. Thus, an integer $\eta_{x,y}\in\mathbb{N}\subset\mathbb{Q}\subset\mathbf{k}$ is defined for each $(x,y)\in\operatorname{Comp}_n\times\operatorname{Comp}_n$.

We observe that the only elements $\eta_{s,t}$ (with $(s,t)\in\operatorname{Comp}_n\times\operatorname{Comp}_n$) appearing in the statements (13.172.1) and (13.172.2) are those which satisfy $t\underset{\text{wll}}{\leq}s$. Hence, if some $(s,t)\in\operatorname{Comp}_n\times\operatorname{Comp}_n$ does not satisfy $t\underset{\text{wll}}{\leq}s$, then the corresponding element $\eta_{s,t}$ does not appear in any of the statements (13.172.1) and (13.172.2); as a consequence, we can arbitrarily change the value of this $\eta_{s,t}$ without running the risk of invalidating (13.172.1) and (13.172.2). Hence, we can WLOG assume that

$$\text{every } (s,t)\in\operatorname{Comp}_n\times\operatorname{Comp}_n \text{ which does not satisfy } t\underset{\text{wll}}{\leq}s \text{ must satisfy } \eta_{s,t}=0 \tag{13.172.3}$$

(otherwise, we can just set all such $\eta_{s,t}$ to 0). Assume this. Thus, the matrix $(\eta_{x,y})_{(x,y)\in\operatorname{Comp}_n\times\operatorname{Comp}_n}$ is triangular. The diagonal entries $\eta_{x,x}$ of this matrix are nonzero elements of $\mathbb{N}$ (because of (13.172.2)) and therefore invertible elements of $\mathbf{k}$ (since every nonzero element of $\mathbb{N}$ is an invertible element of $\mathbf{k}$). Thus, the matrix $(\eta_{x,y})_{(x,y)\in\operatorname{Comp}_n\times\operatorname{Comp}_n}$ (regarded as a matrix in $\mathbf{k}^{\operatorname{Comp}_n\times\operatorname{Comp}_n}$) is invertibly triangular.

Now, every $x\in\operatorname{Comp}_n$ satisfies

$$\sum_{y\in\operatorname{Comp}_n}\eta_{x,y}M_y=\underbrace{\sum_{\substack{y\in\operatorname{Comp}_n;\\ y\underset{\text{wll}}{\leq}x}}\eta_{x,y}M_y}_{\substack{=\mathbf{M}_x\\ \text{(by (13.172.1))}}}+\sum_{\substack{y\in\operatorname{Comp}_n;\\ \text{not } y\underset{\text{wll}}{\leq}x}}\underbrace{\eta_{x,y}}_{\substack{=0\\ \text{(by (13.172.3), applied to } (s,t)=(x,y))}}M_y$$
$$=\mathbf{M}_x+\underbrace{\sum_{\substack{y\in\operatorname{Comp}_n;\\ \text{not } y\underset{\text{wll}}{\leq}x}}}_{=0}0M_y=\mathbf{M}_x.$$

In other words, every $x\in\operatorname{Comp}_n$ satisfies $\mathbf{M}_x=\sum_{y\in\operatorname{Comp}_n}\eta_{x,y}M_y$. In other words, the family $(\mathbf{M}_u)_{u\in\operatorname{Comp}_n}$ expands in the family $(M_u)_{u\in\operatorname{Comp}_n}$ through the matrix $(\eta_{x,y})_{(x,y)\in\operatorname{Comp}_n\times\operatorname{Comp}_n}$. Since the latter matrix is invertibly triangular, we thus conclude that the family $(\mathbf{M}_u)_{u\in\operatorname{Comp}_n}$ expands invertibly triangularly in the family $(M_u)_{u\in\operatorname{Comp}_n}$.

[1147] Here, we have used the fact that the set Comp_n is finite.

[1148] *Proof of (13.173.1):* Let $s\in\{1,2,3,\ldots\}$ and $\mathbf{i}\in\operatorname{SIS}(\ell)$. Then, $\mathbf{i}$ is an ℓ-tuple of positive integers. Write $\mathbf{i}$ in the form $\mathbf{i}=(i_1,i_2,\ldots,i_\ell)$. Then, $\mathbf{x}_{\mathbf{i}}^{\alpha}=x_{i_1}^{\alpha_1}x_{i_2}^{\alpha_2}\cdots x_{i_\ell}^{\alpha_\ell}$ (by the definition of $\mathbf{x}_{\mathbf{i}}^{\alpha}$). Also, by the definition of $\mathbf{x}_{\mathbf{i}}^{\alpha\{s\}}$, we have $\mathbf{x}_{\mathbf{i}}^{\alpha\{s\}}=x_{i_1}^{s\alpha_1}x_{i_2}^{s\alpha_2}\cdots x_{i_\ell}^{s\alpha_\ell}$ (since $\alpha\{s\}=(s\alpha_1,s\alpha_2,\ldots,s\alpha_\ell)$). Now, taking both sides of the equality $\mathbf{x}_{\mathbf{i}}^{\alpha}=x_{i_1}^{\alpha_1}x_{i_2}^{\alpha_2}\cdots x_{i_\ell}^{\alpha_\ell}$ to the s-th power, we obtain

$$(\mathbf{x}_{\mathbf{i}}^{\alpha})^s=\left(x_{i_1}^{\alpha_1}x_{i_2}^{\alpha_2}\cdots x_{i_\ell}^{\alpha_\ell}\right)^s=\left(x_{i_1}^{\alpha_1}\right)^s\left(x_{i_2}^{\alpha_2}\right)^s\cdots\left(x_{i_\ell}^{\alpha_\ell}\right)^s=x_{i_1}^{s\alpha_1}x_{i_2}^{s\alpha_2}\cdots x_{i_\ell}^{s\alpha_\ell}=\mathbf{x}_{\mathbf{i}}^{\alpha\{s\}}.$$

This proves (13.173.1).

(a) Let s be a positive integer.

We begin with the following general observation:

If R is a topological commutative $\mathbf{k}$-algebra, if $\mathbf{I}$ is a countable (finite or not) set, and if $(s_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}} \in R^{\mathbf{I}}$ is a power-summable family of elements of R, then

$$p_s\left((s_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}}\right) = \sum_{\mathbf{i}\in\mathbf{I}} s_{\mathbf{i}}^s. \tag{13.173.2}$$

[1149] Applying this to $R = \mathbf{k}[[\mathbf{x}]]$, $\mathbf{I} = \operatorname{SIS}(\ell)$ and $\mathbf{s}_{\mathbf{i}} = \mathbf{x}_{\mathbf{i}}^{\alpha}$, we obtain

$$p_s\left(\left(\mathbf{x}_{\mathbf{i}}^{\alpha}\right)_{\mathbf{i}\in\operatorname{SIS}(\ell)}\right) = \sum_{\mathbf{i}\in\operatorname{SIS}(\ell)} \underbrace{\left(\mathbf{x}_{\mathbf{i}}^{\alpha}\right)^s}_{\substack{=\mathbf{x}_{\mathbf{i}}^{\alpha\{s\}} \\ \text{(by (13.173.1))}}} = \sum_{\mathbf{i}\in\operatorname{SIS}(\ell)} \mathbf{x}_{\mathbf{i}}^{\alpha\{s\}}.$$

[1149] *Proof of (13.173.2):* Let R be a topological commutative $\mathbf{k}$-algebra. Let $\mathbf{I}$ be a countable set. Let $(s_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}} \in R^{\mathbf{I}}$ be a power-summable family of elements of R. We need to prove that (13.173.2) holds.

We must be in one of the following two cases:

Case 1: The set $\mathbf{I}$ is infinite.

Case 2: The set $\mathbf{I}$ is finite.

Let us first consider Case 1. In this case, the set $\mathbf{I}$ is infinite, and therefore countably infinite (since it is countable). Fix a bijection $\mathrm{j} : \{1,2,3,\ldots\} \to \mathbf{I}$ (such a bijection clearly exists). Then, $p_s\left((s_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}}\right)$ is the result of substituting $s_{\mathrm{j}(1)}$, $s_{\mathrm{j}(2)}$, $s_{\mathrm{j}(3)}$, $\ldots$ for the variables x_1, x_2, x_3, $\ldots$ in p_s (by the definition of $p_s\left((s_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}}\right)$). Thus,

$$\begin{aligned}
&p_s\left((s_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}}\right) \\
&= \left(\text{the result of substituting } s_{\mathrm{j}(1)},\ s_{\mathrm{j}(2)},\ s_{\mathrm{j}(3)},\ \ldots \text{ for the variables } x_1,\ x_2,\ x_3,\ \ldots \text{ in } \underbrace{p_s}_{=\sum_{i\in\{1,2,3,\ldots\}} x_i^s}\right) \\
&= \left(\text{the result of substituting } s_{\mathrm{j}(1)},\ s_{\mathrm{j}(2)},\ s_{\mathrm{j}(3)},\ \ldots \text{ for the variables } x_1,\ x_2,\ x_3,\ \ldots \text{ in } \sum_{i\in\{1,2,3,\ldots\}} x_i^s\right) \\
&= \sum_{i\in\{1,2,3,\ldots\}} s_{\mathrm{j}(i)}^s = \sum_{\mathbf{i}\in\mathbf{I}} s_{\mathbf{i}}^s
\end{aligned}$$

(here, we have substituted $\mathbf{i}$ for $\mathrm{j}(i)$ in the sum, since the map $\mathrm{j} : \{1,2,3,\ldots\} \to \mathbf{I}$ is a bijection). Thus, (13.173.2) is proven in Case 1.

Let us now consider Case 2. In this case, the set $\mathbf{I}$ is finite. Fix a bijection $\mathrm{j} : \{1,2,\ldots,|\mathbf{I}|\} \to \mathbf{I}$ (such a bijection clearly exists). Then, $p_s\left((s_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}}\right)$ is the result of substituting $s_{\mathrm{j}(1)}$, $s_{\mathrm{j}(2)}$, $\ldots$, $s_{\mathrm{j}(|\mathbf{I}|)}$, 0, 0, 0, $\ldots$ for the variables x_1, x_2, x_3, $\ldots$ in p_s (by the definition of $p_s\left((s_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}}\right)$). Thus,

$$\begin{aligned}
&p_s\left((s_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}}\right) \\
&= \left(\text{the result of substituting } s_{\mathrm{j}(1)},\ s_{\mathrm{j}(2)},\ \ldots,\ s_{\mathrm{j}(|\mathbf{I}|)},\ 0,\ 0,\ 0,\ \ldots \text{ for the variables } x_1,\ x_2,\ x_3,\ \ldots \text{ in } \underbrace{\underbrace{p_s}_{=\sum_{i=1}^{\infty} x_i^s}}_{=\sum_{i=1}^{|\mathbf{I}|} x_i^s + \sum_{i=|\mathbf{I}|+1}^{\infty} x_i^s}\right) \\
&= \left(\text{the result of substituting } s_{\mathrm{j}(1)},\ s_{\mathrm{j}(2)},\ \ldots,\ s_{\mathrm{j}(|\mathbf{I}|)},\ 0,\ 0,\ 0,\ \ldots \text{ for the variables } x_1,\ x_2,\ x_3,\ \ldots \text{ in } \sum_{i=1}^{|\mathbf{I}|} x_i^s + \sum_{i=|\mathbf{I}|+1}^{\infty} x_i^s\right) \\
&= \underbrace{\sum_{i=1}^{|\mathbf{I}|}}_{=\sum_{i\in\{1,2,\ldots,|\mathbf{I}|\}}} s_{\mathrm{j}(i)}^s + \sum_{i=\mathrm{j}(|\mathbf{I}|)+1}^{\infty} \underbrace{0^s}_{\substack{=0 \\ (\text{since } s>0)}} = \sum_{i\in\{1,2,\ldots,|\mathbf{I}|\}} s_{\mathrm{j}(i)}^s + \underbrace{\sum_{i=\mathrm{j}(|\mathbf{I}|)+1}^{\infty} 0}_{=0} = \sum_{i\in\{1,2,\ldots,|\mathbf{I}|\}} s_{\mathrm{j}(i)}^s \\
&= \sum_{\mathbf{i}\in\mathbf{I}} s_{\mathbf{i}}^s
\end{aligned}$$

(here, we have substituted $\mathbf{i}$ for $\mathrm{j}(i)$ in the sum, since the map $\mathrm{j} : \{1,2,\ldots,|\mathbf{I}|\} \to \mathbf{I}$ is a bijection). Thus, (13.173.2) is proven in Case 2.

We have thus proven (13.173.2) in each of the Cases 1 and 2. Since these two Cases are the only cases that can occur, we thus conclude that (13.173.2) holds, qed.

Compared with $M_{\alpha\{s\}} = \sum_{\mathbf{i}\in\mathrm{SIS}(\ell)} \mathbf{x}_{\mathbf{i}}^{\alpha\{s\}}$ (by (6.5.1), applied to $\alpha\{s\}$ instead of α), this yields $p_s\left(\left(\mathbf{x}_{\mathbf{i}}^{\alpha}\right)_{\mathbf{i}\in\mathrm{SIS}(\ell)}\right) = M_{\alpha\{s\}}$. Exercise 6.5.4(a) is now solved.

(b) Let $s \in \mathbb{N}$. We shall first make some general statements.

- If R is a topological commutative $\mathbf{k}$-algebra, if $\mathbf{I}$ is a countable (finite or not) totally ordered set, and if $(s_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}} \in R^{\mathbf{I}}$ is a power-summable family of elements of R, then

$$\sum_{\substack{(\mathbf{i}_1,\mathbf{i}_2,\ldots,\mathbf{i}_s)\in\mathbf{I}^s;\\ \mathbf{i}_1<\mathbf{i}_2<\cdots<\mathbf{i}_s}} s_{\mathbf{i}_1}s_{\mathbf{i}_2}\cdots s_{\mathbf{i}_s} = \sum_{\substack{\mathbf{K}\subset\mathbf{I};\\ |\mathbf{K}|=s}} \prod_{\mathbf{i}\in\mathbf{K}} s_{\mathbf{i}}. \tag{13.173.3}$$

[1150]

- If R is a commutative $\mathbf{k}$-algebra, if $\mathbf{I}$ and $\mathbf{J}$ are two sets, if $\mathsf{j} : \mathbf{J} \to \mathbf{I}$ is a bijection, and if $(s_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}} \in R^{\mathbf{I}}$ is a family of elements of R, then

$$\prod_{\mathbf{i}\in\mathsf{j}(\mathbf{K})} s_{\mathbf{i}} = \prod_{\mathbf{i}\in\mathbf{K}} s_{\mathsf{j}(\mathbf{i})} \qquad \text{for every finite subset } \mathbf{K} \text{ of } \mathbf{J}. \tag{13.173.4}$$

[1151]

- If R is a topological commutative $\mathbf{k}$-algebra, if $\mathbf{I}$ and $\mathbf{J}$ are two countable totally ordered sets, if $\mathsf{j} : \mathbf{J} \to \mathbf{I}$ is a (not necessarily order-preserving!) bijection, and if $(s_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}} \in R^{\mathbf{I}}$ is a power-summable family of elements of R, then

$$\sum_{\substack{(\mathbf{i}_1,\mathbf{i}_2,\ldots,\mathbf{i}_s)\in\mathbf{I}^s;\\ \mathbf{i}_1<\mathbf{i}_2<\cdots<\mathbf{i}_s}} s_{\mathbf{i}_1}s_{\mathbf{i}_2}\cdots s_{\mathbf{i}_s} = \sum_{\substack{(\mathbf{j}_1,\mathbf{j}_2,\ldots,\mathbf{j}_s)\in\mathbf{J}^s;\\ \mathbf{j}_1<\mathbf{j}_2<\cdots<\mathbf{j}_s}} s_{\mathsf{j}(\mathbf{j}_1)}s_{\mathsf{j}(\mathbf{j}_2)}\cdots s_{\mathsf{j}(\mathbf{j}_s)}. \tag{13.173.5}$$

[1152]

[1150] *Proof of (13.173.3):* Let R be a topological commutative $\mathbf{k}$-algebra. Let $\mathbf{I}$ be a countable totally ordered set. Let $(s_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}} \in R^{\mathbf{I}}$ be a power-summable family of elements of R. We need to prove that (13.173.3) holds.

The order on $\mathbf{I}$ is total. Hence, the map

$$\{(\mathbf{i}_1,\mathbf{i}_2,\ldots,\mathbf{i}_s)\in\mathbf{I}^s \mid \mathbf{i}_1<\mathbf{i}_2<\cdots<\mathbf{i}_s\} \to \{\mathbf{K}\subset\mathbf{I} \mid |\mathbf{K}|=s\},$$
$$(\mathbf{i}_1,\mathbf{i}_2,\ldots,\mathbf{i}_s) \mapsto \{\mathbf{i}_1,\mathbf{i}_2,\ldots,\mathbf{i}_s\}$$

is a bijection (because every subset $\mathbf{K}$ of $\mathbf{I}$ satisfying $|\mathbf{K}| = s$ can be written in the form $\mathbf{K} = \{\mathbf{i}_1,\mathbf{i}_2,\ldots,\mathbf{i}_s\}$ for a unique s-tuple $(\mathbf{i}_1,\mathbf{i}_2,\ldots,\mathbf{i}_s)\in\mathbf{I}^s$ satisfying $\mathbf{i}_1<\mathbf{i}_2<\cdots<\mathbf{i}_s$).

Now, let $(\mathbf{i}_1,\mathbf{i}_2,\ldots,\mathbf{i}_s)\in\mathbf{I}^s$ be such that $\mathbf{i}_1<\mathbf{i}_2<\cdots<\mathbf{i}_s$. Then, the s-tuple $(\mathbf{i}_1,\mathbf{i}_2,\ldots,\mathbf{i}_s)$ is a list of all elements of the set $\{\mathbf{i}_1,\mathbf{i}_2,\ldots,\mathbf{i}_s\}$, each occurring exactly once (since the elements $\mathbf{i}_1$, $\mathbf{i}_2$, $\ldots$, $\mathbf{i}_s$ are distinct (since $\mathbf{i}_1<\mathbf{i}_2<\cdots<\mathbf{i}_s$)). Hence, $\prod_{\mathbf{i}\in\{\mathbf{i}_1,\mathbf{i}_2,\ldots,\mathbf{i}_s\}} s_{\mathbf{i}} = s_{\mathbf{i}_1}s_{\mathbf{i}_2}\cdots s_{\mathbf{i}_s}$, so that $s_{\mathbf{i}_1}s_{\mathbf{i}_2}\cdots s_{\mathbf{i}_s} = \prod_{\mathbf{i}\in\{\mathbf{i}_1,\mathbf{i}_2,\ldots,\mathbf{i}_s\}} s_{\mathbf{i}}$.

Now, let us forget that we fixed $(\mathbf{i}_1,\mathbf{i}_2,\ldots,\mathbf{i}_s)\in\mathbf{I}^s$. We thus have shown that $s_{\mathbf{i}_1}s_{\mathbf{i}_2}\cdots s_{\mathbf{i}_s} = \prod_{\mathbf{i}\in\{\mathbf{i}_1,\mathbf{i}_2,\ldots,\mathbf{i}_s\}} s_{\mathbf{i}}$ for every $(\mathbf{i}_1,\mathbf{i}_2,\ldots,\mathbf{i}_s)\in\mathbf{I}^s$ be such that $\mathbf{i}_1<\mathbf{i}_2<\cdots<\mathbf{i}_s$. Hence,

$$\sum_{\substack{(\mathbf{i}_1,\mathbf{i}_2,\ldots,\mathbf{i}_s)\in\mathbf{I}^s;\\ \mathbf{i}_1<\mathbf{i}_2<\cdots<\mathbf{i}_s}} \underbrace{s_{\mathbf{i}_1}s_{\mathbf{i}_2}\cdots s_{\mathbf{i}_s}}_{=\prod_{\mathbf{i}\in\{\mathbf{i}_1,\mathbf{i}_2,\ldots,\mathbf{i}_s\}} s_{\mathbf{i}}} = \sum_{\substack{(\mathbf{i}_1,\mathbf{i}_2,\ldots,\mathbf{i}_s)\in\mathbf{I}^s;\\ \mathbf{i}_1<\mathbf{i}_2<\cdots<\mathbf{i}_s}} \prod_{\mathbf{i}\in\{\mathbf{i}_1,\mathbf{i}_2,\ldots,\mathbf{i}_s\}} s_{\mathbf{i}} = \sum_{\substack{\mathbf{K}\subset\mathbf{I};\\ |\mathbf{K}|=s}} \prod_{\mathbf{i}\in\mathbf{K}} s_{\mathbf{i}}$$

(here, we have substituted $\mathbf{K}$ for $\{\mathbf{i}_1,\mathbf{i}_2,\ldots,\mathbf{i}_s\}$ in the sum, since the map

$$\{(\mathbf{i}_1,\mathbf{i}_2,\ldots,\mathbf{i}_s)\in\mathbf{I}^s \mid \mathbf{i}_1<\mathbf{i}_2<\cdots<\mathbf{i}_s\} \to \{\mathbf{K}\subset\mathbf{I} \mid |\mathbf{K}|=s\},$$
$$(\mathbf{i}_1,\mathbf{i}_2,\ldots,\mathbf{i}_s) \mapsto \{\mathbf{i}_1,\mathbf{i}_2,\ldots,\mathbf{i}_s\}$$

is a bijection). This proves (13.173.3).

[1151] *Proof of (13.173.4):* Let R be a commutative $\mathbf{k}$-algebra. Let $\mathbf{I}$ and $\mathbf{J}$ be two sets. Let $\mathsf{j} : \mathbf{J} \to \mathbf{I}$ be a bijection. Let $(s_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}} \in R^{\mathbf{I}}$ be a family of elements of R. Let $\mathbf{K}$ be a finite subset of $\mathbf{J}$. The map j is a bijection, and thus injective. Hence, the map $\mathbf{K} \to \mathsf{j}(\mathbf{K})$, $\mathbf{i} \mapsto \mathsf{j}(\mathbf{i})$ is a bijection. Hence, we can substitute $\mathsf{j}(\mathbf{i})$ for $\mathbf{i}$ in the product $\prod_{\mathbf{i}\in\mathsf{j}(\mathbf{K})} s_{\mathbf{i}}$. As a result, we obtain $\prod_{\mathbf{i}\in\mathsf{j}(\mathbf{K})} s_{\mathbf{i}} = \prod_{\mathbf{i}\in\mathbf{K}} s_{\mathsf{j}(\mathbf{i})}$. This proves (13.173.4).

[1152] *Proof of (13.173.5):* Let R be a topological commutative $\mathbf{k}$-algebra. Let $\mathbf{I}$ and $\mathbf{J}$ be two countable totally ordered sets. Let $\mathsf{j} : \mathbf{J} \to \mathbf{I}$ be a bijection. Let $(s_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}} \in R^{\mathbf{I}}$ be a power-summable family of elements of R.

Notice that $\left(s_{\mathsf{j}(\mathbf{i})}\right)_{\mathbf{i}\in\mathbf{J}} \in R^{\mathbf{J}}$ is a reindexing of the family $(s_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}} \in R^{\mathbf{I}}$ (since $\mathsf{j} : \mathbf{J} \to \mathbf{I}$ is a bijection), and thus a power-summable family of elements of R (since $(s_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}}$ is a power-summable family of elements of R).

The map

$$\{\mathbf{K}\subset\mathbf{J} \mid |\mathbf{K}|=s\} \to \{\mathbf{K}\subset\mathbf{I} \mid |\mathbf{K}|=s\},$$
$$\mathbf{K} \mapsto \mathsf{j}(\mathbf{K})$$

- If R is a topological commutative $\mathbf{k}$-algebra, if $\mathbf{I}$ is a countable (finite or not) totally ordered set, and if $(s_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}} \in R^{\mathbf{I}}$ is a power-summable family of elements of R, then

$$e_s\left((s_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}}\right) = \sum_{\substack{(\mathbf{i}_1,\mathbf{i}_2,\ldots,\mathbf{i}_s)\in\mathbf{I}^s;\\ \mathbf{i}_1<\mathbf{i}_2<\cdots<\mathbf{i}_s}} s_{\mathbf{i}_1}s_{\mathbf{i}_2}\cdots s_{\mathbf{i}_s}. \tag{13.173.7}$$

[1153]

Now, let us return to our proof. Applying (13.173.7) to $R = \mathbf{k}[[\mathbf{x}]]$, $\mathbf{I} = \operatorname{SIS}(\ell)$ and $\mathbf{s_i} = \mathbf{x}_{\mathbf{i}}^{\alpha}$, we obtain

$$e_s\left((\mathbf{x}_{\mathbf{i}}^{\alpha})_{\mathbf{i}\in\operatorname{SIS}(\ell)}\right) = \sum_{\substack{(\mathbf{i}_1,\mathbf{i}_2,\ldots,\mathbf{i}_s)\in(\operatorname{SIS}(\ell))^s;\\ \mathbf{i}_1<\mathbf{i}_2<\cdots<\mathbf{i}_s}} \mathbf{x}_{\mathbf{i}_1}^{\alpha}\mathbf{x}_{\mathbf{i}_2}^{\alpha}\cdots\mathbf{x}_{\mathbf{i}_s}^{\alpha}.$$

Thus,

$$M_{\alpha}^{\langle s\rangle} = e_s\left((\mathbf{x}_{\mathbf{i}}^{\alpha})_{\mathbf{i}\in\operatorname{SIS}(\ell)}\right) = \sum_{\substack{(\mathbf{i}_1,\mathbf{i}_2,\ldots,\mathbf{i}_s)\in(\operatorname{SIS}(\ell))^s;\\ \mathbf{i}_1<\mathbf{i}_2<\cdots<\mathbf{i}_s}} \mathbf{x}_{\mathbf{i}_1}^{\alpha}\mathbf{x}_{\mathbf{i}_2}^{\alpha}\cdots\mathbf{x}_{\mathbf{i}_s}^{\alpha}.$$

This solves Exercise 6.5.4(b).

is a bijection (since j is a bijection from $\mathbf{J}$ to $\mathbf{I}$). Hence, we can substitute $\mathrm{j}(\mathbf{K})$ for $\mathbf{K}$ in the sum $\sum_{\substack{\mathbf{K}\subset\mathbf{I};\\ |\mathbf{K}|=s}} \prod_{\mathbf{i}\in\mathbf{K}} s_{\mathbf{i}}$, and as a result we obtain

$$\sum_{\substack{\mathbf{K}\subset\mathbf{I};\\ |\mathbf{K}|=s}} \prod_{\mathbf{i}\in\mathbf{K}} s_{\mathbf{i}} = \sum_{\substack{\mathbf{K}\subset\mathbf{J};\\ |\mathbf{K}|=s}} \underbrace{\prod_{\mathbf{i}\in\mathrm{j}(\mathbf{K})} s_{\mathbf{i}}}_{\substack{=\prod_{\mathbf{i}\in\mathbf{K}} s_{\mathrm{j}(\mathbf{i})}\\ \text{(by (13.173.4))}}} = \sum_{\substack{\mathbf{K}\subset\mathbf{J};\\ |\mathbf{K}|=s}} \prod_{\mathbf{i}\in\mathbf{K}} s_{\mathrm{j}(\mathbf{i})}.$$

Thus, (13.173.3) becomes

$$\sum_{\substack{(\mathbf{i}_1,\mathbf{i}_2,\ldots,\mathbf{i}_s)\in\mathbf{I}^s;\\ \mathbf{i}_1<\mathbf{i}_2<\cdots<\mathbf{i}_s}} s_{\mathbf{i}_1}s_{\mathbf{i}_2}\cdots s_{\mathbf{i}_s} = \sum_{\substack{\mathbf{K}\subset\mathbf{I};\\ |\mathbf{K}|=s}} \prod_{\mathbf{i}\in\mathbf{K}} s_{\mathbf{i}} = \sum_{\substack{\mathbf{K}\subset\mathbf{J};\\ |\mathbf{K}|=s}} \prod_{\mathbf{i}\in\mathbf{K}} s_{\mathrm{j}(\mathbf{i})}. \tag{13.173.6}$$

But

$$\begin{aligned}\sum_{\substack{(\mathbf{j}_1,\mathbf{j}_2,\ldots,\mathbf{j}_s)\in\mathbf{J}^s;\\ \mathbf{j}_1<\mathbf{j}_2<\cdots<\mathbf{j}_s}} s_{\mathrm{j}(\mathbf{j}_1)}s_{\mathrm{j}(\mathbf{j}_2)}\cdots s_{\mathrm{j}(\mathbf{j}_s)} &= \sum_{\substack{(\mathbf{i}_1,\mathbf{i}_2,\ldots,\mathbf{i}_s)\in\mathbf{J}^s;\\ \mathbf{i}_1<\mathbf{i}_2<\cdots<\mathbf{i}_s}} s_{\mathrm{j}(\mathbf{i}_1)}s_{\mathrm{j}(\mathbf{i}_2)}\cdots s_{\mathrm{j}(\mathbf{i}_s)}\\ &\qquad \left(\text{here, we renamed the summation index } (\mathbf{j}_1,\mathbf{j}_2,\ldots,\mathbf{j}_s) \text{ as } (\mathbf{i}_1,\mathbf{i}_2,\ldots,\mathbf{i}_s)\right)\\ &= \sum_{\substack{\mathbf{K}\subset\mathbf{J};\\ |\mathbf{K}|=s}} \prod_{\mathbf{i}\in\mathbf{K}} s_{\mathrm{j}(\mathbf{i})}\end{aligned}$$

(by (13.173.3), applied to $\mathbf{J}$ and $s_{\mathrm{j}(\mathbf{i})}$ instead of $\mathbf{I}$ and $s_{\mathbf{i}}$). Compared with (13.173.6), this yields

$$\sum_{\substack{(\mathbf{i}_1,\mathbf{i}_2,\ldots,\mathbf{i}_s)\in\mathbf{I}^s;\\ \mathbf{i}_1<\mathbf{i}_2<\cdots<\mathbf{i}_s}} s_{\mathbf{i}_1}s_{\mathbf{i}_2}\cdots s_{\mathbf{i}_s} = \sum_{\substack{(\mathbf{j}_1,\mathbf{j}_2,\ldots,\mathbf{j}_s)\in\mathbf{J}^s;\\ \mathbf{j}_1<\mathbf{j}_2<\cdots<\mathbf{j}_s}} s_{\mathrm{j}(\mathbf{j}_1)}s_{\mathrm{j}(\mathbf{j}_2)}\cdots s_{\mathrm{j}(\mathbf{j}_s)}.$$

This proves (13.173.5).

[1153] *Proof of (13.173.7):* Let R be a topological commutative $\mathbf{k}$-algebra. Let $\mathbf{I}$ be a countable totally ordered set. Let $(s_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}} \in R^{\mathbf{I}}$ be a power-summable family of elements of R. We need to prove that (13.173.7) holds.

The definition of e_s yields

$$e_s = \sum_{\substack{(i_1,i_2,\ldots,i_s)\in\{1,2,3,\ldots\}^s;\\ i_1<i_2<\cdots<i_s}} x_{i_1}x_{i_2}\cdots x_{i_s} = \sum_{\substack{(\mathbf{j}_1,\mathbf{j}_2,\ldots,\mathbf{j}_s)\in\{1,2,3,\ldots\}^s;\\ \mathbf{j}_1<\mathbf{j}_2<\cdots<\mathbf{j}_s}} x_{\mathbf{j}_1}x_{\mathbf{j}_2}\cdots x_{\mathbf{j}_s} \tag{13.173.8}$$

(here, we renamed the summation index $(i_1,i_2,\ldots,i_s)$ as $(\mathbf{j}_1,\mathbf{j}_2,\ldots,\mathbf{j}_s)$).

We must be in one of the following two cases:

Case 1: The set $\mathbf{I}$ is infinite.

Case 2: The set $\mathbf{I}$ is finite.

Let us first consider Case 1. In this case, the set $\mathbf{I}$ is infinite, and therefore countably infinite (since it is countable). Let $\mathbf{J}$ denote the totally ordered set $\{1,2,3,\ldots\}$. Then, (13.173.8) rewrites as

$$e_s = \sum_{\substack{(\mathbf{j}_1,\mathbf{j}_2,\ldots,\mathbf{j}_s)\in\{1,2,3,\ldots\}^s;\\ \mathbf{j}_1<\mathbf{j}_2<\cdots<\mathbf{j}_s}} x_{\mathbf{j}_1}x_{\mathbf{j}_2}\cdots x_{\mathbf{j}_s} = \sum_{\substack{(\mathbf{j}_1,\mathbf{j}_2,\ldots,\mathbf{j}_s)\in\mathbf{J}^s;\\ \mathbf{j}_1<\mathbf{j}_2<\cdots<\mathbf{j}_s}} x_{\mathbf{j}_1}x_{\mathbf{j}_2}\cdots x_{\mathbf{j}_s} \tag{13.173.9}$$

(since $\{1,2,3,\ldots\} = \mathbf{J}$).

Fix a bijection $\mathrm{j} : \{1, 2, 3, \ldots\} \to \mathbf{I}$ (such a bijection clearly exists, since $\mathbf{I}$ is countably infinite). Then, $e_s\left((s_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}}\right)$ is the result of substituting $s_{\mathrm{j}(1)}$, $s_{\mathrm{j}(2)}$, $s_{\mathrm{j}(3)}$, $\ldots$ for the variables x_1, x_2, x_3, $\ldots$ in e_s (by the definition of $e_s\left((s_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}}\right)$). Thus,

$$\begin{aligned}
&e_s\left((s_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}}\right)\\
&= \left(\text{the result of substituting } s_{\mathrm{j}(1)},\ s_{\mathrm{j}(2)},\ s_{\mathrm{j}(3)},\ \ldots \text{ for the variables } x_1,\ x_2,\ x_3,\ \ldots \text{ in } \underbrace{e_s}_{\substack{=\sum_{\substack{(\mathbf{j}_1,\mathbf{j}_2,\ldots,\mathbf{j}_s)\in\mathbf{J}^s;\\ \mathbf{j}_1<\mathbf{j}_2<\cdots<\mathbf{j}_s}} x_{\mathbf{j}_1}x_{\mathbf{j}_2}\cdots x_{\mathbf{j}_s}\\ \text{(by (13.173.9))}}}\right)\\
&= \left(\text{the result of substituting } s_{\mathrm{j}(1)},\ s_{\mathrm{j}(2)},\ s_{\mathrm{j}(3)},\ \ldots \text{ for the variables } x_1,\ x_2,\ x_3,\ \ldots \text{ in } \sum_{\substack{(\mathbf{j}_1,\mathbf{j}_2,\ldots,\mathbf{j}_s)\in\mathbf{J}^s;\\ \mathbf{j}_1<\mathbf{j}_2<\cdots<\mathbf{j}_s}} x_{\mathbf{j}_1}x_{\mathbf{j}_2}\cdots x_{\mathbf{j}_s}\right)\\
&= \sum_{\substack{(\mathbf{j}_1,\mathbf{j}_2,\ldots,\mathbf{j}_s)\in\mathbf{J}^s;\\ \mathbf{j}_1<\mathbf{j}_2<\cdots<\mathbf{j}_s}} s_{\mathrm{j}(\mathbf{j}_1)}s_{\mathrm{j}(\mathbf{j}_2)}\cdots s_{\mathrm{j}(\mathbf{j}_s)}.
\end{aligned} \tag{13.173.10}$$

But $\mathbf{J}$ is a countable totally ordered set (since $\mathbf{J} = \{1, 2, 3, \ldots\}$). Also, j is a bijection $\{1, 2, 3, \ldots\} \to \mathbf{I}$. In other words, j is a bijection $\mathbf{J} \to \mathbf{I}$ (since $\mathbf{J} = \{1, 2, 3, \ldots\}$). Now, (13.173.10) becomes

$$e_s\left((s_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}}\right) = \sum_{\substack{(\mathbf{j}_1,\mathbf{j}_2,\ldots,\mathbf{j}_s)\in\mathbf{J}^s;\\ \mathbf{j}_1<\mathbf{j}_2<\cdots<\mathbf{j}_s}} s_{\mathrm{j}(\mathbf{j}_1)}s_{\mathrm{j}(\mathbf{j}_2)}\cdots s_{\mathrm{j}(\mathbf{j}_s)} = \sum_{\substack{(\mathbf{i}_1,\mathbf{i}_2,\ldots,\mathbf{i}_s)\in\mathbf{I}^s;\\ \mathbf{i}_1<\mathbf{i}_2<\cdots<\mathbf{i}_s}} s_{\mathbf{i}_1}s_{\mathbf{i}_2}\cdots s_{\mathbf{i}_s}$$

(by (13.173.5)). Thus, (13.173.7) is proven in Case 1.

Let us now consider Case 2. In this case, the set $\mathbf{I}$ is finite. Let $\mathbf{J}$ denote the totally ordered set $\{1, 2, \ldots, |\mathbf{I}|\}$. Then, $\mathbf{J} = \{1, 2, \ldots, |\mathbf{I}|\} \subset \{1, 2, 3, \ldots\}$ and $\{1, 2, 3, \ldots\} \setminus \underbrace{\mathbf{J}}_{=\{1,2,\ldots,|\mathbf{I}|\}} = \{1, 2, 3, \ldots\} \setminus \{1, 2, \ldots, |\mathbf{I}|\} = \{|\mathbf{I}| + 1, |\mathbf{I}| + 2, |\mathbf{I}| + 3, \ldots\}$.

Fix a bijection $\mathrm{j} : \{1, 2, \ldots, |\mathbf{I}|\} \to \mathbf{I}$ (such a bijection clearly exists, since $\mathbf{I}$ is finite). Then, $e_s\left((s_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}}\right)$ is the result of substituting $s_{\mathrm{j}(1)}$, $s_{\mathrm{j}(2)}$, $\ldots$, $s_{\mathrm{j}(|\mathbf{I}|)}$, $0, 0, 0, \ldots$ for the variables x_1, x_2, x_3, $\ldots$ in e_s (by the definition of $e_s\left((s_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}}\right)$). Thus,

$$\begin{aligned}
&e_s\left((s_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}}\right)\\
&= \left(\text{the result of substituting } s_{\mathrm{j}(1)},\ s_{\mathrm{j}(2)},\ \ldots,\ s_{\mathrm{j}(|\mathbf{I}|)},\ 0,\ 0,\ 0,\ \ldots \text{ for the variables } x_1,\ x_2,\ x_3,\ \ldots \text{ in } \underbrace{e_s}_{\substack{=\sum_{\substack{(\mathbf{j}_1,\mathbf{j}_2,\ldots,\mathbf{j}_s)\in\{1,2,3,\ldots\}^s;\\ \mathbf{j}_1<\mathbf{j}_2<\cdots<\mathbf{j}_s}} x_{\mathbf{j}_1}x_{\mathbf{j}_2}\cdots x_{\mathbf{j}_s}\\ \text{(by (13.173.8))}}}\right)\\
&= \left(\text{the result of substituting } s_{\mathrm{j}(1)},\ s_{\mathrm{j}(2)},\ \ldots,\ s_{\mathrm{j}(|\mathbf{I}|)},\ 0,\ 0,\ 0,\ \ldots \text{ for the variables } x_1,\ x_2,\ x_3,\ \ldots \text{ in } \sum_{\substack{(\mathbf{j}_1,\mathbf{j}_2,\ldots,\mathbf{j}_s)\in\{1,2,3,\ldots\}^s;\\ \mathbf{j}_1<\mathbf{j}_2<\cdots<\mathbf{j}_s}} x_{\mathbf{j}_1}x_{\mathbf{j}_2}\cdots x_{\mathbf{j}_s}\right)\\
&= \sum_{\substack{(\mathbf{j}_1,\mathbf{j}_2,\ldots,\mathbf{j}_s)\in\{1,2,3,\ldots\}^s;\\ \mathbf{j}_1<\mathbf{j}_2<\cdots<\mathbf{j}_s}} \left(\text{the result of substituting } s_{\mathrm{j}(1)},\ s_{\mathrm{j}(2)},\ \ldots,\ s_{\mathrm{j}(|\mathbf{I}|)},\ 0,\ 0,\ 0,\ \ldots \text{ for the variables } x_1,\ x_2,\ x_3,\ \ldots \text{ in } x_{\mathbf{j}_1}x_{\mathbf{j}_2}\cdots x_{\mathbf{j}_s}\right).
\end{aligned} \tag{13.173.11}$$

Now, let us fix some $(\mathbf{j}_1, \mathbf{j}_2, \ldots, \mathbf{j}_s) \in \mathbf{J}^s$. Then, every $k \in \{1, 2, \ldots, s\}$ satisfies $\mathbf{j}_k \in \mathbf{J} = \{1, 2, \ldots, |\mathbf{I}|\}$. Hence, for every $k \in \{1, 2, \ldots, s\}$, the substitution of $s_{\mathrm{j}(1)}$, $s_{\mathrm{j}(2)}$, $\ldots$, $s_{\mathrm{j}(|\mathbf{I}|)}$, $0, 0, 0, \ldots$ for the variables x_1, x_2, x_3, $\ldots$ transforms the indeterminate $x_{\mathbf{j}_k}$ into $s_{\mathrm{j}(\mathbf{j}_k)}$. Consequently, the substitution of $s_{\mathrm{j}(1)}$, $s_{\mathrm{j}(2)}$, $\ldots$, $s_{\mathrm{j}(|\mathbf{I}|)}$, $0, 0, 0, \ldots$ for the variables x_1, x_2, x_3, $\ldots$ transforms the product $x_{\mathbf{j}_1}x_{\mathbf{j}_2}\cdots x_{\mathbf{j}_s}$ into $s_{\mathrm{j}(\mathbf{j}_1)}s_{\mathrm{j}(\mathbf{j}_2)}\cdots s_{\mathrm{j}(\mathbf{j}_s)}$. In other words,

$$\begin{aligned}
&\left(\text{the result of substituting } s_{\mathrm{j}(1)},\ s_{\mathrm{j}(2)},\ \ldots,\ s_{\mathrm{j}(|\mathbf{I}|)},\ 0,\ 0,\ 0,\ \ldots \text{ for the variables } x_1,\ x_2,\ x_3,\ \ldots \text{ in } x_{\mathbf{j}_1}x_{\mathbf{j}_2}\cdots x_{\mathbf{j}_s}\right)\\
&= s_{\mathrm{j}(\mathbf{j}_1)}s_{\mathrm{j}(\mathbf{j}_2)}\cdots s_{\mathrm{j}(\mathbf{j}_s)}.
\end{aligned} \tag{13.173.12}$$

(c) Every $\mathbf{i} \in \operatorname{SIS}(\ell)$ satisfies $(\mathbf{x}_{\mathbf{i}}^{\alpha})^n = \mathbf{x}_{\mathbf{i}}^{\alpha\{n\}}$ (by (13.173.1), applied to n instead of s). In other words, every $\mathbf{i} \in \operatorname{SIS}(\ell)$ satisfies $\mathbf{x}_{\mathbf{i}}^{\alpha\{n\}} = (\mathbf{x}_{\mathbf{i}}^{\alpha})^n$. Now, by the definition of $M_{\alpha\{n\}}^{\langle s\rangle}$, we have

$$M_{\alpha\{n\}}^{\langle s\rangle} = e_s\left(\left(\underbrace{\mathbf{x}_{\mathbf{i}}^{\alpha\{n\}}}_{=(\mathbf{x}_{\mathbf{i}}^{\alpha})^n}\right)_{\mathbf{i}\in\operatorname{SIS}(\ell)}\right) = e_s\left(\left((\mathbf{x}_{\mathbf{i}}^{\alpha})^n\right)_{\mathbf{i}\in\operatorname{SIS}(\ell)}\right). \tag{13.173.15}$$

We now notice that if R is a topological commutative $\mathbf{k}$-algebra, if $\mathbf{I}$ is a countable set, and if $(s_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}} \in R^{\mathbf{I}}$ is a power-summable family of elements of R, then

$$e_s^{\langle n\rangle}\left((\mathbf{s}_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}}\right) = e_s\left((\mathbf{s}_{\mathbf{i}}^n)_{\mathbf{i}\in\mathbf{I}}\right). \tag{13.173.16}$$

Now, let us forget that we fixed $(\mathbf{j}_1, \mathbf{j}_2, \ldots, \mathbf{j}_s) \in \mathbf{J}^s$. We thus have shown that (13.173.12) holds for every $(\mathbf{j}_1, \mathbf{j}_2, \ldots, \mathbf{j}_s) \in \mathbf{J}^s$.

On the other hand, let us fix some $(\mathbf{j}_1, \mathbf{j}_2, \ldots, \mathbf{j}_s) \in \{1, 2, 3, \ldots\}^s \setminus \mathbf{J}^s$. Since $(\mathbf{j}_1, \mathbf{j}_2, \ldots, \mathbf{j}_s) \in \{1, 2, 3, \ldots\}^s \setminus \mathbf{J}^s$, we must have $(\mathbf{j}_1, \mathbf{j}_2, \ldots, \mathbf{j}_s) \in \{1, 2, 3, \ldots\}^s$ but $(\mathbf{j}_1, \mathbf{j}_2, \ldots, \mathbf{j}_s) \notin \mathbf{J}^s$. Since $(\mathbf{j}_1, \mathbf{j}_2, \ldots, \mathbf{j}_s) \notin \mathbf{J}^s$, there exists some $k \in \{1, 2, \ldots, s\}$ such that $\mathbf{j}_k \notin \mathbf{J}$. Fix this k. We have $\mathbf{j}_k \in \{1, 2, 3, \ldots\}$ but $\mathbf{j}_k \notin \mathbf{J}$; therefore, $\mathbf{j}_k \in \{1, 2, 3, \ldots\} \setminus \mathbf{J} = \{|\mathbf{I}|+1, |\mathbf{I}|+2, |\mathbf{I}|+3, \ldots\}$. In other words, $\mathbf{j}_k > |\mathbf{I}|$. Hence, the substitution of $s_{\mathbf{j}(1)}$, $s_{\mathbf{j}(2)}$, $\ldots$, $s_{\mathbf{j}(|\mathbf{I}|)}$, $0, 0, 0, \ldots$ for the variables x_1, x_2, x_3, $\ldots$ transforms the indeterminate $x_{\mathbf{j}_k}$ into 0. In other words,

$$\begin{aligned}&\left(\text{the result of substituting } s_{\mathbf{j}(1)}, s_{\mathbf{j}(2)}, \ldots, s_{\mathbf{j}(|\mathbf{I}|)}, 0, 0, 0, \ldots \text{ for the variables } x_1, x_2, x_3, \ldots \text{ in } x_{\mathbf{j}_k}\right)\\ &= 0.\end{aligned}$$

Now,

$$\begin{aligned}&\left(\text{the result of substituting } s_{\mathbf{j}(1)}, s_{\mathbf{j}(2)}, \ldots, s_{\mathbf{j}(|\mathbf{I}|)}, 0, 0, 0, \ldots \text{ for the variables } x_1, x_2, x_3, \ldots \text{ in } \underbrace{x_{\mathbf{j}_1} x_{\mathbf{j}_2} \cdots x_{\mathbf{j}_s}}_{=\prod_{z\in\{1,2,\ldots,s\}} x_{\mathbf{j}_z}}\right)\\
&= \left(\text{the result of substituting } s_{\mathbf{j}(1)}, s_{\mathbf{j}(2)}, \ldots, s_{\mathbf{j}(|\mathbf{I}|)}, 0, 0, 0, \ldots \text{ for the variables } x_1, x_2, x_3, \ldots \text{ in } \prod_{z\in\{1,2,\ldots,s\}} x_{\mathbf{j}_z}\right)\\
&= \prod_{z\in\{1,2,\ldots,s\}} \left(\text{the result of substituting } s_{\mathbf{j}(1)}, s_{\mathbf{j}(2)}, \ldots, s_{\mathbf{j}(|\mathbf{I}|)}, 0, 0, 0, \ldots \text{ for the variables } x_1, x_2, x_3, \ldots \text{ in } x_{\mathbf{j}_z}\right)\\
&= \underbrace{\left(\text{the result of substituting } s_{\mathbf{j}(1)}, s_{\mathbf{j}(2)}, \ldots, s_{\mathbf{j}(|\mathbf{I}|)}, 0, 0, 0, \ldots \text{ for the variables } x_1, x_2, x_3, \ldots \text{ in } x_{\mathbf{j}_k}\right)}_{=0}\\
&\qquad \cdot \prod_{\substack{z\in\{1,2,\ldots,s\};\\ z\neq k}} \left(\text{the result of substituting } s_{\mathbf{j}(1)}, s_{\mathbf{j}(2)}, \ldots, s_{\mathbf{j}(|\mathbf{I}|)}, 0, 0, 0, \ldots \text{ for the variables } x_1, x_2, x_3, \ldots \text{ in } x_{\mathbf{j}_z}\right)\\
&\qquad \left(\text{here, we have split off the factor for } z = k \text{ from the product}\right)\\
&= 0.\end{aligned} \tag{13.173.13}$$

Now, let us forget that we fixed $(\mathbf{j}_1, \mathbf{j}_2, \ldots, \mathbf{j}_s) \in \{1, 2, 3, \ldots\}^s \setminus \mathbf{J}^s$. We thus have shown that (13.173.13) holds for every $(\mathbf{j}_1, \mathbf{j}_2, \ldots, \mathbf{j}_s) \in \{1, 2, 3, \ldots\}^s \setminus \mathbf{J}^s$.

But $\mathbf{J}^s \subset \{1, 2, 3, \ldots\}^s$ (since $\mathbf{J} \subset \{1, 2, 3, \ldots\}$). Hence, the set $\{1, 2, 3, \ldots\}^s$ is the union of its two disjoint subsets $\mathbf{J}^s$ and $\{1, 2, 3, \ldots\}^s \setminus \mathbf{J}^s$. Now, (13.173.11) becomes

[1154] Applying this to $R = \mathbf{k}[[\mathbf{x}]]$, $\mathbf{I} = \operatorname{SIS}(\ell)$ and $s_{\mathbf{i}} = \mathbf{x}_{\mathbf{i}}^{\alpha}$, we obtain

$$e_s^{\langle n\rangle}\left(\left(\mathbf{x}_{\mathbf{i}}^{\alpha}\right)_{\mathbf{i}\in\operatorname{SIS}(\ell)}\right) = e_s\left(\left(\left(\mathbf{x}_{\mathbf{i}}^{\alpha}\right)^n\right)_{\mathbf{i}\in\operatorname{SIS}(\ell)}\right).$$

Compared with (13.173.15), this yields $M_{\alpha\{n\}}^{\langle s\rangle} = e_s^{\langle n\rangle}\left(\left(\mathbf{x}_{\mathbf{i}}^{\alpha}\right)_{\mathbf{i}\in\operatorname{SIS}(\ell)}\right)$. This solves Exercise 6.5.4(c).

(d) The first sentence of Proposition 2.4.1 yields that the family $(e_1, e_2, e_3, \ldots)$ generates the $\mathbf{k}$-algebra Λ. Thus, every element of Λ can be written as a polynomial in the elements $e_1, e_2, e_3, \ldots$. Applying this to

---

$$\begin{aligned}
&e_s\left((s_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}}\right)\\
&= \sum_{\substack{(\mathbf{j}_1,\mathbf{j}_2,\ldots,\mathbf{j}_s)\in\{1,2,3,\ldots\}^s;\\ \mathbf{j}_1<\mathbf{j}_2<\cdots<\mathbf{j}_s}} \left(\text{the result of substituting } s_{\mathrm{j}(1)},\ s_{\mathrm{j}(2)},\ \ldots,\ s_{\mathrm{j}(|\mathbf{I}|)},\ 0,\ 0,\ 0,\ \ldots \text{ for the variables } x_1,\ x_2,\ x_3,\ \ldots \text{ in } x_{\mathbf{j}_1}x_{\mathbf{j}_2}\cdots x_{\mathbf{j}_s}\right)\\
&= \sum_{\substack{(\mathbf{j}_1,\mathbf{j}_2,\ldots,\mathbf{j}_s)\in\mathbf{J}^s;\\ \mathbf{j}_1<\mathbf{j}_2<\cdots<\mathbf{j}_s}} \underbrace{\left(\text{the result of substituting } s_{\mathrm{j}(1)},\ s_{\mathrm{j}(2)},\ \ldots,\ s_{\mathrm{j}(|\mathbf{I}|)},\ 0,\ 0,\ 0,\ \ldots \text{ for the variables } x_1,\ x_2,\ x_3,\ \ldots \text{ in } x_{\mathbf{j}_1}x_{\mathbf{j}_2}\cdots x_{\mathbf{j}_s}\right)}_{\substack{=s_{\mathrm{j}(\mathbf{j}_1)}s_{\mathrm{j}(\mathbf{j}_2)}\cdots s_{\mathrm{j}(\mathbf{j}_s)}\\ \text{(by (13.173.12))}}}\\
&\qquad + \sum_{\substack{(\mathbf{j}_1,\mathbf{j}_2,\ldots,\mathbf{j}_s)\in\{1,2,3,\ldots\}^s\setminus\mathbf{J}^s;\\ \mathbf{j}_1<\mathbf{j}_2<\cdots<\mathbf{j}_s}}\\
&\qquad\qquad \underbrace{\left(\text{the result of substituting } s_{\mathrm{j}(1)},\ s_{\mathrm{j}(2)},\ \ldots,\ s_{\mathrm{j}(|\mathbf{I}|)},\ 0,\ 0,\ 0,\ \ldots \text{ for the variables } x_1,\ x_2,\ x_3,\ \ldots \text{ in } x_{\mathbf{j}_1}x_{\mathbf{j}_2}\cdots x_{\mathbf{j}_s}\right)}_{\substack{=0\\ \text{(by (13.173.13))}}}\\
&\qquad \left(\text{since the set } \{1,2,3,\ldots\}^s \text{ is the union of its two disjoint subsets } \mathbf{J}^s \text{ and } \{1,2,3,\ldots\}^s\setminus\mathbf{J}^s\right)\\
&= \sum_{\substack{(\mathbf{j}_1,\mathbf{j}_2,\ldots,\mathbf{j}_s)\in\mathbf{J}^s;\\ \mathbf{j}_1<\mathbf{j}_2<\cdots<\mathbf{j}_s}} s_{\mathrm{j}(\mathbf{j}_1)}s_{\mathrm{j}(\mathbf{j}_2)}\cdots s_{\mathrm{j}(\mathbf{j}_s)} + \underbrace{\sum_{\substack{(\mathbf{j}_1,\mathbf{j}_2,\ldots,\mathbf{j}_s)\in\{1,2,3,\ldots\}^s\setminus\mathbf{J}^s;\\ \mathbf{j}_1<\mathbf{j}_2<\cdots<\mathbf{j}_s}} 0}_{=0}\\
&= \sum_{\substack{(\mathbf{j}_1,\mathbf{j}_2,\ldots,\mathbf{j}_s)\in\mathbf{J}^s;\\ \mathbf{j}_1<\mathbf{j}_2<\cdots<\mathbf{j}_s}} s_{\mathrm{j}(\mathbf{j}_1)}s_{\mathrm{j}(\mathbf{j}_2)}\cdots s_{\mathrm{j}(\mathbf{j}_s)}.
\end{aligned} \tag{13.173.14}$$

But $\mathbf{J}$ is a countable totally ordered set (since $\mathbf{J} = \{1, 2, \ldots, |\mathbf{I}|\}$). Also, j is a bijection $\{1, 2, \ldots, |\mathbf{I}|\} \to \mathbf{I}$. In other words, j is a bijection $\mathbf{J} \to \mathbf{I}$ (since $\mathbf{J} = \{1, 2, \ldots, |\mathbf{I}|\}$). Now, (13.173.14) becomes

$$e_s\left((s_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}}\right) = \sum_{\substack{(\mathbf{j}_1,\mathbf{j}_2,\ldots,\mathbf{j}_s)\in\mathbf{J}^s;\\ \mathbf{j}_1<\mathbf{j}_2<\cdots<\mathbf{j}_s}} s_{\mathrm{j}(\mathbf{j}_1)}s_{\mathrm{j}(\mathbf{j}_2)}\cdots s_{\mathrm{j}(\mathbf{j}_s)} = \sum_{\substack{(\mathbf{i}_1,\mathbf{i}_2,\ldots,\mathbf{i}_s)\in\mathbf{I}^s;\\ \mathbf{i}_1<\mathbf{i}_2<\cdots<\mathbf{i}_s}} s_{\mathbf{i}_1}s_{\mathbf{i}_2}\cdots s_{\mathbf{i}_s}$$

(by (13.173.5)). Thus, (13.173.7) is proven in Case 2.

We have thus proven (13.173.7) in each of the Cases 1 and 2. Since these two Cases are the only cases that can occur, we thus conclude that (13.173.7) holds, qed.

[1154] *Proof of (13.173.16):* Let R be a topological commutative $\mathbf{k}$-algebra. Let $\mathbf{I}$ be a countable set. Let $(s_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}} \in R^{\mathbf{I}}$ be a power-summable family of elements of R.

We must be in one of the following two cases:

Case 1: The set $\mathbf{I}$ is infinite.

Case 2: The set $\mathbf{I}$ is finite.

Let us first consider Case 1. In this case, the set $\mathbf{I}$ is infinite, and therefore countably infinite (since it is countable). Fix a bijection $\mathrm{j} : \{1, 2, 3, \ldots\} \to \mathbf{I}$ (such a bijection clearly exists, since $\mathbf{I}$ is countably infinite). Then, $e_s^{\langle n\rangle}\left((s_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}}\right)$ is the result of substituting $s_{\mathrm{j}(1)}, s_{\mathrm{j}(2)}, s_{\mathrm{j}(3)}, \ldots$ for the variables $x_1, x_2, x_3, \ldots$ in $e_s^{\langle n\rangle}$ (by the definition of $e_s^{\langle n\rangle}\left((s_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}}\right)$). Thus,

$$\begin{aligned}
&e_s^{\langle n\rangle}\left((s_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}}\right)\\
&= \left(\text{the result of substituting } s_{\mathrm{j}(1)},\ s_{\mathrm{j}(2)},\ s_{\mathrm{j}(3)},\ \ldots \text{ for the variables } x_1,\ x_2,\ x_3,\ \ldots \text{ in } \underbrace{e_s^{\langle n\rangle}}_{=\sum_{i_1<i_2<\cdots<i_s} x_{i_1}^n x_{i_2}^n\cdots x_{i_s}^n}\right)\\
&= \left(\text{the result of substituting } s_{\mathrm{j}(1)},\ s_{\mathrm{j}(2)},\ s_{\mathrm{j}(3)},\ \ldots \text{ for the variables } x_1,\ x_2,\ x_3,\ \ldots \text{ in } \sum_{i_1<i_2<\cdots<i_s} x_{i_1}^n x_{i_2}^n\cdots x_{i_s}^n\right)\\
&= \sum_{i_1<i_2<\cdots<i_s} s_{\mathrm{j}(\mathbf{j}_1)}^n s_{\mathrm{j}(\mathbf{j}_2)}^n\cdots s_{\mathrm{j}(\mathbf{j}_s)}^n.
\end{aligned} \tag{13.173.17}$$

the element $e_s^{\langle n\rangle}$ of Λ, we conclude that $e_s^{\langle n\rangle}$ can be written as a polynomial in the elements $e_1, e_2, e_3, \ldots$. In other words, there exists a polynomial $Q \in \mathbf{k}[z_1, z_2, z_3, \ldots]$ such that $e_s^{\langle n\rangle} = Q(e_1, e_2, e_3, \ldots)$. Consider this polynomial Q.

Let us now define a map $\Phi : \Lambda \to \mathbf{k}[[\mathbf{x}]]$ by

$$\Phi(f) = f\left(\left(\mathbf{x}_{\mathbf{i}}^{\alpha}\right)_{\mathbf{i}\in\operatorname{SIS}(\ell)}\right) \qquad \text{for every } f \in \Lambda$$

(where $f\left(\left(\mathbf{x}_{\mathbf{i}}^{\alpha}\right)_{\mathbf{i}\in\operatorname{SIS}(\ell)}\right)$ is defined as in Definition 6.5.1(b)). This map Φ is a $\mathbf{k}$-algebra homomorphism (since it amounts to a substitution of certain elements for the variables in a power series). Therefore, it commutes with polynomials, i.e., it satisfies

$$\Phi(R(f_1, f_2, f_3, \ldots)) = R(\Phi(f_1), \Phi(f_2), \Phi(f_3), \ldots)$$

On the other hand, $e_s\left(\left(\mathbf{s}_{\mathbf{i}}^n\right)_{\mathbf{i}\in\mathbf{I}}\right)$ is the result of substituting $s_{\mathrm{j}(1)}^n$, $s_{\mathrm{j}(2)}^n$, $s_{\mathrm{j}(3)}^n$, $\ldots$ for the variables x_1, x_2, x_3, $\ldots$ in e_s (by the definition of $e_s\left(\left(s_{\mathbf{i}}^n\right)_{\mathbf{i}\in\mathbf{I}}\right)$). Thus,

$$\begin{aligned}
&e_s\left(\left(\mathbf{s}_{\mathbf{i}}^n\right)_{\mathbf{i}\in\mathbf{I}}\right)\\
&= \left(\text{the result of substituting } s_{\mathrm{j}(1)}^n,\ s_{\mathrm{j}(2)}^n,\ s_{\mathrm{j}(3)}^n,\ \ldots \text{ for the variables } x_1,\ x_2,\ x_3,\ \ldots \text{ in } \underbrace{e_s}_{=\sum_{i_1<i_2<\cdots<i_s} x_{i_1}x_{i_2}\cdots x_{i_s}}\right)\\
&= \left(\text{the result of substituting } s_{\mathrm{j}(1)}^n,\ s_{\mathrm{j}(2)}^n,\ s_{\mathrm{j}(3)}^n,\ \ldots \text{ for the variables } x_1,\ x_2,\ x_3,\ \ldots \text{ in } \sum_{i_1<i_2<\cdots<i_s} x_{i_1}x_{i_2}\cdots x_{i_s}\right)\\
&= \sum_{i_1<i_2<\cdots<i_s} s_{\mathrm{j}(\mathbf{j}_1)}^n s_{\mathrm{j}(\mathbf{j}_2)}^n \cdots s_{\mathrm{j}(\mathbf{j}_s)}^n.
\end{aligned}$$

Compared with (13.173.17), this yields $e_s^{\langle n\rangle}\left((\mathbf{s}_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}}\right) = e_s\left(\left(\mathbf{s}_{\mathbf{i}}^n\right)_{\mathbf{i}\in\mathbf{I}}\right)$. Thus, (13.173.16) is proven in Case 1.

Let us now consider Case 2. In this case, the set $\mathbf{I}$ is finite. Fix a bijection $\mathrm{j} : \{1, 2, \ldots, |\mathbf{I}|\} \to \mathbf{I}$ (such a bijection clearly exists, since $\mathbf{I}$ is finite).

Define an infinite sequence $(t_1, t_2, t_3, \ldots)$ of elements of R by

$$(t_1, t_2, t_3, \ldots) = \left(s_{\mathrm{j}(1)}, s_{\mathrm{j}(2)}, \ldots, s_{\mathrm{j}(|\mathbf{I}|)}, 0, 0, 0, \ldots\right). \tag{13.173.18}$$

Then,

$$(t_1^n, t_2^n, t_3^n, \ldots) = \left(s_{\mathrm{j}(1)}^n, s_{\mathrm{j}(2)}^n, \ldots, s_{\mathrm{j}(|\mathbf{I}|)}^n, \underbrace{0^n, 0^n, 0^n, \ldots}_{\substack{=(0,0,0,\ldots)\\ (\text{since } 0^n=0 \text{ (since } n \text{ is positive))}}}\right) = \left(s_{\mathrm{j}(1)}^n, s_{\mathrm{j}(2)}^n, \ldots, s_{\mathrm{j}(|\mathbf{I}|)}^n, 0, 0, 0, \ldots\right). \tag{13.173.19}$$

Recall that $e_s^{\langle n\rangle}\left((\mathbf{s}_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}}\right)$ is the result of substituting $s_{\mathrm{j}(1)}$, $s_{\mathrm{j}(2)}$, $\ldots$, $s_{\mathrm{j}(|\mathbf{I}|)}$, 0, 0, 0, $\ldots$ for the variables x_1, x_2, x_3, $\ldots$ in $e_s^{\langle n\rangle}$ (by the definition of $e_s^{\langle n\rangle}\left((s_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}}\right)$). Thus,

$$\begin{aligned}
&e_s^{\langle n\rangle}\left((s_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}}\right)\\
&= \left(\text{the result of substituting } \underbrace{s_{\mathrm{j}(1)},\ s_{\mathrm{j}(2)},\ \ldots,\ s_{\mathrm{j}(|\mathbf{I}|)},\ 0,\ 0,\ 0,\ \ldots}_{\substack{=(t_1,t_2,t_3,\ldots)\\ (\text{by } (13.173.18))}} \text{ for the variables } x_1,\ x_2,\ x_3,\ \ldots \text{ in } e_s^{\langle n\rangle}\right)\\
&= \left(\text{the result of substituting } t_1,\ t_2,\ t_3,\ \ldots \text{ for the variables } x_1,\ x_2,\ x_3,\ \ldots \text{ in } \underbrace{e_s^{\langle n\rangle}}_{=\sum_{i_1<i_2<\cdots<i_s} x_{i_1}^n x_{i_2}^n\cdots x_{i_s}^n}\right)\\
&= \left(\text{the result of substituting } t_1,\ t_2,\ t_3,\ \ldots \text{ for the variables } x_1,\ x_2,\ x_3,\ \ldots \text{ in } \sum_{i_1<i_2<\cdots<i_s} x_{i_1}^n x_{i_2}^n\cdots x_{i_s}^n\right)\\
&= \sum_{i_1<i_2<\cdots<i_s} t_{i_1}^n t_{i_2}^n \cdots t_{i_s}^n. \qquad (13.173.20)
\end{aligned}$$

for every $f_1, f_2, f_3, \ldots \in \Lambda$ and every polynomial $R \in \mathbf{k}[z_1, z_2, z_3, \ldots]$. Applying this to $f_i = e_i$ and $R = Q$, we obtain

$$\Phi\left(Q\left(e_1, e_2, e_3, \ldots\right)\right) = Q\left(\Phi\left(e_1\right), \Phi\left(e_2\right), \Phi\left(e_3\right), \ldots\right). \tag{13.173.21}$$

However, for every $j \in \{1, 2, 3, \ldots\}$, we have

$$\begin{aligned}\Phi\left(e_j\right) &= e_j\left(\left(\mathbf{x}_{\mathbf{i}}^{\alpha}\right)_{\mathbf{i}\in\operatorname{SIS}(\ell)}\right) && \left(\text{by the definition of } \Phi\left(e_j\right)\right)\\ &= M_{\alpha}^{\langle j\rangle} && \left(\text{since } M_{\alpha}^{\langle j\rangle} = e_j\left(\left(\mathbf{x}_{\mathbf{i}}^{\alpha}\right)_{\mathbf{i}\in\operatorname{SIS}(\ell)}\right) \text{ (by the definition of } M_{\alpha}^{\langle j\rangle}\text{)}\right).\end{aligned}$$

Thus, $\left(\Phi\left(e_1\right), \Phi\left(e_2\right), \Phi\left(e_3\right), \ldots\right) = \left(M_{\alpha}^{\langle 1\rangle}, M_{\alpha}^{\langle 2\rangle}, M_{\alpha}^{\langle 3\rangle}, \ldots\right)$, so that

$$Q\left(\Phi\left(e_1\right), \Phi\left(e_2\right), \Phi\left(e_3\right), \ldots\right) = Q\left(M_{\alpha}^{\langle 1\rangle}, M_{\alpha}^{\langle 2\rangle}, M_{\alpha}^{\langle 3\rangle}, \ldots\right).$$

Hence, (13.173.21) becomes

$$\Phi\left(Q\left(e_1, e_2, e_3, \ldots\right)\right) = Q\left(\Phi\left(e_1\right), \Phi\left(e_2\right), \Phi\left(e_3\right), \ldots\right) = Q\left(M_{\alpha}^{\langle 1\rangle}, M_{\alpha}^{\langle 2\rangle}, M_{\alpha}^{\langle 3\rangle}, \ldots\right).$$

Compared with

$$\begin{aligned}\Phi\left(\underbrace{Q\left(e_1, e_2, e_3, \ldots\right)}_{\substack{= e_s^{\langle n\rangle} \\ \left(\text{since } e_s^{\langle n\rangle} = Q(e_1, e_2, e_3, \ldots)\right)}}\right) &= \Phi\left(e_s^{\langle n\rangle}\right) = e_s^{\langle n\rangle}\left(\left(\mathbf{x}_{\mathbf{i}}^{\alpha}\right)_{\mathbf{i}\in\operatorname{SIS}(\ell)}\right) && \left(\text{by the definition of } \Phi\left(e_s^{\langle n\rangle}\right)\right)\\ &= M_{\alpha\{n\}}^{\langle s\rangle} && \left(\text{by Exercise 6.5.4(c)}\right),\end{aligned}$$

this yields $M_{\alpha\{n\}}^{\langle s\rangle} = Q\left(M_{\alpha}^{\langle 1\rangle}, M_{\alpha}^{\langle 2\rangle}, M_{\alpha}^{\langle 3\rangle}, \ldots\right)$. Thus, there exists a polynomial $P \in \mathbf{k}[z_1, z_2, z_3, \ldots]$ such that $M_{\alpha\{n\}}^{\langle s\rangle} = P\left(M_{\alpha}^{\langle 1\rangle}, M_{\alpha}^{\langle 2\rangle}, M_{\alpha}^{\langle 3\rangle}, \ldots\right)$ (namely, $P = Q$). This solves Exercise 6.5.4(d).

On the other hand, $e_s\left(\left(\mathbf{s}_{\mathbf{i}}^n\right)_{\mathbf{i}\in\mathbf{I}}\right)$ is the result of substituting $s_{\mathbf{j}(1)}^n$, $s_{\mathbf{j}(2)}^n$, $\ldots$, $s_{\mathbf{j}(|\mathbf{I}|)}^n$, 0, 0, 0, $\ldots$ for the variables x_1, x_2, x_3, $\ldots$ in e_s (by the definition of $e_s\left(\left(s_{\mathbf{i}}^n\right)_{\mathbf{i}\in\mathbf{I}}\right)$). Thus,

$$\begin{aligned}&e_s\left(\left(\mathbf{s}_{\mathbf{i}}^n\right)_{\mathbf{i}\in\mathbf{I}}\right)\\ &= \left(\text{the result of substituting } \underbrace{s_{\mathbf{j}(1)}^n,\ s_{\mathbf{j}(2)}^n,\ \ldots,\ s_{\mathbf{j}(|\mathbf{I}|)}^n,\ 0,\ 0,\ 0,\ \ldots}_{\substack{=\left(t_1^n, t_2^n, t_3^n, \ldots\right) \\ \text{(by (13.173.19))}}} \text{ for the variables } x_1,\ x_2,\ x_3,\ \ldots \text{ in } e_s\right)\\ &= \left(\text{the result of substituting } t_1^n,\ t_2^n,\ t_3^n,\ \ldots \text{ for the variables } x_1,\ x_2,\ x_3,\ \ldots \text{ in } \underbrace{e_s}_{=\sum_{i_1<i_2<\cdots<i_s} x_{i_1}x_{i_2}\cdots x_{i_s}}\right)\\ &= \left(\text{the result of substituting } t_1^n,\ t_2^n,\ t_3^n,\ \ldots \text{ for the variables } x_1,\ x_2,\ x_3,\ \ldots \text{ in } \sum_{i_1<i_2<\cdots<i_s} x_{i_1}x_{i_2}\cdots x_{i_s}\right)\\ &= \sum_{i_1<i_2<\cdots<i_s} t_{i_1}^n t_{i_2}^n \cdots t_{i_s}^n.\end{aligned}$$

Compared with (13.173.20), this yields $e_s^{\langle n\rangle}\left(\left(\mathbf{s}_{\mathbf{i}}\right)_{\mathbf{i}\in\mathbf{I}}\right) = e_s\left(\left(\mathbf{s}_{\mathbf{i}}^n\right)_{\mathbf{i}\in\mathbf{I}}\right)$. Thus, (13.173.16) is proven in Case 2.

We have thus proven (13.173.16) in each of the Cases 1 and 2. Since these two Cases are the only cases that can occur, we thus conclude that (13.173.16) holds, qed.

13.174. **Solution to Exercise 6.5.5.** *Solution to Exercise 6.5.5.* The definition of $M_{(1)}^{\langle s\rangle}$ yields $M_{(1)}^{\langle s\rangle} = e_s\left(\left(\mathbf{x}_{\mathbf{i}}^{(1)}\right)_{\mathbf{i}\in\operatorname{SIS}(1)}\right)$.

But $\operatorname{SIS}(1)$ is defined as the set of all strictly increasing 1-tuples of positive integers. Clearly, such 1-tuples are in bijection with positive integers; the bijection sends a positive integer i to the strictly increasing 1-tuple (i). This bijection shows that the family $\left(\mathbf{x}_{\mathbf{i}}^{(1)}\right)_{\mathbf{i}\in\operatorname{SIS}(1)}$ is a reparametrization of the family $\left(\mathbf{x}_{(i)}^{(1)}\right)_{i\in\{1,2,3,\ldots\}}$. Thus, $e_s\left(\left(\mathbf{x}_{\mathbf{i}}^{(1)}\right)_{\mathbf{i}\in\operatorname{SIS}(1)}\right) = e_s\left(\left(\mathbf{x}_{(i)}^{(1)}\right)_{i\in\{1,2,3,\ldots\}}\right)$. Thus,

$$M_{(1)}^{\langle s\rangle} = e_s\left(\left(\mathbf{x}_{\mathbf{i}}^{(1)}\right)_{\mathbf{i}\in\operatorname{SIS}(1)}\right) = e_s\left(\left(\underbrace{\mathbf{x}_{(i)}^{(1)}}_{\substack{=x_i^1 \\ \text{(by the definition} \\ \text{of } \mathbf{x}_{(i)}^{(1)})}}\right)_{i\in\{1,2,3,\ldots\}}\right) = e_s\left(\left(\underbrace{x_i^1}_{=x_i}\right)_{i\in\{1,2,3,\ldots\}}\right)$$
$$= e_s\left((x_i)_{i\in\{1,2,3,\ldots\}}\right) = e_s.$$

This solves Exercise 6.5.5.

13.175. **Solution to Exercise 6.5.7.** *Solution to Exercise 6.5.7.*

Proof of Proposition 6.5.6. We first recall a general fact from algebra. Namely, if $\mathfrak{C}$ and $\mathfrak{D}$ are two commutative rings, if $\varphi : \mathfrak{C} \to \mathfrak{D}$ is a ring homomorphism, and if u and v are two nonnegative integers, then we can define a homomorphism $\varphi^{u\times v} : \mathfrak{C}^{u\times v} \to \mathfrak{D}^{u\times v}$ of additive groups by sending every matrix $(c_{i,j})_{i=1,2,\ldots,u;\ j=1,2,\ldots,v} \in \mathfrak{C}^{u\times v}$ to the matrix $(\varphi(c_{i,j}))_{i=1,2,\ldots,u;\ j=1,2,\ldots,v} \in \mathfrak{D}^{u\times v}$. This homomorphism $\varphi^{u\times v}$ is the map from $\mathfrak{C}^{u\times v}$ to $\mathfrak{D}^{u\times v}$ canonically induced by φ, and it has many structure-preserving properties (for instance, it respects the multiplication of matrices, in the sense that we have $\varphi^{u\times v}(X)\cdot\varphi^{v\times w}(Y) = \varphi^{u\times w}(XY)$ whenever $X \in \mathfrak{C}^{u\times v}$ and $Y \in \mathfrak{C}^{v\times w}$). We furthermore have

$$\det\left(\varphi^{u\times u}(X)\right) = \varphi(\det X) \tag{13.175.1}$$

for any two commutative rings $\mathfrak{C}$ and $\mathfrak{D}$, any ring homomorphism $\varphi : \mathfrak{C} \to \mathfrak{D}$, any nonnegative integer u and any matrix $X \in \mathfrak{C}^{u\times u}$. (This is because the determinant of a matrix is a polynomial in its entries, and polynomials commute with ring homomorphisms.)

We shall now construct a particular $\mathbf{k}$-algebra homomorphism $\Lambda \to \mathbf{k}[[\mathbf{x}]]$ which will help us in our proof. Namely, we define a map $\Phi : \Lambda \to \mathbf{k}[[\mathbf{x}]]$ by

$$\Phi(f) = f\left((\mathbf{x}_{\mathbf{i}}^{\alpha})_{\mathbf{i}\in\operatorname{SIS}(\ell)}\right) \qquad \text{for every } f \in \Lambda$$

(where $f\left((\mathbf{x}_{\mathbf{i}}^{\alpha})_{\mathbf{i}\in\operatorname{SIS}(\ell)}\right)$ is defined as in Definition 6.5.1(b)). This map Φ is a $\mathbf{k}$-algebra homomorphism (since it amounts to a substitution of certain elements for the variables in a power series). Every $s \in \mathbb{N}$ satisfies

$$\begin{aligned} \Phi(e_s) &= e_s\left((\mathbf{x}_{\mathbf{i}}^{\alpha})_{\mathbf{i}\in\operatorname{SIS}(\ell)}\right) && \text{(by the definition of } \Phi(e_s)) \\ &= M_{\alpha}^{\langle s\rangle} && \left(\text{since } M_{\alpha}^{\langle s\rangle} \text{ was defined as } e_s\left((\mathbf{x}_{\mathbf{i}}^{\alpha})_{\mathbf{i}\in\operatorname{SIS}(\ell)}\right)\right). \end{aligned} \tag{13.175.2}$$

Every positive integer s satisfies

$$\begin{aligned} \Phi(p_s) &= p_s\left((\mathbf{x}_{\mathbf{i}}^{\alpha})_{\mathbf{i}\in\operatorname{SIS}(\ell)}\right) && \text{(by the definition of } \Phi(p_s)) \\ &= M_{\alpha\{s\}} && \text{(by Exercise 6.5.4(a))}. \end{aligned} \tag{13.175.3}$$

(a) Define a matrix $A_n = (a_{i,j})_{i,j=1,2,\ldots,n}$ as in Exercise 2.9.13(a). Then, Exercise 2.9.13(a) yields $\det(A_n) = n!e_n$. Applying the map Φ to both sides of this equality, we obtain

$$\Phi(\det(A_n)) = \Phi(n!e_n) = n! \underbrace{\Phi(e_n)}_{\substack{=M_\alpha^{\langle n\rangle} \\ \text{(by (13.175.2), applied} \\ \text{to } s=n)}} = n!M_\alpha^{\langle n\rangle}.$$

But (13.175.1) (applied to $\mathfrak{C} = \Lambda$, $\mathfrak{D} = \mathbf{k}[[\mathbf{x}]]$, $\varphi = \Phi$, $u = n$ and $X = A_n$) yields

$$\det\left(\Phi^{n\times n}(A_n)\right) = \Phi(\det(A_n)) = n!M_\alpha^{\langle n\rangle}. \tag{13.175.4}$$

But

$$\Phi^{n\times n}\left(\underbrace{A_n}_{=(a_{i,j})_{i,j=1,2,\ldots,n}}\right) = \Phi^{n\times n}\left((a_{i,j})_{i,j=1,2,\ldots,n}\right) = (\Phi(a_{i,j}))_{i,j=1,2,\ldots,n} \tag{13.175.5}$$

(by the definition of $\Phi^{n\times n}\left((a_{i,j})_{i,j=1,2,\ldots,n}\right)$). But every $(i,j) \in \{1,2,\ldots,n\}^2$ satisfies

$$\begin{aligned}
\Phi(a_{i,j}) &= \Phi\left(\begin{cases} p_{i-j+1}, & \text{if } i \geq j; \\ i, & \text{if } i = j-1; \\ 0, & \text{if } i < j-1 \end{cases}\right) \qquad \left(\text{since } a_{i,j} = \begin{cases} p_{i-j+1}, & \text{if } i \geq j; \\ i, & \text{if } i = j-1; \\ 0, & \text{if } i < j-1 \end{cases}\right) \\
&= \begin{cases} \Phi(p_{i-j+1}), & \text{if } i \geq j; \\ \Phi(i), & \text{if } i = j-1; \\ \Phi(0), & \text{if } i < j-1 \end{cases} = \begin{cases} \Phi(p_{i-j+1}), & \text{if } i \geq j; \\ i, & \text{if } i = j-1; \\ 0, & \text{if } i < j-1 \end{cases} \\
&\qquad \left(\begin{array}{c} \text{since } \Phi(i) = i \text{ in the case when } i = j-1 \text{ (because } \Phi \text{ is a} \\ \mathbf{k}\text{-algebra homomorphism), and because } \Phi(0) = 0 \text{ in the} \\ \text{case when } i < j-1 \text{ (for the same reason)} \end{array}\right) \\
&= \begin{cases} M_{\alpha\{i-j+1\}}, & \text{if } i \geq j; \\ i, & \text{if } i = j-1; \\ 0, & \text{if } i < j-1 \end{cases} \\
&\qquad \left(\begin{array}{c} \text{since } \Phi(p_{i-j+1}) = M_{\alpha\{i-j+1\}} \text{ in the case when } i \geq j \\ \text{(by (13.175.3), applied to } s = i-j+1) \end{array}\right) \\
&= a_{i,j}^{\langle\alpha\rangle} \qquad \left(\text{since } a_{i,j}^{\langle\alpha\rangle} = \begin{cases} M_{\alpha\{i-j+1\}}, & \text{if } i \geq j; \\ i, & \text{if } i = j-1; \\ 0, & \text{if } i < j-1 \end{cases}\right).
\end{aligned}$$

Hence, $\left(\underbrace{\Phi(a_{i,j})}_{=a_{i,j}^{\langle\alpha\rangle}}\right)_{i,j=1,2,\ldots,n} = \left(a_{i,j}^{\langle\alpha\rangle}\right)_{i,j=1,2,\ldots,n} = A_n^{\langle\alpha\rangle}$. Thus, (13.175.5) becomes

$$\Phi^{n\times n}(A_n) = (\Phi(a_{i,j}))_{i,j=1,2,\ldots,n} = A_n^{\langle\alpha\rangle}.$$

Thus, (13.175.4) rewrites as $\det\left(A_n^{\langle\alpha\rangle}\right) = n!M_\alpha^{\langle n\rangle}$. This proves Proposition 6.5.6(a).

(b) The proof of Proposition 6.5.6(b) proceeds similarly to our proof of Proposition 6.5.6(a) above, as long as the obvious changes are made (one needs to consider B_n, $b_{i,j}$, $B_n^{\langle\alpha\rangle}$ and $b_{i,j}^{\langle\alpha\rangle}$ instead of A_n, $a_{i,j}$, $A_n^{\langle\alpha\rangle}$ and $a_{i,j}^{\langle\alpha\rangle}$). □

13.176. **Solution to Exercise 6.5.9.** *Solution to Exercise 6.5.9.*

Proof of Corollary 6.5.8. (a) We WLOG assume that $s \in \mathbb{N}$ (because otherwise, s is negative and thus satisfies $M_\alpha^{\langle s\rangle} = 0 \in \operatorname{QSym}$).

Define a matrix $A_s^{\langle\alpha\rangle} = \left(a_{i,j}^{\langle\alpha\rangle}\right)_{i,j=1,2,\ldots,s}$ as in Proposition 6.5.6(a) (but for s instead of n). Then, $A_s^{\langle\alpha\rangle} \in \operatorname{QSym}^{s\times s}$ (since every $(i,j) \in \{1,2,\ldots,s\}^2$ satisfies $a_{i,j}^{\langle\alpha\rangle} \in \operatorname{QSym}$ (as follows from the definition of $a_{i,j}^{\langle\alpha\rangle}$)). Hence, $\det\left(A_s^{\langle\alpha\rangle}\right) \in \operatorname{QSym}$. But Proposition 6.5.6(a) (applied to $n = s$) yields $\det\left(A_s^{\langle\alpha\rangle}\right) = s!M_\alpha^{\langle s\rangle}$. Hence, $s!M_\alpha^{\langle s\rangle} = \det\left(A_s^{\langle\alpha\rangle}\right) \in \operatorname{QSym}$.

Now, let us recall that there is a canonical ring homomorphism $\varphi : \mathbb{Z} \to \mathbf{k}$. This homomorphism gives rise to a ring homomorphism $\varphi[[\mathbf{x}]] : \mathbb{Z}[[\mathbf{x}]] \to \mathbf{k}[[\mathbf{x}]]$, and this latter homomorphism $\varphi[[\mathbf{x}]]$ sends $\operatorname{QSym}_{\mathbb{Z}}$ to $\operatorname{QSym}_{\mathbf{k}}$; that is, we have $(\varphi[[\mathbf{x}]])(\operatorname{QSym}_{\mathbb{Z}}) \subset \operatorname{QSym}_{\mathbf{k}}$. Moreover, it is clear that the ring homomorphism $\varphi[[\mathbf{x}]]$ sends the element $M_\alpha^{\langle s\rangle}$ of $\mathbb{Z}[[\mathbf{x}]]$ to the element $M_\alpha^{\langle s\rangle}$ of $\mathbf{k}[[\mathbf{x}]]$ (because the definition of $M_\alpha^{\langle s\rangle}$ is functorial in the base ring $\mathbf{k}$). Therefore, if we can prove that the element $M_\alpha^{\langle s\rangle}$ of $\mathbb{Z}[[\mathbf{x}]]$ belongs to $\operatorname{QSym}_{\mathbb{Z}}$, then it will automatically follow that the element $M_\alpha^{\langle s\rangle}$ of $\mathbf{k}[[\mathbf{x}]]$ belongs to $(\varphi[[\mathbf{x}]])(\operatorname{QSym}_{\mathbb{Z}}) \subset \operatorname{QSym}_{\mathbf{k}}$; this will complete the proof of Corollary 6.5.8(a). Hence, in order to prove Corollary 6.5.8(a), it only remains to prove that the element $M_\alpha^{\langle s\rangle}$ of $\mathbb{Z}[[\mathbf{x}]]$ belongs to $\operatorname{QSym}_{\mathbb{Z}}$. In other words, it only remains to prove Corollary 6.5.8(a) in the case of $\mathbf{k} = \mathbb{Z}$. Hence, in proving Corollary 6.5.8(a), we can WLOG assume that $\mathbf{k} = \mathbb{Z}$. Assume this. Since $\mathbf{k} = \mathbb{Z}$, we have $\operatorname{QSym} = \operatorname{QSym}_{\mathbb{Z}}$.

If a positive integer N and an element f of $\mathbb{Z}[[\mathbf{x}]]$ satisfy $Nf \in \operatorname{QSym}_{\mathbb{Z}}$, then f also lies in $\operatorname{QSym}_{\mathbb{Z}}$ (because N is not a zero-divisor in $\mathbf{k} = \mathbb{Z}$, and therefore f is obtained from the power series Nf by dividing all coefficients by N). Applying this to $N = s!$ and $f = M_\alpha^{\langle s\rangle}$, we obtain $M_\alpha^{\langle s\rangle} \in \operatorname{QSym}_{\mathbb{Z}}$ (since $s!M_\alpha^{\langle s\rangle} \in \operatorname{QSym} = \operatorname{QSym}_{\mathbb{Z}}$). In other words, $M_\alpha^{\langle s\rangle} \in \operatorname{QSym}$ (since $\mathbf{k} = \mathbb{Z}$). This completes the proof of Corollary 6.5.8(a).

(b) Recall that $M_\alpha^{\langle s\rangle} = e_s\left(\left(\mathbf{x}_{\mathbf{i}}^\alpha\right)_{\mathbf{i}\in\operatorname{SIS}(\ell)}\right)$. Thus, $M_\alpha^{\langle s\rangle}$ is a homogeneous power series of degree $s|\alpha|$ (since each $\mathbf{x}_{\mathbf{i}}^\alpha$ is a monomial of degree $|\alpha|$, and since e_s is a power series of degree s). Combined with $M_\alpha^{\langle s\rangle} \in \operatorname{QSym}$ (which follows from Corollary 6.5.8(a)), this yields $M_\alpha^{\langle s\rangle} \in \operatorname{QSym}_{s|\alpha|}$. This proves Corollary 6.5.8(b). □

13.177. **Solution to Exercise 6.5.12.** *Solution to Exercise 6.5.12.*

Proof of Remark 6.5.11. (a) Write the composition α in the form $(\alpha_1, \alpha_2, \ldots, \alpha_\ell)$. Then, $\operatorname{red}\alpha = \left(\frac{\alpha_1}{\gcd\alpha}, \frac{\alpha_2}{\gcd\alpha}, \ldots, \frac{\alpha_\ell}{\gcd\alpha}\right)$ (by the definition of $\operatorname{red}\alpha$). Hence, the definition of $(\operatorname{red}\alpha)\{\gcd\alpha\}$ yields

$$(\operatorname{red}\alpha)\{\gcd\alpha\} = \left((\gcd\alpha)\cdot\frac{\alpha_1}{\gcd\alpha}, (\gcd\alpha)\cdot\frac{\alpha_2}{\gcd\alpha}, \ldots, (\gcd\alpha)\cdot\frac{\alpha_\ell}{\gcd\alpha}\right) = (\alpha_1, \alpha_2, \ldots, \alpha_\ell) = \alpha,$$

so that Remark 6.5.11(a) is proven.

(c) Write the composition α in the form $(\alpha_1, \alpha_2, \ldots, \alpha_\ell)$. Then, $\operatorname{red}\alpha = \left(\frac{\alpha_1}{\gcd\alpha}, \frac{\alpha_2}{\gcd\alpha}, \ldots, \frac{\alpha_\ell}{\gcd\alpha}\right)$ (by the definition of $\operatorname{red}\alpha$). Hence, the definition of $\gcd(\operatorname{red}\alpha)$ yields

$$\begin{aligned}\gcd(\operatorname{red}\alpha) = \gcd\left(\frac{\alpha_1}{\gcd\alpha}, \frac{\alpha_2}{\gcd\alpha}, \ldots, \frac{\alpha_\ell}{\gcd\alpha}\right) &= \frac{\gcd(\alpha_1, \alpha_2, \ldots, \alpha_\ell)}{\gcd\alpha} = \frac{\gcd(\alpha_1, \alpha_2, \ldots, \alpha_\ell)}{\gcd(\alpha_1, \alpha_2, \ldots, \alpha_\ell)} \\ &\qquad \left(\text{since } \gcd\alpha = \gcd(\alpha_1, \alpha_2, \ldots, \alpha_\ell) \text{ (by the definition of } \gcd\alpha)\right) \\ &= 1.\end{aligned}$$

In other words, the composition $\operatorname{red}\alpha$ is reduced. This proves Remark 6.5.11(c).

(d) Assume that α is reduced. Write the composition α in the form $(\alpha_1, \alpha_2, \ldots, \alpha_\ell)$. Then, $\operatorname{red}\alpha = \left(\frac{\alpha_1}{\gcd\alpha}, \frac{\alpha_2}{\gcd\alpha}, \ldots, \frac{\alpha_\ell}{\gcd\alpha}\right)$ (by the definition of $\operatorname{red}\alpha$). But $\gcd\alpha = 1$ (since α is reduced). Now,

$$\begin{aligned}\operatorname{red}\alpha &= \left(\frac{\alpha_1}{\gcd\alpha}, \frac{\alpha_2}{\gcd\alpha}, \ldots, \frac{\alpha_\ell}{\gcd\alpha}\right) = \left(\frac{\alpha_1}{1}, \frac{\alpha_2}{1}, \ldots, \frac{\alpha_\ell}{1}\right) \qquad (\text{since } \gcd\alpha = 1)\\ &= (\alpha_1, \alpha_2, \ldots, \alpha_\ell) = \alpha.\end{aligned}$$

This proves Remark 6.5.11(d).

(e) Let $s \in \{1, 2, 3, \ldots\}$. Write the composition α in the form $(\alpha_1, \alpha_2, \ldots, \alpha_\ell)$. Then, $\operatorname{red}\alpha = \left(\frac{\alpha_1}{\gcd\alpha}, \frac{\alpha_2}{\gcd\alpha}, \ldots, \frac{\alpha_\ell}{\gcd\alpha}\right)$ (by the definition of $\operatorname{red}\alpha$). Also, $\alpha\{s\} = (s\alpha_1, s\alpha_2, \ldots, s\alpha_\ell)$ (by the definition of $\alpha\{s\}$). Now, $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_\ell)$, so that $\ell(\alpha) = \ell$. Hence, $\ell = \ell(\alpha) > 0$ (since α is nonempty). Since $\alpha\{s\} = (s\alpha_1, s\alpha_2, \ldots, s\alpha_\ell)$, we have $\ell(\alpha\{s\}) = \ell > 0$, so that the composition $\alpha\{s\}$ is nonempty.

By the definition of $\gcd\alpha$, we have $\gcd\alpha = \gcd(\alpha_1, \alpha_2, \ldots, \alpha_\ell)$. By the definition of $\gcd(\alpha\{s\})$, we have

$$\begin{aligned}\gcd(\alpha\{s\}) &= \gcd(s\alpha_1, s\alpha_2, \ldots, s\alpha_\ell) \qquad (\text{since } \alpha\{s\} = (s\alpha_1, s\alpha_2, \ldots, s\alpha_\ell))\\ &= s\underbrace{\gcd(\alpha_1, \alpha_2, \ldots, \alpha_\ell)}_{=\gcd\alpha} = s\gcd\alpha.\end{aligned}$$

Now, recall that $\alpha\{s\} = (s\alpha_1, s\alpha_2, \ldots, s\alpha_\ell)$. Hence, the definition of $\operatorname{red}(\alpha\{s\})$ yields

$$\begin{aligned}\operatorname{red}(\alpha\{s\}) &= \left(\frac{s\alpha_1}{\gcd(\alpha\{s\})}, \frac{s\alpha_2}{\gcd(\alpha\{s\})}, \ldots, \frac{s\alpha_\ell}{\gcd(\alpha\{s\})}\right)\\ &= \left(\frac{s\alpha_1}{s\gcd\alpha}, \frac{s\alpha_2}{s\gcd\alpha}, \ldots, \frac{s\alpha_\ell}{s\gcd\alpha}\right) \qquad (\text{since } \gcd(\alpha\{s\}) = s\gcd\alpha)\\ &= \left(\frac{\alpha_1}{\gcd\alpha}, \frac{\alpha_2}{\gcd\alpha}, \ldots, \frac{\alpha_\ell}{\gcd\alpha}\right) = \operatorname{red}\alpha.\end{aligned}$$

This proves Remark 6.5.11(e).

(f) Write the composition α in the form $(\alpha_1, \alpha_2, \ldots, \alpha_\ell)$. Then, $\operatorname{red}\alpha = \left(\frac{\alpha_1}{\gcd\alpha}, \frac{\alpha_2}{\gcd\alpha}, \ldots, \frac{\alpha_\ell}{\gcd\alpha}\right)$ (by the definition of $\operatorname{red}\alpha$). Thus,

$$|\operatorname{red}\alpha| = \frac{\alpha_1}{\gcd\alpha} + \frac{\alpha_2}{\gcd\alpha} + \cdots + \frac{\alpha_\ell}{\gcd\alpha} = \frac{\alpha_1 + \alpha_2 + \cdots + \alpha_\ell}{\gcd\alpha},$$

so that $(\gcd\alpha)|\operatorname{red}\alpha| = \alpha_1 + \alpha_2 + \cdots + \alpha_\ell = |\alpha|$ (since $|\alpha| = \alpha_1 + \alpha_2 + \cdots + \alpha_\ell$ (because $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_\ell)$)). This proves Remark 6.5.11(f).

(b) We have $(\operatorname{red}\alpha)\{\gcd\alpha\} = \alpha$ (by Remark 6.5.11(a)).

The property of a composition to be Lyndon does not change if all entries of the composition are multiplied by a fixed positive integer m (because this property only depends on the relative order of the parts of the composition). In other words, if β is a composition and m is a positive integer, then β is Lyndon if and only if $\beta\{m\}$ is Lyndon. Applying this to $\beta = \operatorname{red}\alpha$ and $m = \gcd\alpha$, we conclude that $\operatorname{red}\alpha$ is Lyndon if and only if $(\operatorname{red}\alpha)\{\gcd\alpha\}$ is Lyndon. In other words, $\operatorname{red}\alpha$ is Lyndon if and only if α is Lyndon (because $(\operatorname{red}\alpha)\{\gcd\alpha\} = \alpha$). This proves Remark 6.5.11(b). □

13.178. **Solution to Exercise 6.5.15.** *Solution to Exercise 6.5.15.*

Proof of Lemma 6.5.14. For every $\alpha \in \mathfrak{L}$, the pair $(\operatorname{red}\alpha, \gcd\alpha)$ is a well-defined element of $\mathfrak{RL} \times \{1, 2, 3, \ldots\}$ [1155]. Hence, we can define a map $\mathbf{R} : \mathfrak{L} \to \mathfrak{RL} \times \{1, 2, 3, \ldots\}$ by

$$(\mathbf{R}(\alpha) = (\operatorname{red}\alpha, \gcd\alpha) \qquad \text{for every } \alpha \in \mathfrak{L}).$$

Consider this $\mathbf{R}$.

For every $(w, s) \in \mathfrak{RL} \times \{1, 2, 3, \ldots\}$, the element $w\{s\}$ is a well-defined element of $\mathfrak{L}$ [1156]. Hence, we can define a map $\mathbf{M} : \mathfrak{RL} \times \{1, 2, 3, \ldots\} \to \mathfrak{L}$ by

$$(\mathbf{M}(w, s) = w\{s\} \qquad \text{for every } (w, s) \in \mathfrak{RL} \times \{1, 2, 3, \ldots\}).$$

Consider this $\mathbf{M}$.

It is now easy to see that $\mathbf{R} \circ \mathbf{M} = \operatorname{id}$ [1157] and $\mathbf{M} \circ \mathbf{R} = \operatorname{id}$ [1158]. Therefore, the maps $\mathbf{M}$ and $\mathbf{R}$ are mutually inverse. Hence, the map $\mathbf{M}$ is a bijection. Hence, the family $\left(M_{\operatorname{red}(\mathbf{M}(w,s))}^{\langle\gcd(\mathbf{M}(w,s))\rangle}\right)_{(w,s)\in\mathfrak{RL}\times\{1,2,3,\ldots\}}$ is a reindexing of the family $\left(M_{\operatorname{red}\alpha}^{\langle\gcd\alpha\rangle}\right)_{\alpha\in\mathfrak{L}}$. Since every $(w, s) \in \mathfrak{RL} \times \{1, 2, 3, \ldots\}$ satisfies

$$\gcd(\mathbf{M}(w, s)) = s \qquad \text{and} \qquad \operatorname{red}(\mathbf{M}(w, s)) = w$$

[1155] *Proof.* Let $\alpha \in \mathfrak{L}$. Then, α is a Lyndon word (since $\mathfrak{L}$ is the set of all Lyndon words), and thus nonempty. Hence, α is a nonempty composition, so that $\operatorname{red}\alpha$ is a well-defined composition, and $\gcd\alpha$ is a well-defined positive integer. Remark 6.5.11(b) yields that the composition α is Lyndon if and only if the composition $\operatorname{red}\alpha$ is Lyndon. Since α is Lyndon, this yields that $\operatorname{red}\alpha$ is Lyndon. Also, $\operatorname{red}\alpha$ is reduced (by Remark 6.5.11(c)). Thus, $\operatorname{red}\alpha$ is a reduced Lyndon composition. In other words, $\operatorname{red}\alpha \in \mathfrak{RL}$ (since $\mathfrak{RL}$ is the set of all reduced Lyndon compositions). Combined with $\gcd\alpha \in \{1, 2, 3, \ldots\}$ (since $\gcd\alpha$ is a well-defined positive integer), this yields $(\operatorname{red}\alpha, \gcd\alpha) \in \mathfrak{RL} \times \{1, 2, 3, \ldots\}$, qed.

[1156] *Proof.* Let $(w, s) \in \mathfrak{RL} \times \{1, 2, 3, \ldots\}$. Then, $w \in \mathfrak{RL}$ and $s \in \{1, 2, 3, \ldots\}$. Since $\mathfrak{RL}$ is the set of all reduced Lyndon compositions, we see that w is a reduced Lyndon composition (since $w \in \mathfrak{RL}$). Thus, w is nonempty (since w is Lyndon).

We shall now show that $w\{s\} \in \mathfrak{L}$.

Remark 6.5.11(e) (applied to $\alpha = w$) yields that the composition $w\{s\}$ is nonempty and satisfies $\operatorname{red}(w\{s\}) = \operatorname{red}w$ and $\gcd(w\{s\}) = s\gcd w$. We have $\operatorname{red}(w\{s\}) = \operatorname{red}w = w$ (by Remark 6.5.11(d), applied to $\alpha = w$). Now, recall that the composition w is Lyndon. Hence, the composition $\operatorname{red}(w\{s\})$ is Lyndon (since $\operatorname{red}(w\{s\}) = w$).

Remark 6.5.11(b) (applied to $\alpha = w\{s\}$) yields that the composition $w\{s\}$ is Lyndon if and only if the composition $\operatorname{red}(w\{s\})$ is Lyndon. Since the composition $\operatorname{red}(w\{s\})$ is Lyndon, this yields that the composition $w\{s\}$ is Lyndon. In other words, $w\{s\} \in \mathfrak{L}$ (since $\mathfrak{L}$ is the set of all Lyndon words), qed.

[1157] *Proof.* Let $(w, s) \in \mathfrak{RL} \times \{1, 2, 3, \ldots\}$. Thus, $w \in \mathfrak{RL}$ and $s \in \{1, 2, 3, \ldots\}$. Since $\mathfrak{RL}$ is the set of all reduced Lyndon compositions, we see that w is a reduced Lyndon composition (since $w \in \mathfrak{RL}$). Thus, w is nonempty (since w is Lyndon).

Remark 6.5.11(e) (applied to $\alpha = w$) yields that the composition $w\{s\}$ is nonempty and satisfies $\operatorname{red}(w\{s\}) = \operatorname{red}w$ and $\gcd(w\{s\}) = s\gcd w$. But $\gcd w = 1$ (since w is reduced), so that $\gcd(w\{s\}) = s\underbrace{\gcd w}_{=1} = s$. Also, $\operatorname{red}(w\{s\}) = \operatorname{red}w = w$ (by Remark 6.5.11(d), applied to $\alpha = w$). Now,

$$\begin{aligned}(\mathbf{R} \circ \mathbf{M})(w, s) &= \mathbf{R}\Bigg(\underbrace{\mathbf{M}(w, s)}_{\substack{=w\{s\} \\ \text{(by the definition of } \mathbf{M}(w,s))}}\Bigg) = \mathbf{R}(w\{s\}) = \Bigg(\underbrace{\operatorname{red}(w\{s\})}_{=w}, \underbrace{\gcd(w\{s\})}_{=s}\Bigg) \\ &\qquad \text{(by the definition of } \mathbf{R}(w\{s\})) \\ &= (w, s) = \operatorname{id}(w, s).\end{aligned}$$

Now, let us forget that we fixed (w, s). We thus have shown that $(\mathbf{R} \circ \mathbf{M})(w, s) = \operatorname{id}(w, s)$ for every $(w, s) \in \mathfrak{RL} \times \{1, 2, 3, \ldots\}$. In other words, $\mathbf{R} \circ \mathbf{M} = \operatorname{id}$, qed.

[1158] *Proof.* Let $\alpha \in \mathfrak{L}$. Then, α is a Lyndon word (since $\mathfrak{L}$ is the set of all Lyndon words), and thus nonempty. Hence, α is a nonempty composition. Now,

$$\begin{aligned}(\mathbf{M} \circ \mathbf{R})(\alpha) &= \mathbf{M}\Bigg(\underbrace{\mathbf{R}(\alpha)}_{\substack{=(\operatorname{red}\alpha, \gcd\alpha) \\ \text{(by the definition of } \mathbf{R}(\alpha))}}\Bigg) = \mathbf{M}(\operatorname{red}\alpha, \gcd\alpha) = (\operatorname{red}\alpha)\{\gcd\alpha\} \qquad \text{(by the definition of } \mathbf{M}(\operatorname{red}\alpha, \gcd\alpha)) \\ &= \alpha \qquad \text{(by Remark 6.5.11(a))} \\ &= \operatorname{id}(\alpha).\end{aligned}$$

Now, let us forget that we fixed α. We thus have proven that $(\mathbf{M} \circ \mathbf{R})(\alpha) = \operatorname{id}(\alpha)$ for every $\alpha \in \mathfrak{L}$. Thus, $\mathbf{M} \circ \mathbf{R} = \operatorname{id}$, qed.

[1159], this rewrites as follows: The family $\left(M_w^{\langle s\rangle}\right)_{(w,s)\in\mathfrak{RL}\times\{1,2,3,\ldots\}}$ is a reindexing of the family $\left(M_{\operatorname{red}\alpha}^{\langle\gcd\alpha\rangle}\right)_{\alpha\in\mathfrak{L}}$. This proves Lemma 6.5.14. □

13.179. **Solution to Exercise 6.5.17.** *Solution to Exercise 6.5.17.* We shall give two proofs of Lemma 6.5.16.

First proof of Lemma 6.5.16. We first introduce some notation.

For every $m\in\mathbb{Z}$, let $\mathcal{F}_m$ denote the $\mathbf{k}$-submodule of QSym spanned by $(M_\beta)_{\beta\in\operatorname{Comp};\ \ell(\beta)\leq m}$. Then, $1\in\mathcal{F}_0$ [1160]. Furthermore,

$$\mathcal{F}_u\mathcal{F}_v\subset\mathcal{F}_{u+v}\qquad\text{for every } u\in\mathbb{Z}\text{ and } v\in\mathbb{Z}\tag{13.179.1}$$

[1161]. It is now easy to see that

$$(\mathcal{F}_m)^k\subset\mathcal{F}_{km}\qquad\text{for every } m\in\mathbb{Z}\text{ and } k\in\mathbb{N}\tag{13.179.2}$$

(where $(\mathcal{F}_m)^k$ means $\underbrace{\mathcal{F}_m\mathcal{F}_m\cdots\mathcal{F}_m}_{k\text{ times}}$) [1162].

Write the composition α in the form $\alpha=(\alpha_1,\alpha_2,\ldots,\alpha_\ell)$; then, $\ell=\ell(\alpha)$.

[1159] *Proof.* We have $\mathbf{R}\circ\mathbf{M}=\operatorname{id}$, thus $\underbrace{(\mathbf{R}\circ\mathbf{M})}_{=\operatorname{id}}(w,s)=\operatorname{id}(w,s)=(w,s)$ and therefore

$$(w,s)=(\mathbf{R}\circ\mathbf{M})(w,s)=\mathbf{R}(\mathbf{M}(w,s))=(\operatorname{red}(\mathbf{M}(w,s)),\gcd(\mathbf{M}(w,s)))\qquad(\text{by the definition of }\mathbf{R}(\mathbf{M}(w,s))).$$

Hence, $(\operatorname{red}(\mathbf{M}(w,s)),\gcd(\mathbf{M}(w,s)))=(w,s)$. In other words, $\operatorname{red}(\mathbf{M}(w,s))=w$ and $\gcd(\mathbf{M}(w,s))=s$, qed.

[1160] *Proof.* We know that $\mathcal{F}_0$ is the $\mathbf{k}$-submodule of QSym spanned by $(M_\beta)_{\beta\in\operatorname{Comp};\ \ell(\beta)\leq 0}$ (by the definition of $\mathcal{F}_0$). Hence, in particular, $M_\beta\in\mathcal{F}_0$ for every $\beta\in\operatorname{Comp}$ satisfying $\ell(\beta)\leq 0$. Applying this to $\beta=\varnothing$, we obtain $M_\varnothing\in\mathcal{F}_0$ (since $\ell(\varnothing)=0$). Since $M_\varnothing=1$, this rewrites as $1\in\mathcal{F}_0$, qed.

[1161] *Proof of (13.179.1):* Let $u\in\mathbb{Z}$ and $v\in\mathbb{Z}$. Let $p\in\mathcal{F}_u$ and $q\in\mathcal{F}_v$. We are going to prove that $pq\in\mathcal{F}_{u+v}$.

Notice that $\mathcal{F}_{u+v}$ is a $\mathbf{k}$-submodule of QSym. Hence, the claim that $pq\in\mathcal{F}_{u+v}$ is $\mathbf{k}$-linear in p.

We know that p belongs to $\mathcal{F}_u$. In other words, p is a $\mathbf{k}$-linear combination of the elements $(M_\beta)_{\beta\in\operatorname{Comp};\ \ell(\beta)\leq u}$ (since $\mathcal{F}_u$ is the $\mathbf{k}$-submodule of QSym spanned by $(M_\beta)_{\beta\in\operatorname{Comp};\ \ell(\beta)\leq u}$). Hence, we can WLOG assume that p is one of the elements $(M_\beta)_{\beta\in\operatorname{Comp};\ \ell(\beta)\leq u}$ (because the claim that we are proving – namely, the claim that $pq\in\mathcal{F}_{u+v}$ – is $\mathbf{k}$-linear in p). Assume this. Similarly, assume that q is one of the elements $(M_\beta)_{\beta\in\operatorname{Comp};\ \ell(\beta)\leq v}$.

There exists a $\varphi\in\operatorname{Comp}$ satisfying $\ell(\varphi)\leq u$ and $p=M_\varphi$ (since p is one of the elements $(M_\beta)_{\beta\in\operatorname{Comp};\ \ell(\beta)\leq u}$). Similarly, there exists a $\psi\in\operatorname{Comp}$ satisfying $\ell(\psi)\leq v$ and $q=M_\psi$. Consider these φ and ψ. Corollary 6.4.7 (applied to φ and ψ instead of α and β) shows that $M_\varphi M_\psi$ is a sum of terms of the form M_δ with $\delta\in\mathfrak{A}^*$ satisfying $\ell(\delta)\leq\ell(\varphi)+\ell(\psi)$. Since every $\delta\in\mathfrak{A}^*$ satisfying $\ell(\delta)\leq\ell(\varphi)+\ell(\psi)$ also satisfies $\ell(\delta)\leq u+v$ (because $\underbrace{\ell(\varphi)}_{\leq u}+\underbrace{\ell(\psi)}_{\leq v}\leq u+v$), this yields that $M_\varphi M_\psi$ is a sum of terms of the form M_δ with $\delta\in\mathfrak{A}^*$ satisfying $\ell(\delta)\leq u+v$. In particular, $M_\varphi M_\psi$ is a $\mathbf{k}$-linear combination of $(M_\delta)_{\delta\in\mathfrak{A}^*;\ \ell(\delta)\leq u+v}$. In other words, $M_\varphi M_\psi$ is a $\mathbf{k}$-linear combination of $(M_\delta)_{\delta\in\operatorname{Comp};\ \ell(\delta)\leq u+v}$ (since $\mathfrak{A}^*=\operatorname{Comp}$). In other words, $M_\varphi M_\psi$ is a $\mathbf{k}$-linear combination of $(M_\beta)_{\beta\in\operatorname{Comp};\ \ell(\beta)\leq u+v}$ (here, we renamed the index δ as β). In other words, $M_\varphi M_\psi$ belongs to the $\mathbf{k}$-submodule of QSym spanned by $(M_\beta)_{\beta\in\operatorname{Comp};\ \ell(\beta)\leq u+v}$. In other words, $M_\varphi M_\psi$ belongs to $\mathcal{F}_{u+v}$ (since $\mathcal{F}_{u+v}$ was defined as the $\mathbf{k}$-submodule of QSym spanned by $(M_\beta)_{\beta\in\operatorname{Comp};\ \ell(\beta)\leq u+v}$). Hence, $M_\varphi M_\psi\in\mathcal{F}_{u+v}$, so that $\underbrace{p}_{=M_\varphi}\underbrace{q}_{=M_\psi}=M_\varphi M_\psi\in\mathcal{F}_{u+v}$.

Now, let us forget that we fixed p and q. We thus have proven that $pq\in\mathcal{F}_{u+v}$ for any $p\in\mathcal{F}_u$ and $q\in\mathcal{F}_v$. Since $\mathcal{F}_{u+v}$ is a $\mathbf{k}$-submodule of QSym, this yields that $\mathcal{F}_u\mathcal{F}_v\subset\mathcal{F}_{u+v}$.

[1162] *Proof of (13.179.2):* Let $m\in\mathbb{Z}$. We shall prove that (13.179.2) holds for every $k\in\mathbb{N}$. Indeed, we shall prove this by induction over k:

Induction base: We have

$$\begin{aligned}(\mathcal{F}_m)^0&=\mathbf{k}\cdot\underbrace{1}_{\in\mathcal{F}_0}\subset\mathbf{k}\cdot\mathcal{F}_0\subset\mathcal{F}_0\qquad(\text{since }\mathcal{F}_0\text{ is a }\mathbf{k}\text{-module})\\&=\mathcal{F}_{0m}\qquad(\text{since }0=0m).\end{aligned}$$

In other words, (13.179.2) holds for $k=0$. This completes the induction base.

Induction step: Let $K\in\mathbb{N}$. Assume that (13.179.2) holds for $k=K$. We must then show that (13.179.2) holds for $k=K+1$.

Define $A_s^{\langle\alpha\rangle}$ and $a_{i,j}^{\langle\alpha\rangle}$ as in Proposition 6.5.6(a) (but with s instead of n). Then,

$$a_{i,j}^{\langle\alpha\rangle} = \begin{cases} M_{\alpha\{i-j+1\}}, & \text{if } i \geq j; \\ i, & \text{if } i = j-1; \\ 0, & \text{if } i < j-1 \end{cases} \qquad \text{for all } (i,j) \in \{1,2,\ldots,s\}^2 ,$$

and we have $A_s^{\langle\alpha\rangle} = \left(a_{i,j}^{\langle\alpha\rangle}\right)_{i,j=1,2,\ldots,s}$. Proposition 6.5.6(a) (applied to s instead of n) yields $\det\left(A_s^{\langle\alpha\rangle}\right) = s! M_\alpha^{\langle s\rangle}$. Since $A_s^{\langle\alpha\rangle} = \left(a_{i,j}^{\langle\alpha\rangle}\right)_{i,j=1,2,\ldots,s}$, this rewrites as

$$\det\left(\left(a_{i,j}^{\langle\alpha\rangle}\right)_{i,j=1,2,\ldots,s}\right) = s! M_\alpha^{\langle s\rangle}. \tag{13.179.3}$$

On the other hand, it is easy to see that

$$a_{i,j}^{\langle\alpha\rangle} \in \mathbf{k} \cdot 1 \qquad \text{for every } (i,j) \in \{1,2,\ldots,s\}^2 \text{ satisfying } i < j. \tag{13.179.4}$$

[1163] Furthermore,

$$a_{i,j}^{\langle\alpha\rangle} \in \mathcal{F}_\ell \qquad \text{for every } (i,j) \in \{1,2,\ldots,s\}^2 . \tag{13.179.5}$$

We have $(\mathcal{F}_m)^K \subset \mathcal{F}_{Km}$ (since (13.179.2) holds for $k = K$). Thus,

$$\begin{aligned} (\mathcal{F}_m)^{K+1} &= \underbrace{(\mathcal{F}_m)^K}_{\subset \mathcal{F}_{Km}} \cdot \mathcal{F}_m \subset \mathcal{F}_{Km} \cdot \mathcal{F}_m \subset \mathcal{F}_{Km+m} && \text{(by (13.179.1), applied to } u = Km \text{ and } v = m) \\ &= \mathcal{F}_{(K+1)m} && \text{(since } Km + m = (K+1)\,m) . \end{aligned}$$

In other words, (13.179.2) holds for $k = K+1$. This completes the induction step. The proof of (13.179.2) is thus complete.

[1163] *Proof of (13.179.4):* Let $(i,j) \in \{1,2,\ldots,s\}^2$ be such that $i < j$. We need to prove that $a_{i,j}^{\langle\alpha\rangle} \in \mathbf{k} \cdot 1$. This is clear in the case when $i = j-1$ (because in this case, we have

$$a_{i,j}^{\langle\alpha\rangle} = \begin{cases} M_{\alpha\{i-j+1\}}, & \text{if } i \geq j; \\ i, & \text{if } i = j-1; \\ 0, & \text{if } i < j-1 \end{cases} = i \qquad (\text{since } i = j-1)$$
$$\in \mathbf{k} \cdot 1$$

). Hence, for the rest of this proof, we can WLOG assume that $i = j-1$. Assume this. Thus, $i \neq j-1$.

We have $i < j$. Thus, $i \leq j-1$ (since i and j are integers), so that $i < j-1$ (since $i \neq j-1$). Now,

$$a_{i,j}^{\langle\alpha\rangle} = \begin{cases} M_{\alpha\{i-j+1\}}, & \text{if } i \geq j; \\ i, & \text{if } i = j-1; \\ 0, & \text{if } i < j-1 \end{cases} = 0 \qquad (\text{since } i < j-1)$$
$$\in \mathbf{k} \cdot 1,$$

and thus (13.179.4) is proven.

[1164] Finally,

$$a_{i,i}^{\langle\alpha\rangle} = M_\alpha \qquad \text{for every } i \in \{1, 2, \ldots, s\}. \tag{13.179.6}$$

[1165]

But recall that every commutative ring A, every $m \in \mathbb{N}$ and every matrix $(u_{i,j})_{i,j=1,2,\ldots,m} \in A^{m\times m}$ satisfy

$$\det\left((u_{i,j})_{i,j=1,2,\ldots,m}\right) = \sum_{\sigma\in\mathfrak{S}_m} (-1)^\sigma \prod_{i=1}^m u_{i,\sigma(i)}.$$

[1166] Applying this equality to $A = \operatorname{QSym}$, $m = s$ and $u_{i,j} = a_{i,j}^{\langle\alpha\rangle}$, we obtain

$$\begin{aligned}
\det\left(\left(a_{i,j}^{\langle\alpha\rangle}\right)_{i,j=1,2,\ldots,s}\right) &= \sum_{\sigma\in\mathfrak{S}_s} (-1)^\sigma \prod_{i=1}^s a_{i,\sigma(i)}^{\langle\alpha\rangle} \\
&= \underbrace{(-1)^{\operatorname{id}}}_{=1} \prod_{i=1}^s \underbrace{a_{i,\operatorname{id}(i)}^{\langle\alpha\rangle}}_{\substack{=a_{i,i}^{\langle\alpha\rangle}=M_\alpha \\ \text{(by (13.179.6))}}} + \sum_{\substack{\sigma\in\mathfrak{S}_s;\\ \sigma\neq\operatorname{id}}} (-1)^\sigma \prod_{i=1}^s a_{i,\sigma(i)}^{\langle\alpha\rangle} \\
&\qquad \left(\text{here, we have split off the addend for } \sigma = \operatorname{id} \text{ from the sum}\right) \\
&= \underbrace{\prod_{i=1}^s M_\alpha}_{=M_\alpha^s} + \sum_{\substack{\sigma\in\mathfrak{S}_s;\\ \sigma\neq\operatorname{id}}} (-1)^\sigma \prod_{i=1}^s a_{i,\sigma(i)}^{\langle\alpha\rangle} = M_\alpha^s + \sum_{\substack{\sigma\in\mathfrak{S}_s;\\ \sigma\neq\operatorname{id}}} (-1)^\sigma \prod_{i=1}^s a_{i,\sigma(i)}^{\langle\alpha\rangle}.
\end{aligned}$$

Compared with (13.179.3), this yields

$$s! M_\alpha^{\langle s\rangle} = M_\alpha^s + \sum_{\substack{\sigma\in\mathfrak{S}_s;\\ \sigma\neq\operatorname{id}}} (-1)^\sigma \prod_{i=1}^s a_{i,\sigma(i)}^{\langle\alpha\rangle}.$$

[1164] *Proof of (13.179.5):* Let $(i, j) \in \{1, 2, \ldots, s\}^2$. We need to show that $a_{i,j}^{\langle\alpha\rangle} \in \mathcal{F}_\ell$.

First, recall that $\mathcal{F}_\ell$ is the $\mathbf{k}$-submodule of QSym spanned by $(M_\beta)_{\beta\in\operatorname{Comp};\ \ell(\beta)\leq\ell}$ (by the definition of $\mathcal{F}_\ell$). In particular, $M_\beta \in \mathcal{F}_\ell$ for every $\beta \in \operatorname{Comp}$ satisfying $\ell(\beta) \leq \ell$. Applied to $\beta = \varnothing$, this yields that $M_\varnothing \in \mathcal{F}_\ell$ (since $\varnothing \in \operatorname{Comp}$ satisfies $\ell(\varnothing) = 0 \leq \ell(\alpha) = \ell$). Since $M_\varnothing = 1$, this rewrites as $1 \in \mathcal{F}_\ell$, whence $\mathbf{k} \cdot \underbrace{1}_{\in\mathcal{F}_\ell} \subset \mathbf{k} \cdot \mathcal{F}_\ell \subset \mathcal{F}_\ell$ (since $\mathcal{F}_\ell$ is a $\mathbf{k}$-module).

Now, if $i < j$, then (13.179.4) yields $a_{i,j}^{\langle\alpha\rangle} \in \mathbf{k} \cdot 1 \subset \mathcal{F}_\ell$. Hence, (13.179.5) is proven in the case when $i < j$. For the rest of our proof of (13.179.5), we can thus WLOG assume that we don't have $i < j$. Assume this.

We have $i \geq j$ (since we don't have $i < j$) and $\ell(\alpha\{i - j + 1\}) = \ell(\alpha)$ (since $\ell(\alpha\{k\}) = \ell(\alpha)$ for every positive integer k). Now, recall that $M_\beta \in \mathcal{F}_\ell$ for every $\beta \in \operatorname{Comp}$ satisfying $\ell(\beta) \leq \ell$. Applied to $\beta = \alpha\{i - j + 1\}$, this yields that $M_{\alpha\{i-j+1\}} \in \mathcal{F}_\ell$ (since $\ell(\alpha\{i - j + 1\}) = \ell(\alpha) = \ell$). Now,

$$\begin{aligned}
a_{i,j}^{\langle\alpha\rangle} &= \begin{cases} M_{\alpha\{i-j+1\}}, & \text{if } i \geq j; \\ i, & \text{if } i = j - 1; \\ 0, & \text{if } i < j - 1 \end{cases} = M_{\alpha\{i-j+1\}} \qquad (\text{since } i \geq j) \\
&\in \mathcal{F}_\ell,
\end{aligned}$$

and this proves (13.179.5).

[1165] *Proof of (13.179.6):* Let $i \in \{1, 2, \ldots, s\}$. Then, the definition of $a_{i,i}^{\langle\alpha\rangle}$ yields

$$\begin{aligned}
a_{i,i}^{\langle\alpha\rangle} &= \begin{cases} M_{\alpha\{i-i+1\}}, & \text{if } i \geq i; \\ i, & \text{if } i = i - 1; \\ 0, & \text{if } i < i - 1 \end{cases} = M_{\alpha\{i-i+1\}} \qquad (\text{since } i \geq i) \\
&= M_\alpha \qquad \left(\text{since } \alpha\left\{\underbrace{i - i + 1}_{=1}\right\} = \alpha\{1\} = \alpha\right),
\end{aligned}$$

qed.

[1166] This is simply the explicit formula for the determinant of a matrix as a sum over permutations.

Subtracting M_α^s from both sides of this equality, we obtain

$$s!M_\alpha^{\langle s\rangle} - M_\alpha^s = \sum_{\substack{\sigma\in\mathfrak{S}_s;\\ \sigma\neq\mathrm{id}}} (-1)^\sigma \prod_{i=1}^s a_{i,\sigma(i)}^{\langle\alpha\rangle}. \tag{13.179.7}$$

But it is easy to see that

$$\prod_{i=1}^s a_{i,\sigma(i)}^{\langle\alpha\rangle} \in \mathcal{F}_{(s-1)\ell} \qquad \text{for every } \sigma\in\mathfrak{S}_s \text{ satisfying } \sigma\neq\mathrm{id} \tag{13.179.8}$$

[1167]. Hence, (13.179.7) becomes

$$s!M_\alpha^{\langle s\rangle} - M_\alpha^s = \sum_{\substack{\sigma\in\mathfrak{S}_s;\\ \sigma\neq\mathrm{id}}} (-1)^\sigma \underbrace{\prod_{i=1}^s a_{i,\sigma(i)}^{\langle\alpha\rangle}}_{\substack{\in\mathcal{F}_{(s-1)\ell}\\ \text{(by (13.179.8))}}} \in \sum_{\substack{\sigma\in\mathfrak{S}_s;\\ \sigma\neq\mathrm{id}}} (-1)^\sigma \mathcal{F}_{(s-1)\ell} \subset \mathcal{F}_{(s-1)\ell} \tag{13.179.9}$$

(since $\mathcal{F}_{(s-1)\ell}$ is a $\mathbf{k}$-module).

But $\mathcal{F}_{(s-1)\ell}$ is the $\mathbf{k}$-submodule of QSym spanned by $(M_\beta)_{\beta\in\mathrm{Comp};\ \ell(\beta)\leq(s-1)\ell}$ (because this is how $\mathcal{F}_{(s-1)\ell}$ was defined). In other words,

$$\mathcal{F}_{(s-1)\ell} = \sum_{\substack{\beta\in\mathrm{Comp};\\ \ell(\beta)\leq(s-1)\ell}} \mathbf{k}M_\beta.$$

Hence, (13.179.9) becomes

$$s!M_\alpha^{\langle s\rangle} - M_\alpha^s \in \mathcal{F}_{(s-1)\ell} = \sum_{\substack{\beta\in\mathrm{Comp};\\ \ell(\beta)\leq(s-1)\ell}} \mathbf{k}M_\beta. \tag{13.179.10}$$

Also, $s!M_\alpha^{\langle s\rangle} - M_\alpha^s \in \mathcal{F}_{(s-1)\ell} \subset \mathrm{QSym}$, so that $s!M_\alpha^{\langle s\rangle} \in \underbrace{M_\alpha^s}_{\in\mathrm{QSym}} + \mathrm{QSym} \subset \mathrm{QSym}+\mathrm{QSym} \subset \mathrm{QSym}$.

Recall that $M_\alpha^{\langle s\rangle} = e_s\left((\mathbf{x}_\mathbf{i}^\alpha)_{\mathbf{i}\in\mathrm{SIS}(\ell)}\right)$. Thus, $M_\alpha^{\langle s\rangle}$ is a homogeneous power series of degree $s|\alpha|$ (since each $\mathbf{x}_\mathbf{i}^\alpha$ is a monomial of degree $|\alpha|$, and since e_s is a power series of degree s). Hence, $s!M_\alpha^{\langle s\rangle}$ is a homogeneous power series of degree $s|\alpha|$ as well. Thus, $s!M_\alpha^{\langle s\rangle} \in \mathrm{QSym}_{s|\alpha|}$ (since $s!M_\alpha^{\langle s\rangle} \in \mathrm{QSym}$). Also, $M_\alpha^s \in \mathrm{QSym}_{s|\alpha|}$ (since $M_\alpha \in \mathrm{QSym}_{|\alpha|}$). Thus, $\underbrace{s!M_\alpha^{\langle s\rangle}}_{\in\mathrm{QSym}_{s|\alpha|}} - \underbrace{M_\alpha^s}_{\in\mathrm{QSym}_{s|\alpha|}} \in \mathrm{QSym}_{s|\alpha|} - \mathrm{QSym}_{s|\alpha|} \subset \mathrm{QSym}_{s|\alpha|}$ (since $\mathrm{QSym}_{s|\alpha|}$ is a $\mathbf{k}$-module).

Now, let π denote the projection from the direct sum $\mathrm{QSym} = \bigoplus_{k\in\mathbb{N}} \mathrm{QSym}_k$ onto its $(s|\alpha|)$-th homogeneous component $\mathrm{QSym}_{s|\alpha|}$. Notice that $\pi\left(s!M_\alpha^{\langle s\rangle} - M_\alpha^s\right) = s!M_\alpha^{\langle s\rangle} - M_\alpha^s$ (since $s!M_\alpha^{\langle s\rangle} - M_\alpha^s \in \mathrm{QSym}_{s|\alpha|}$).

[1167] *Proof of (13.179.8):* Let $\sigma\in\mathfrak{S}_s$ be such that $\sigma\neq\mathrm{id}$. Then, there exists a $k\in\{1,2,\ldots,s\}$ such that $\sigma(k) > k$. Consider this k.

We have $k<\sigma(k)$ (since $\sigma(k)>k$) and thus $a_{k,\sigma(k)}^{\langle\alpha\rangle} \in \mathbf{k}\cdot 1$ (by (13.179.4), applied to $i=k$ and $j=\sigma(k)$). But

$$\underbrace{\prod_{i=1}^s}_{=\prod_{i\in\{1,2,\ldots,s\}}} a_{i,\sigma(i)}^{\langle\alpha\rangle} = \prod_{i\in\{1,2,\ldots,s\}} a_{i,\sigma(i)}^{\langle\alpha\rangle} = \underbrace{a_{k,\sigma(k)}^{\langle\alpha\rangle}}_{\in\mathbf{k}\cdot 1} \cdot \prod_{\substack{i\in\{1,2,\ldots,s\};\\ i\neq k}} \underbrace{a_{i,\sigma(i)}^{\langle\alpha\rangle}}_{\substack{\in\mathcal{F}_\ell\\ \text{(by (13.179.5), applied to } j=\sigma(i))}}$$

(here, we have split off the factor for $i=k$ from the product)

$$\in \mathbf{k}\cdot 1\cdot \underbrace{\prod_{\substack{i\in\{1,2,\ldots,s\};\\ i\neq k}} \mathcal{F}_\ell}_{\substack{=(\mathcal{F}_\ell)^{s-1}\subset\mathcal{F}_{(s-1)\ell}\\ \text{(by (13.179.2), applied to}\\ \ell \text{ and } s-1 \text{ instead of } m \text{ and } k)}} \subset \mathbf{k}\cdot \underbrace{1\cdot\mathcal{F}_{(s-1)\ell}}_{=\mathcal{F}_{(s-1)\ell}} = \mathbf{k}\cdot\mathcal{F}_{(s-1)\ell} \subset \mathcal{F}_{(s-1)\ell}$$

(since $\mathcal{F}_{(s-1)\ell}$ is a $\mathbf{k}$-module). This proves (13.179.8).

We have

$$\pi\left(M_{\beta}\right)=0 \qquad \text{for every } \beta\in\operatorname{Comp}\setminus\operatorname{Comp}_{s\left|\alpha\right|} \tag{13.179.11}$$

[1168].

Now, applying the map π to both sides of the relation (13.179.10), we obtain

$$\begin{aligned}
\pi\left(s!M_{\alpha}^{\langle s\rangle}-M_{\alpha}^{s}\right) &\in \pi\left(\sum_{\substack{\beta\in\operatorname{Comp};\\ \ell(\beta)\leq(s-1)\ell}}\mathbf{k}M_{\beta}\right)=\sum_{\substack{\beta\in\operatorname{Comp};\\ \ell(\beta)\leq(s-1)\ell}}\mathbf{k}\pi\left(M_{\beta}\right) \qquad \left(\text{since the map }\pi\text{ is }\mathbf{k}\text{-linear}\right)\\
&=\sum_{\substack{\beta\in\operatorname{Comp}_{s\left|\alpha\right|};\\ \ell(\beta)\leq(s-1)\ell}}\mathbf{k}\underbrace{\pi\left(M_{\beta}\right)}_{\substack{=M_{\beta}\\ \text{(since }M_{\beta}\in\operatorname{QSym}_{s\left|\alpha\right|}\\ \text{(because }\beta\in\operatorname{Comp}_{s\left|\alpha\right|}\text{),}\\ \text{whereas }\pi\text{ is a projection}\\ \text{onto }\operatorname{QSym}_{s\left|\alpha\right|}\text{)}}}+\sum_{\substack{\beta\in\operatorname{Comp}\setminus\operatorname{Comp}_{s\left|\alpha\right|};\\ \ell(\beta)\leq(s-1)\ell}}\mathbf{k}\underbrace{\pi\left(M_{\beta}\right)}_{\substack{=0\\ \text{(by (13.179.11))}}}\\
&\qquad\left(\begin{array}{c}\text{since the set }\operatorname{Comp}\text{ is the union of its two}\\ \text{disjoint subsets }\operatorname{Comp}_{s\left|\alpha\right|}\text{ and }\operatorname{Comp}\setminus\operatorname{Comp}_{s\left|\alpha\right|}\end{array}\right)\\
&=\sum_{\substack{\beta\in\operatorname{Comp}_{s\left|\alpha\right|};\\ \ell(\beta)\leq(s-1)\ell}}\mathbf{k}M_{\beta}+\underbrace{\sum_{\substack{\beta\in\operatorname{Comp}\setminus\operatorname{Comp}_{s\left|\alpha\right|};\\ \ell(\beta)\leq(s-1)\ell}}\mathbf{k}0}_{=0}=\sum_{\substack{\beta\in\operatorname{Comp}_{s\left|\alpha\right|};\\ \ell(\beta)\leq(s-1)\ell}}\mathbf{k}M_{\beta}=\sum_{\substack{\beta\in\operatorname{Comp}_{s\left|\alpha\right|};\\ \ell(\beta)\leq(s-1)\ell(\alpha)}}\mathbf{k}M_{\beta}
\end{aligned}$$

(since $\ell=\ell\left(\alpha\right)$). Since $\pi\left(s!M_{\alpha}^{\langle s\rangle}-M_{\alpha}^{s}\right)=s!M_{\alpha}^{\langle s\rangle}-M_{\alpha}^{s}$, this rewrites as $s!M_{\alpha}^{\langle s\rangle}-M_{\alpha}^{s}\in\sum_{\substack{\beta\in\operatorname{Comp}_{s\left|\alpha\right|};\\ \ell(\beta)\leq(s-1)\ell(\alpha)}}\mathbf{k}M_{\beta}$. This proves Lemma 6.5.16. □

Second proof of Lemma 6.5.16. Let us first notice that if γ and β are two compositions, then

$$\left(\text{the coefficient of the monomial }\mathbf{x}^{\gamma}\text{ in }M_{\beta}\right)=\delta_{\gamma,\beta}. \tag{13.179.12}$$

(This follows from the definition of M_{β}.)

Write the composition α in the form $\alpha=\left(\alpha_1,\alpha_2,\ldots,\alpha_\ell\right)$; then, $\ell=\ell\left(\alpha\right)$. Let us now fix a total order on the set $\operatorname{SIS}\left(\ell\right)$ (for example, the lexicographic order). Exercise 6.5.4(b) yields

$$M_{\alpha}^{\langle s\rangle}=\sum_{\substack{(\mathbf{i}_1,\mathbf{i}_2,\ldots,\mathbf{i}_s)\in(\operatorname{SIS}(\ell))^s;\\ \mathbf{i}_1<\mathbf{i}_2<\cdots<\mathbf{i}_s}}\mathbf{x}_{\mathbf{i}_1}^{\alpha}\mathbf{x}_{\mathbf{i}_2}^{\alpha}\cdots\mathbf{x}_{\mathbf{i}_s}^{\alpha}. \tag{13.179.13}$$

Now, for every $\left(\mathbf{i}_1,\mathbf{i}_2,\ldots,\mathbf{i}_s\right)\in\left(\operatorname{SIS}\left(\ell\right)\right)^s$ such that $\mathbf{i}_1$, $\mathbf{i}_2$, $\ldots$, $\mathbf{i}_s$ are distinct, there exists a **unique** $\sigma\in\mathfrak{S}_s$ satisfying $\mathbf{i}_{\sigma(1)}<\mathbf{i}_{\sigma(2)}<\cdots<\mathbf{i}_{\sigma(s)}$ (because there is exactly one way to sort the ℓ-tuple $\left(\mathbf{i}_1,\mathbf{i}_2,\ldots,\mathbf{i}_s\right)$ into increasing order[1169]). Hence, we can split the sum $\sum_{\substack{(\mathbf{i}_1,\mathbf{i}_2,\ldots,\mathbf{i}_s)\in(\operatorname{SIS}(\ell))^s;\\ \mathbf{i}_1,\ \mathbf{i}_2,\ \ldots,\ \mathbf{i}_s\text{ are distinct}}}\mathbf{x}_{\mathbf{i}_1}^{\alpha}\mathbf{x}_{\mathbf{i}_2}^{\alpha}\cdots\mathbf{x}_{\mathbf{i}_s}^{\alpha}$ into several subsums,

[1168] *Proof of (13.179.11):* Let $\beta\in\operatorname{Comp}\setminus\operatorname{Comp}_{s\left|\alpha\right|}$. Then, $\beta\in\operatorname{Comp}$ but $\beta\notin\operatorname{Comp}_{s\left|\alpha\right|}$. In other words, β is a composition with size $\left|\beta\right|\neq s\left|\alpha\right|$. As a consequence, M_{β} is a homogeneous element of QSym of degree $\left|\beta\right|\neq s\left|\alpha\right|$. Therefore, $\pi\left(M_{\beta}\right)=0$ (since π is the projection from the direct sum $\operatorname{QSym}=\bigoplus_{k\in\mathbb{N}}\operatorname{QSym}_k$ onto its $\left(s\left|\alpha\right|\right)$-th homogeneous component $\operatorname{QSym}_{s\left|\alpha\right|}$). This proves (13.179.11).

[1169] Here, we are using that the order on $\operatorname{SIS}\left(\ell\right)$ is total.

one for each $\sigma \in \mathfrak{S}_s$, as follows:

$$\begin{aligned}
&\sum_{\substack{(\mathbf{i}_1,\mathbf{i}_2,\ldots,\mathbf{i}_s)\in(\operatorname{SIS}(\ell))^s;\\ \mathbf{i}_1,\ \mathbf{i}_2,\ \ldots,\ \mathbf{i}_s \text{ are distinct}}} \mathbf{x}_{\mathbf{i}_1}^{\alpha}\mathbf{x}_{\mathbf{i}_2}^{\alpha}\cdots\mathbf{x}_{\mathbf{i}_s}^{\alpha} \\
&= \sum_{\sigma\in\mathfrak{S}_s} \underbrace{\sum_{\substack{(\mathbf{i}_1,\mathbf{i}_2,\ldots,\mathbf{i}_s)\in(\operatorname{SIS}(\ell))^s;\\ \mathbf{i}_1,\ \mathbf{i}_2,\ \ldots,\ \mathbf{i}_s \text{ are distinct;}\\ \mathbf{i}_{\sigma(1)}<\mathbf{i}_{\sigma(2)}<\cdots<\mathbf{i}_{\sigma(s)}}}}_{\substack{=\sum_{\substack{(\mathbf{i}_1,\mathbf{i}_2,\ldots,\mathbf{i}_s)\in(\operatorname{SIS}(\ell))^s;\\ \mathbf{i}_{\sigma(1)}<\mathbf{i}_{\sigma(2)}<\cdots<\mathbf{i}_{\sigma(s)}}} \\ \text{(since the condition that } \mathbf{i}_1,\ \mathbf{i}_2,\ \ldots,\ \mathbf{i}_s \text{ are distinct} \\ \text{follows from the condition that } \mathbf{i}_{\sigma(1)}<\mathbf{i}_{\sigma(2)}<\cdots<\mathbf{i}_{\sigma(s)} \\ \text{(because if } \mathbf{i}_{\sigma(1)}<\mathbf{i}_{\sigma(2)}<\cdots<\mathbf{i}_{\sigma(s)}\text{, then } \mathbf{i}_{\sigma(1)},\ \mathbf{i}_{\sigma(2)},\ \ldots,\ \mathbf{i}_{\sigma(s)} \text{ are} \\ \text{distinct, and thus } \mathbf{i}_1,\ \mathbf{i}_2,\ \ldots,\ \mathbf{i}_s \text{ are distinct} \\ \text{(since } \sigma \text{ is a permutation of } \{1,2,\ldots,s\})))}} \mathbf{x}_{\mathbf{i}_1}^{\alpha}\mathbf{x}_{\mathbf{i}_2}^{\alpha}\cdots\mathbf{x}_{\mathbf{i}_s}^{\alpha} \\
&= \sum_{\sigma\in\mathfrak{S}_s} \sum_{\substack{(\mathbf{i}_1,\mathbf{i}_2,\ldots,\mathbf{i}_s)\in(\operatorname{SIS}(\ell))^s;\\ \mathbf{i}_{\sigma(1)}<\mathbf{i}_{\sigma(2)}<\cdots<\mathbf{i}_{\sigma(s)}}} \mathbf{x}_{\mathbf{i}_1}^{\alpha}\mathbf{x}_{\mathbf{i}_2}^{\alpha}\cdots\mathbf{x}_{\mathbf{i}_s}^{\alpha}.
\end{aligned} \tag{13.179.14}$$

But every $\sigma \in \mathfrak{S}_s$ and every $(\mathbf{i}_1, \mathbf{i}_2, \ldots, \mathbf{i}_s) \in (\operatorname{SIS}(\ell))^s$ satisfy

$$\mathbf{x}_{\mathbf{i}_1}^{\alpha}\mathbf{x}_{\mathbf{i}_2}^{\alpha}\cdots\mathbf{x}_{\mathbf{i}_s}^{\alpha} = \mathbf{x}_{\mathbf{i}_{\sigma(1)}}^{\alpha}\mathbf{x}_{\mathbf{i}_{\sigma(2)}}^{\alpha}\cdots\mathbf{x}_{\mathbf{i}_{\sigma(s)}}^{\alpha} \tag{13.179.15}$$

[1170]. Now, (13.179.14) becomes

$$\begin{aligned}
&\sum_{\substack{(\mathbf{i}_1,\mathbf{i}_2,\ldots,\mathbf{i}_s)\in(\operatorname{SIS}(\ell))^s;\\ \mathbf{i}_1,\ \mathbf{i}_2,\ \ldots,\ \mathbf{i}_s \text{ are distinct}}} \mathbf{x}_{\mathbf{i}_1}^{\alpha}\mathbf{x}_{\mathbf{i}_2}^{\alpha}\cdots\mathbf{x}_{\mathbf{i}_s}^{\alpha} \\
&= \sum_{\sigma\in\mathfrak{S}_s} \sum_{\substack{(\mathbf{i}_1,\mathbf{i}_2,\ldots,\mathbf{i}_s)\in(\operatorname{SIS}(\ell))^s;\\ \mathbf{i}_{\sigma(1)}<\mathbf{i}_{\sigma(2)}<\cdots<\mathbf{i}_{\sigma(s)}}} \underbrace{\mathbf{x}_{\mathbf{i}_1}^{\alpha}\mathbf{x}_{\mathbf{i}_2}^{\alpha}\cdots\mathbf{x}_{\mathbf{i}_s}^{\alpha}}_{\substack{=\mathbf{x}_{\mathbf{i}_{\sigma(1)}}^{\alpha}\mathbf{x}_{\mathbf{i}_{\sigma(2)}}^{\alpha}\cdots\mathbf{x}_{\mathbf{i}_{\sigma(s)}}^{\alpha} \\ \text{(by (13.179.15))}}} \\
&= \sum_{\sigma\in\mathfrak{S}_s} \underbrace{\sum_{\substack{(\mathbf{i}_1,\mathbf{i}_2,\ldots,\mathbf{i}_s)\in(\operatorname{SIS}(\ell))^s;\\ \mathbf{i}_{\sigma(1)}<\mathbf{i}_{\sigma(2)}<\cdots<\mathbf{i}_{\sigma(s)}}} \mathbf{x}_{\mathbf{i}_{\sigma(1)}}^{\alpha}\mathbf{x}_{\mathbf{i}_{\sigma(2)}}^{\alpha}\cdots\mathbf{x}_{\mathbf{i}_{\sigma(s)}}^{\alpha}}_{\substack{=\sum_{\substack{(\mathbf{i}_1,\mathbf{i}_2,\ldots,\mathbf{i}_s)\in(\operatorname{SIS}(\ell))^s;\\ \mathbf{i}_1<\mathbf{i}_2<\cdots<\mathbf{i}_s}} \mathbf{x}_{\mathbf{i}_1}^{\alpha}\mathbf{x}_{\mathbf{i}_2}^{\alpha}\cdots\mathbf{x}_{\mathbf{i}_s}^{\alpha} \\ \text{(here, we have substituted } (\mathbf{i}_1,\mathbf{i}_2,\ldots,\mathbf{i}_s) \\ \text{for } \left(\mathbf{i}_{\sigma(1)},\mathbf{i}_{\sigma(2)},\ldots,\mathbf{i}_{\sigma(s)}\right) \text{ in the sum} \\ \text{(since } \sigma \text{ is a permutation of } \{1,2,\ldots,s\}))}} \\
&= \sum_{\sigma\in\mathfrak{S}_s} \underbrace{\sum_{\substack{(\mathbf{i}_1,\mathbf{i}_2,\ldots,\mathbf{i}_s)\in(\operatorname{SIS}(\ell))^s;\\ \mathbf{i}_1<\mathbf{i}_2<\cdots<\mathbf{i}_s}} \mathbf{x}_{\mathbf{i}_1}^{\alpha}\mathbf{x}_{\mathbf{i}_2}^{\alpha}\cdots\mathbf{x}_{\mathbf{i}_s}^{\alpha}}_{\substack{=M_{\alpha}^{\langle s\rangle} \\ \text{(by (13.179.13))}}} = \sum_{\sigma\in\mathfrak{S}_s} M_{\alpha}^{\langle s\rangle} = \underbrace{\left|\mathfrak{S}_s\right|}_{=s!} M_{\alpha}^{\langle s\rangle} = s! M_{\alpha}^{\langle s\rangle}.
\end{aligned} \tag{13.179.16}$$

[1170] *Proof of (13.179.15):* Let $\sigma \in \mathfrak{S}_s$ and $(\mathbf{i}_1, \mathbf{i}_2, \ldots, \mathbf{i}_s) \in (\operatorname{SIS}(\ell))^s$. Then, the product $\mathbf{x}_{\mathbf{i}_{\sigma(1)}}^{\alpha}\mathbf{x}_{\mathbf{i}_{\sigma(2)}}^{\alpha}\cdots\mathbf{x}_{\mathbf{i}_{\sigma(s)}}^{\alpha}$ is obtained from the product $\mathbf{x}_{\mathbf{i}_1}^{\alpha}\mathbf{x}_{\mathbf{i}_2}^{\alpha}\cdots\mathbf{x}_{\mathbf{i}_s}^{\alpha}$ by rearranging the factors according to the permutation σ. Since rearranging the factors of a product in QSym does not change the value of the product (because the algebra QSym is commutative), this yields that $\mathbf{x}_{\mathbf{i}_{\sigma(1)}}^{\alpha}\mathbf{x}_{\mathbf{i}_{\sigma(2)}}^{\alpha}\cdots\mathbf{x}_{\mathbf{i}_{\sigma(s)}}^{\alpha} = \mathbf{x}_{\mathbf{i}_1}^{\alpha}\mathbf{x}_{\mathbf{i}_2}^{\alpha}\cdots\mathbf{x}_{\mathbf{i}_s}^{\alpha}$. This proves (13.179.15).

On the other hand, taking both sides of the identity (6.5.1) to the s-th power, we obtain

$$\begin{aligned} M_\alpha^s &= \left(\sum_{\mathbf{i} \in \operatorname{SIS}(\ell)} \mathbf{x}_{\mathbf{i}}^\alpha \right)^s = \sum_{(\mathbf{i}_1, \mathbf{i}_2, \ldots, \mathbf{i}_s) \in (\operatorname{SIS}(\ell))^s} \mathbf{x}_{\mathbf{i}_1}^\alpha \mathbf{x}_{\mathbf{i}_2}^\alpha \cdots \mathbf{x}_{\mathbf{i}_s}^\alpha \qquad \text{(by the product rule)} \\ &= \underbrace{\sum_{\substack{(\mathbf{i}_1, \mathbf{i}_2, \ldots, \mathbf{i}_s) \in (\operatorname{SIS}(\ell))^s; \\ \mathbf{i}_1,\ \mathbf{i}_2,\ \ldots,\ \mathbf{i}_s \text{ are distinct}}} \mathbf{x}_{\mathbf{i}_1}^\alpha \mathbf{x}_{\mathbf{i}_2}^\alpha \cdots \mathbf{x}_{\mathbf{i}_s}^\alpha}_{\substack{= s! M_\alpha^{\langle s \rangle} \\ \text{(by (13.179.16))}}} + \sum_{\substack{(\mathbf{i}_1, \mathbf{i}_2, \ldots, \mathbf{i}_s) \in (\operatorname{SIS}(\ell))^s; \\ \mathbf{i}_1,\ \mathbf{i}_2,\ \ldots,\ \mathbf{i}_s \text{ are not distinct}}} \mathbf{x}_{\mathbf{i}_1}^\alpha \mathbf{x}_{\mathbf{i}_2}^\alpha \cdots \mathbf{x}_{\mathbf{i}_s}^\alpha \\ &= s! M_\alpha^{\langle s \rangle} + \sum_{\substack{(\mathbf{i}_1, \mathbf{i}_2, \ldots, \mathbf{i}_s) \in (\operatorname{SIS}(\ell))^s; \\ \mathbf{i}_1,\ \mathbf{i}_2,\ \ldots,\ \mathbf{i}_s \text{ are not distinct}}} \mathbf{x}_{\mathbf{i}_1}^\alpha \mathbf{x}_{\mathbf{i}_2}^\alpha \cdots \mathbf{x}_{\mathbf{i}_s}^\alpha . \end{aligned}$$

Hence,

$$M_\alpha^s - s! M_\alpha^{\langle s \rangle} = \sum_{\substack{(\mathbf{i}_1, \mathbf{i}_2, \ldots, \mathbf{i}_s) \in (\operatorname{SIS}(\ell))^s; \\ \mathbf{i}_1,\ \mathbf{i}_2,\ \ldots,\ \mathbf{i}_s \text{ are not distinct}}} \mathbf{x}_{\mathbf{i}_1}^\alpha \mathbf{x}_{\mathbf{i}_2}^\alpha \cdots \mathbf{x}_{\mathbf{i}_s}^\alpha . \tag{13.179.17}$$

But $M_\alpha^{\langle s \rangle} \in \operatorname{QSym}$ (by Corollary 6.5.8(a)), so that $\underbrace{M_\alpha^s}_{\in \operatorname{QSym}} - s! \underbrace{M_\alpha^{\langle s \rangle}}_{\in \operatorname{QSym}} \in \operatorname{QSym} - s! \operatorname{QSym} \subset \operatorname{QSym}$. Hence, $M_\alpha^s - s! M_\alpha^{\langle s \rangle}$ is a $\mathbf{k}$-linear combination of the family $(M_\beta)_{\beta \in \operatorname{Comp}}$ (since this family $(M_\beta)_{\beta \in \operatorname{Comp}}$ is a basis of the $\mathbf{k}$-module QSym). In other words, there exists a family $(c_\beta)_{\beta \in \operatorname{Comp}} \in \mathbf{k}^{\operatorname{Comp}}$ of elements of $\mathbf{k}$ such that (all but finitely many $\beta \in \operatorname{Comp}$ satisfy $c_\beta = 0$) and

$$s! M_\alpha^{\langle s \rangle} - M_\alpha^s = \sum_{\beta \in \operatorname{Comp}} c_\beta M_\beta . \tag{13.179.18}$$

Consider this family $(c_\beta)_{\beta \in \operatorname{Comp}} \in \mathbf{k}^{\operatorname{Comp}}$. Every composition γ satisfies

$$\begin{aligned} &\left(\text{the coefficient of the monomial } \mathbf{x}^\gamma \text{ in } \underbrace{s! M_\alpha^{\langle s \rangle} - M_\alpha^s}_{\substack{= \sum_{\beta \in \operatorname{Comp}} c_\beta M_\beta \\ \text{(by (13.179.18))}}} \right) \\ &= \left(\text{the coefficient of the monomial } \mathbf{x}^\gamma \text{ in } \sum_{\beta \in \operatorname{Comp}} c_\beta M_\beta \right) \\ &= \sum_{\beta \in \operatorname{Comp}} c_\beta \underbrace{(\text{the coefficient of the monomial } \mathbf{x}^\gamma \text{ in } M_\beta)}_{\substack{= \delta_{\gamma, \beta} \\ \text{(by (13.179.12))}}} \\ &= \sum_{\beta \in \operatorname{Comp}} c_\beta \delta_{\gamma, \beta} = c_\gamma . \end{aligned}$$

Hence, every composition γ satisfies

$$c_\gamma = \left(\text{the coefficient of the monomial } \mathbf{x}^\gamma \text{ in } s! M_\alpha^{\langle s \rangle} - M_\alpha^s \right) .$$

Thus, every composition γ satisfies

$$\begin{aligned}
-c_\gamma &= -\left(\text{the coefficient of the monomial } \mathbf{x}^\gamma \text{ in } s!M_\alpha^{\langle s\rangle} - M_\alpha^s\right) \\
&= \left(\text{the coefficient of the monomial } \mathbf{x}^\gamma \text{ in } \underbrace{-\left(s!M_\alpha^{\langle s\rangle} - M_\alpha^s\right)}_{\substack{=M_\alpha^s - s!M_\alpha^{\langle s\rangle} \\ = \sum_{\substack{(\mathbf{i}_1,\mathbf{i}_2,\ldots,\mathbf{i}_s)\in(\operatorname{SIS}(\ell))^s; \\ \mathbf{i}_1,\ \mathbf{i}_2,\ \ldots,\ \mathbf{i}_s \text{ are not distinct}}} \mathbf{x}_{\mathbf{i}_1}^\alpha \mathbf{x}_{\mathbf{i}_2}^\alpha \cdots \mathbf{x}_{\mathbf{i}_s}^\alpha \\ \text{(by (13.179.17))}}}\right) \\
&= \left(\text{the coefficient of the monomial } \mathbf{x}^\gamma \text{ in } \sum_{\substack{(\mathbf{i}_1,\mathbf{i}_2,\ldots,\mathbf{i}_s)\in(\operatorname{SIS}(\ell))^s; \\ \mathbf{i}_1,\ \mathbf{i}_2,\ \ldots,\ \mathbf{i}_s \text{ are not distinct}}} \mathbf{x}_{\mathbf{i}_1}^\alpha \mathbf{x}_{\mathbf{i}_2}^\alpha \cdots \mathbf{x}_{\mathbf{i}_s}^\alpha\right) \\
&= \sum_{\substack{(\mathbf{i}_1,\mathbf{i}_2,\ldots,\mathbf{i}_s)\in(\operatorname{SIS}(\ell))^s; \\ \mathbf{i}_1,\ \mathbf{i}_2,\ \ldots,\ \mathbf{i}_s \text{ are not distinct}}} \left(\text{the coefficient of the monomial } \mathbf{x}^\gamma \text{ in } \mathbf{x}_{\mathbf{i}_1}^\alpha \mathbf{x}_{\mathbf{i}_2}^\alpha \cdots \mathbf{x}_{\mathbf{i}_s}^\alpha\right). \qquad (13.179.19)
\end{aligned}$$

We now notice that if γ is a composition satisfying $|\gamma| \neq s\,|\alpha|$, then

$$\left(\text{the coefficient of the monomial } \mathbf{x}^\gamma \text{ in } \mathbf{x}_{\mathbf{i}_1}^\alpha \mathbf{x}_{\mathbf{i}_2}^\alpha \cdots \mathbf{x}_{\mathbf{i}_s}^\alpha\right) = 0 \qquad (13.179.20)$$

for every $(\mathbf{i}_1, \mathbf{i}_2, \ldots, \mathbf{i}_s) \in (\operatorname{SIS}(\ell))^s$ [1171]. Also, if γ is a composition satisfying $\ell(\gamma) > (s-1)\,\ell(\alpha)$, then

$$\left(\text{the coefficient of the monomial } \mathbf{x}^\gamma \text{ in } \mathbf{x}_{\mathbf{i}_1}^\alpha \mathbf{x}_{\mathbf{i}_2}^\alpha \cdots \mathbf{x}_{\mathbf{i}_s}^\alpha\right) = 0 \qquad (13.179.22)$$

[1171] *Proof of (13.179.20):* Let γ be a composition satisfying $|\gamma| \neq s\,|\alpha|$. Let $(\mathbf{i}_1, \mathbf{i}_2, \ldots, \mathbf{i}_s) \in (\operatorname{SIS}(\ell))^s$. Then, the monomial $\mathbf{x}_{\mathbf{i}_1}^\alpha \mathbf{x}_{\mathbf{i}_2}^\alpha \cdots \mathbf{x}_{\mathbf{i}_s}^\alpha$ has degree

$$\deg\left(\mathbf{x}_{\mathbf{i}_1}^\alpha \mathbf{x}_{\mathbf{i}_2}^\alpha \cdots \mathbf{x}_{\mathbf{i}_s}^\alpha\right) = \deg\left(\mathbf{x}_{\mathbf{i}_1}^\alpha\right) + \deg\left(\mathbf{x}_{\mathbf{i}_2}^\alpha\right) + \cdots + \deg\left(\mathbf{x}_{\mathbf{i}_s}^\alpha\right).$$

However, for every $\mathbf{i} \in \operatorname{SIS}(\ell)$, the monomial $\mathbf{x}_\mathbf{i}^\alpha$ is a monomial of degree $|\alpha|$. Thus, for every $\mathbf{i} \in \operatorname{SIS}(\ell)$, we have

$$\deg\left(\mathbf{x}_\mathbf{i}^\alpha\right) = |\alpha|. \qquad (13.179.21)$$

Hence, for every $k \in \{1, 2, \ldots, s\}$, we have $\deg\left(\mathbf{x}_{\mathbf{i}_k}^\alpha\right) = |\alpha|$ (by (13.179.21), applied to $\mathbf{i} = \mathbf{i}_k$). Adding up these equalities over all $k \in \{1, 2, \ldots, s\}$, we obtain $\deg\left(\mathbf{x}_{\mathbf{i}_1}^\alpha\right) + \deg\left(\mathbf{x}_{\mathbf{i}_2}^\alpha\right) + \cdots + \deg\left(\mathbf{x}_{\mathbf{i}_s}^\alpha\right) = \underbrace{|\alpha| + |\alpha| + \cdots + |\alpha|}_{s \text{ times}} = s\,|\alpha|$. Hence,

$$\deg\left(\mathbf{x}_{\mathbf{i}_1}^\alpha \mathbf{x}_{\mathbf{i}_2}^\alpha \cdots \mathbf{x}_{\mathbf{i}_s}^\alpha\right) = \deg\left(\mathbf{x}_{\mathbf{i}_1}^\alpha\right) + \deg\left(\mathbf{x}_{\mathbf{i}_2}^\alpha\right) + \cdots + \deg\left(\mathbf{x}_{\mathbf{i}_s}^\alpha\right) = s\,|\alpha|,$$

so that $s\,|\alpha| = \deg\left(\mathbf{x}_{\mathbf{i}_1}^\alpha \mathbf{x}_{\mathbf{i}_2}^\alpha \cdots \mathbf{x}_{\mathbf{i}_s}^\alpha\right)$. But $\deg\left(\mathbf{x}^\gamma\right) = |\gamma| \neq s\,|\alpha| = \deg\left(\mathbf{x}_{\mathbf{i}_1}^\alpha \mathbf{x}_{\mathbf{i}_2}^\alpha \cdots \mathbf{x}_{\mathbf{i}_s}^\alpha\right)$. In other words, the monomials $\mathbf{x}^\gamma$ and $\mathbf{x}_{\mathbf{i}_1}^\alpha \mathbf{x}_{\mathbf{i}_2}^\alpha \cdots \mathbf{x}_{\mathbf{i}_s}^\alpha$ have different degrees; thus, these monomials are distinct. Therefore, $\left(\text{the coefficient of the monomial } \mathbf{x}^\gamma \text{ in } \mathbf{x}_{\mathbf{i}_1}^\alpha \mathbf{x}_{\mathbf{i}_2}^\alpha \cdots \mathbf{x}_{\mathbf{i}_s}^\alpha\right) = 0$. This proves (13.179.20).

for every $(\mathbf{i}_1, \mathbf{i}_2, \ldots, \mathbf{i}_s) \in (\operatorname{SIS}(\ell))^s$ satisfying ($\mathbf{i}_1$, $\mathbf{i}_2$, $\ldots$, $\mathbf{i}_s$ are not distinct) [1172]. Now, if γ is a composition satisfying $|\gamma| \neq s|\alpha|$, then

$$-c_\gamma = \sum_{\substack{(\mathbf{i}_1,\mathbf{i}_2,\ldots,\mathbf{i}_s)\in(\operatorname{SIS}(\ell))^s;\\ \mathbf{i}_1,\ \mathbf{i}_2,\ \ldots,\ \mathbf{i}_s \text{ are not distinct}}} \underbrace{\left(\text{the coefficient of the monomial } \mathbf{x}^\gamma \text{ in } \mathbf{x}_{\mathbf{i}_1}^\alpha \mathbf{x}_{\mathbf{i}_2}^\alpha \cdots \mathbf{x}_{\mathbf{i}_s}^\alpha\right)}_{\substack{=0\\ (\text{by } (13.179.20))}} \qquad (\text{by } (13.179.19))$$
$$= \sum_{\substack{(\mathbf{i}_1,\mathbf{i}_2,\ldots,\mathbf{i}_s)\in(\operatorname{SIS}(\ell))^s;\\ \mathbf{i}_1,\ \mathbf{i}_2,\ \ldots,\ \mathbf{i}_s \text{ are not distinct}}} 0 = 0.$$

[1172] *Proof of (13.179.22):* Let γ be a composition satisfying $\ell(\gamma) > (s-1)\ell(\alpha)$. Let $(\mathbf{i}_1, \mathbf{i}_2, \ldots, \mathbf{i}_s) \in (\operatorname{SIS}(\ell))^s$ be such that ($\mathbf{i}_1$, $\mathbf{i}_2$, $\ldots$, $\mathbf{i}_s$ are not distinct).

For every monomial $\mathbf{m}$ in the variables $x_1, x_2, x_3, \ldots$, we denote by $\operatorname{Supp}\mathbf{m}$ the set of all variables which occur in the monomial $\mathbf{m}$ (where we say that a variable *occurs* in $\mathbf{m}$ if and only if its exponent in $\mathbf{m}$ is positive). For instance, $\operatorname{Supp} 1 = \varnothing$ and $\operatorname{Supp}\left(x_2^3 x_3 x_5^4\right) = \{x_2, x_3, x_5\}$ and $\operatorname{Supp}\left(\underbrace{x_3 x_4^0}_{=x_3}\right) = \operatorname{Supp}(x_3) = \{x_3\}$. It is very easy to see that $\operatorname{Supp}(\mathbf{mn}) = (\operatorname{Supp}\mathbf{m}) \cup (\operatorname{Supp}\mathbf{n})$ for any two monomials $\mathbf{m}$ and $\mathbf{n}$. More generally, if $k \in \mathbb{N}$, and if $\mathbf{m}_1$, $\mathbf{m}_2$, $\ldots$, $\mathbf{m}_k$ are k monomials, then

$$(13.179.23) \qquad \operatorname{Supp}(\mathbf{m}_1\mathbf{m}_2\cdots\mathbf{m}_k) = (\operatorname{Supp}(\mathbf{m}_1)) \cup (\operatorname{Supp}(\mathbf{m}_2)) \cup \cdots \cup (\operatorname{Supp}(\mathbf{m}_k)).$$

On the other hand,

$$(13.179.24) \qquad \left|\operatorname{Supp}\left(\mathbf{x}^\beta\right)\right| = \ell(\beta) \qquad \text{for every composition } \beta.$$

[*Proof of (13.179.24):* Let β be a composition. Write β in the form $\beta = \left(\beta_1, \beta_2, \ldots, \beta_{\ell(\beta)}\right)$. Then, $\mathbf{x}^\beta = x_1^{\beta_1} x_2^{\beta_2} \cdots x_{\ell(\beta)}^{\beta_{\ell(\beta)}}$. Hence, the variables x_1, x_2, $\ldots$, $x_{\ell(\beta)}$ all occur in the monomial $\mathbf{x}^\beta$ (since their exponents β_1, β_2, $\ldots$, $\beta_{\ell(\beta)}$ are positive), and no other variables do. In other words, the set of all variables which occur in the monomial $\mathbf{x}^\beta$ is $\left\{x_1, x_2, \ldots, x_{\ell(\beta)}\right\}$. In other words, $\operatorname{Supp}\left(\mathbf{x}^\beta\right)$ is $\left\{x_1, x_2, \ldots, x_{\ell(\beta)}\right\}$ (since $\operatorname{Supp}\left(\mathbf{x}^\beta\right)$ is the set of all variables which occur in the monomial $\mathbf{x}^\beta$ (by the definition of $\operatorname{Supp}\left(\mathbf{x}^\beta\right)$)). In other words, $\operatorname{Supp}\left(\mathbf{x}^\beta\right) = \left\{x_1, x_2, \ldots, x_{\ell(\beta)}\right\}$, so that $\left|\operatorname{Supp}\left(\mathbf{x}^\beta\right)\right| = \left|\left\{x_1, x_2, \ldots, x_{\ell(\beta)}\right\}\right| = \ell(\beta)$. This proves (13.179.24).]

Also,

$$(13.179.25) \qquad \left|\operatorname{Supp}\left(\mathbf{x}_\mathbf{i}^\alpha\right)\right| = \ell \qquad \text{for every } \mathbf{i} \in \operatorname{SIS}(\ell).$$

[*Proof of (13.179.25):* Let $\mathbf{i} \in \operatorname{SIS}(\ell)$. Then, $\mathbf{i}$ is a strictly increasing ℓ-tuple of positive integers. In other words, we can write $\mathbf{i}$ in the form $\mathbf{i} = (i_1, i_2, \ldots, i_\ell)$ for some positive integers i_1, i_2, $\ldots$, i_ℓ satisfying $i_1 < i_2 < \cdots < i_\ell$. Consider these i_1, i_2, $\ldots$, i_ℓ. Notice that i_1, i_2, $\ldots$, i_ℓ are distinct (since $i_1 < i_2 < \cdots < i_\ell$), so that the variables x_{i_1}, x_{i_2}, $\ldots$, x_{i_ℓ} are distinct.

The definition of $\mathbf{x}_\mathbf{i}^\alpha$ yields $\mathbf{x}_\mathbf{i}^\alpha = x_{i_1}^{\alpha_1} x_{i_2}^{\alpha_2} \cdots x_{i_\ell}^{\alpha_\ell}$. Thus, the variables x_{i_1}, x_{i_2}, $\ldots$, x_{i_ℓ} all occur in the monomial $\mathbf{x}_\mathbf{i}^\alpha$ (since their exponents α_1, α_2, $\ldots$, α_ℓ are positive), and no other variables do. In other words, the set of all variables which occur in the monomial $\mathbf{x}_\mathbf{i}^\alpha$ is $\left\{x_{i_1}, x_{i_2}, \ldots, x_{i_\ell}\right\}$. In other words, $\operatorname{Supp}\left(\mathbf{x}_\mathbf{i}^\alpha\right)$ is $\left\{x_{i_1}, x_{i_2}, \ldots, x_{i_\ell}\right\}$ (since $\operatorname{Supp}\left(\mathbf{x}_\mathbf{i}^\alpha\right)$ is the set of all variables which occur in the monomial $\mathbf{x}_\mathbf{i}^\alpha$ (by the definition of $\operatorname{Supp}\left(\mathbf{x}_\mathbf{i}^\alpha\right)$)). In other words, $\operatorname{Supp}\left(\mathbf{x}_\mathbf{i}^\alpha\right) = \left\{x_{i_1}, x_{i_2}, \ldots, x_{i_\ell}\right\}$, so that $\left|\operatorname{Supp}\left(\mathbf{x}_\mathbf{i}^\alpha\right)\right| = \left|\left\{x_{i_1}, x_{i_2}, \ldots, x_{i_\ell}\right\}\right| = \ell$ (since the variables x_{i_1}, x_{i_2}, $\ldots$, x_{i_ℓ} are distinct). This proves (13.179.25).]

We know that $\mathbf{i}_1$, $\mathbf{i}_2$, $\ldots$, $\mathbf{i}_s$ are not distinct. In other words, (at least) two of the elements $\mathbf{i}_1$, $\mathbf{i}_2$, $\ldots$, $\mathbf{i}_s$ are equal. In other words, there exist two distinct elements u and v of $\{1, 2, \ldots, s\}$ such that $\mathbf{i}_u = \mathbf{i}_v$. Consider these u and v. Notice that $u \in \{1, 2, \ldots, s\}$, so that $1 \leq u \leq s$ and thus $s \geq 1$.

Since u and v are distinct, we have $u \in \{1, 2, \ldots, s\} \setminus \{v\}$; thus, $\operatorname{Supp}\left(\mathbf{x}_{\mathbf{i}_u}^\alpha\right)$ is an addend of the union $\bigcup_{j\in\{1,2,\ldots,s\}\setminus\{v\}} \operatorname{Supp}\left(\mathbf{x}_{\mathbf{i}_j}^\alpha\right)$. Consequently, $\operatorname{Supp}\left(\mathbf{x}_{\mathbf{i}_u}^\alpha\right) \subset \bigcup_{j\in\{1,2,\ldots,s\}\setminus\{v\}} \operatorname{Supp}\left(\mathbf{x}_{\mathbf{i}_j}^\alpha\right)$. Since $\mathbf{i}_u = \mathbf{i}_v$, this rewrites as

$$(13.179.26) \qquad \operatorname{Supp}\left(\mathbf{x}_{\mathbf{i}_v}^\alpha\right) \subset \bigcup_{j\in\{1,2,\ldots,s\}\setminus\{v\}} \operatorname{Supp}\left(\mathbf{x}_{\mathbf{i}_j}^\alpha\right).$$

Now, (13.179.23) (applied to $k = s$ and $\mathbf{m}_j = \mathbf{x}_{\mathbf{i}_j}^\alpha$) yields

$$\operatorname{Supp}\left(\mathbf{x}_{\mathbf{i}_1}^\alpha \mathbf{x}_{\mathbf{i}_2}^\alpha \cdots \mathbf{x}_{\mathbf{i}_s}^\alpha\right) = \left(\operatorname{Supp}\left(\mathbf{x}_{\mathbf{i}_1}^\alpha\right)\right) \cup \left(\operatorname{Supp}\left(\mathbf{x}_{\mathbf{i}_2}^\alpha\right)\right) \cup \cdots \cup \left(\operatorname{Supp}\left(\mathbf{x}_{\mathbf{i}_s}^\alpha\right)\right)$$
$$= \bigcup_{j\in\{1,2,\ldots,s\}} \operatorname{Supp}\left(\mathbf{x}_{\mathbf{i}_j}^\alpha\right) = \left(\bigcup_{j\in\{1,2,\ldots,s\}\setminus\{v\}} \operatorname{Supp}\left(\mathbf{x}_{\mathbf{i}_j}^\alpha\right)\right) \cup \underbrace{\operatorname{Supp}\left(\mathbf{x}_{\mathbf{i}_v}^\alpha\right)}_{\substack{\subset \bigcup_{j\in\{1,2,\ldots,s\}\setminus\{v\}} \operatorname{Supp}\left(\mathbf{x}_{\mathbf{i}_j}^\alpha\right)\\ (\text{by } (13.179.26))}}$$

(here, we have split off the addend for $j = v$ from the union)

$$\subset \left(\bigcup_{j\in\{1,2,\ldots,s\}\setminus\{v\}} \operatorname{Supp}\left(\mathbf{x}_{\mathbf{i}_j}^\alpha\right)\right) \cup \left(\bigcup_{j\in\{1,2,\ldots,s\}\setminus\{v\}} \operatorname{Supp}\left(\mathbf{x}_{\mathbf{i}_j}^\alpha\right)\right)$$
$$= \bigcup_{j\in\{1,2,\ldots,s\}\setminus\{v\}} \operatorname{Supp}\left(\mathbf{x}_{\mathbf{i}_j}^\alpha\right).$$

Hence, if γ is a composition satisfying $|\gamma| \neq s\,|\alpha|$, then

$$c_\gamma = 0. \tag{13.179.28}$$

Also, if γ is a composition satisfying $\ell(\gamma) > (s-1)\,\ell(\alpha)$, then

$$\begin{aligned}
-c_\gamma &= \sum_{\substack{(\mathbf{i}_1,\mathbf{i}_2,\ldots,\mathbf{i}_s)\in(\mathrm{SIS}(\ell))^s;\\ \mathbf{i}_1,\ \mathbf{i}_2,\ \ldots,\ \mathbf{i}_s \text{ are not distinct}}} \underbrace{\left(\text{the coefficient of the monomial } \mathbf{x}^\gamma \text{ in } \mathbf{x}_{\mathbf{i}_1}^\alpha \mathbf{x}_{\mathbf{i}_2}^\alpha \cdots \mathbf{x}_{\mathbf{i}_s}^\alpha\right)}_{\substack{=0\\ (\text{by (13.179.22)})}} \qquad (\text{by (13.179.19)})\\
&= \sum_{\substack{(\mathbf{i}_1,\mathbf{i}_2,\ldots,\mathbf{i}_s)\in(\mathrm{SIS}(\ell))^s;\\ \mathbf{i}_1,\ \mathbf{i}_2,\ \ldots,\ \mathbf{i}_s \text{ are not distinct}}} 0 = 0.
\end{aligned}$$

Hence, if γ is a composition satisfying $\ell(\gamma) > (s-1)\,\ell(\alpha)$, then

$$c_\gamma = 0. \tag{13.179.29}$$

Now, (13.179.18) becomes

$$\begin{aligned}
s!M_\alpha^{\langle s\rangle} - M_\alpha^s &= \sum_{\beta\in\mathrm{Comp}} c_\beta M_\beta = \underbrace{\sum_{\substack{\beta\in\mathrm{Comp};\\ |\beta|=s|\alpha|}}}_{\substack{=\sum_{\beta\in\mathrm{Comp}_{s|\alpha|}}\\ (\text{since the elements } \beta\in\mathrm{Comp}\\ \text{satisfying } |\beta|=s|\alpha| \text{ are exactly}\\ \text{the elements of } \mathrm{Comp}_{s|\alpha|})}} c_\beta M_\beta + \sum_{\substack{\beta\in\mathrm{Comp};\\ |\beta|\neq s|\alpha|}} \underbrace{c_\beta}_{\substack{=0\\ (\text{by (13.179.28)},\\ \text{applied to } \gamma=\beta)}} M_\beta\\
&\qquad \left(\begin{array}{c}\text{since every } \beta\in\mathrm{Comp} \text{ satisfies exactly one of the two}\\ \text{statements } |\beta| = s\,|\alpha| \text{ and } |\beta|\neq s\,|\alpha|\end{array}\right)\\
&= \sum_{\beta\in\mathrm{Comp}_{s|\alpha|}} c_\beta M_\beta + \underbrace{\sum_{\substack{\beta\in\mathrm{Comp};\\ |\beta|\neq s|\alpha|}} 0 M_\beta}_{=0} = \sum_{\beta\in\mathrm{Comp}_{s|\alpha|}} c_\beta M_\beta\\
&= \sum_{\substack{\beta\in\mathrm{Comp}_{s|\alpha|};\\ \ell(\beta)\leq(s-1)\ell(\alpha)}} c_\beta M_\beta + \sum_{\substack{\beta\in\mathrm{Comp}_{s|\alpha|};\\ \ell(\beta)>(s-1)\ell(\alpha)}} \underbrace{c_\beta}_{\substack{=0\\ (\text{by (13.179.29)},\\ \text{applied to } \gamma=\beta)}} M_\beta\\
&\qquad \left(\begin{array}{c}\text{since every } \beta\in\mathrm{Comp}_{s|\alpha|} \text{ satisfies exactly one of the two}\\ \text{statements } \ell(\beta)\leq(s-1)\,\ell(\alpha) \text{ and } \ell(\beta)>(s-1)\,\ell(\alpha)\end{array}\right)\\
&= \sum_{\substack{\beta\in\mathrm{Comp}_{s|\alpha|};\\ \ell(\beta)\leq(s-1)\ell(\alpha)}} c_\beta M_\beta + \underbrace{\sum_{\substack{\beta\in\mathrm{Comp}_{s|\alpha|};\\ \ell(\beta)>(s-1)\ell(\alpha)}} 0 M_\beta}_{=0} = \sum_{\substack{\beta\in\mathrm{Comp}_{s|\alpha|};\\ \ell(\beta)\leq(s-1)\ell(\alpha)}} \underbrace{c_\beta}_{\in\mathbf{k}} M_\beta \in \sum_{\substack{\beta\in\mathrm{Comp}_{s|\alpha|};\\ \ell(\beta)\leq(s-1)\ell(\alpha)}} \mathbf{k}M_\beta.
\end{aligned}$$

Thus,

$$\begin{aligned}
\left|\mathrm{Supp}\left(\mathbf{x}_{\mathbf{i}_1}^\alpha \mathbf{x}_{\mathbf{i}_2}^\alpha \cdots \mathbf{x}_{\mathbf{i}_s}^\alpha\right)\right| &\leq \left|\bigcup_{j\in\{1,2,\ldots,s\}\setminus\{v\}} \mathrm{Supp}\left(\mathbf{x}_{\mathbf{i}_j}^\alpha\right)\right| \leq \sum_{j\in\{1,2,\ldots,s\}\setminus\{v\}} \underbrace{\left|\mathrm{Supp}\left(\mathbf{x}_{\mathbf{i}_j}^\alpha\right)\right|}_{\substack{=\ell\\ (\text{by (13.179.25), applied to } \mathbf{i}=\mathbf{i}_j)}}\\
&= \sum_{j\in\{1,2,\ldots,s\}\setminus\{v\}} \ell = \underbrace{|\{1,2,\ldots,s\}\setminus\{v\}|}_{=s-1} \underbrace{\ell}_{=\ell(\alpha)} = (s-1)\,\ell(\alpha).
\end{aligned}$$

Thus,

$$(s-1)\,\ell(\alpha) \geq \left|\mathrm{Supp}\left(\mathbf{x}_{\mathbf{i}_1}^\alpha \mathbf{x}_{\mathbf{i}_2}^\alpha \cdots \mathbf{x}_{\mathbf{i}_s}^\alpha\right)\right|. \tag{13.179.27}$$

But (13.179.24) (applied to $\beta = \gamma$) yields $|\mathrm{Supp}(\mathbf{x}^\gamma)| = \ell(\gamma) > (s-1)\,\ell(\alpha) \geq \left|\mathrm{Supp}\left(\mathbf{x}_{\mathbf{i}_1}^\alpha \mathbf{x}_{\mathbf{i}_2}^\alpha \cdots \mathbf{x}_{\mathbf{i}_s}^\alpha\right)\right|$ (by (13.179.27)). Hence, $|\mathrm{Supp}(\mathbf{x}^\gamma)| \neq \left|\mathrm{Supp}\left(\mathbf{x}_{\mathbf{i}_1}^\alpha \mathbf{x}_{\mathbf{i}_2}^\alpha \cdots \mathbf{x}_{\mathbf{i}_s}^\alpha\right)\right|$, so that $\mathbf{x}^\gamma \neq \mathbf{x}_{\mathbf{i}_1}^\alpha \mathbf{x}_{\mathbf{i}_2}^\alpha \cdots \mathbf{x}_{\mathbf{i}_s}^\alpha$. Hence, $\mathbf{x}^\gamma$ and $\mathbf{x}_{\mathbf{i}_1}^\alpha \mathbf{x}_{\mathbf{i}_2}^\alpha \cdots \mathbf{x}_{\mathbf{i}_s}^\alpha$ are two distinct monomials. Therefore, $\left(\text{the coefficient of the monomial } \mathbf{x}^\gamma \text{ in } \mathbf{x}_{\mathbf{i}_1}^\alpha \mathbf{x}_{\mathbf{i}_2}^\alpha \cdots \mathbf{x}_{\mathbf{i}_s}^\alpha\right) = 0$. This proves (13.179.22).

This proves Lemma 6.5.16 once again. □

13.180. **Solution to Exercise 6.5.20.** *Solution to Exercise 6.5.20.*

Proof of Corollary 6.5.19. We know that u is a Lyndon word and satisfies $u = u$. Hence, (u) is the CFL factorization of the word u.

Define the notion $\operatorname{mult}_w z$ (for any Lyndon word w and any word z) as in Theorem 6.2.2(b).

Theorem 6.2.2(e) yields that the lexicographically highest element of the multiset $u \sqcup\!\sqcup v$ is uv, and the multiplicity with which this word uv appears in the multiset $u \sqcup\!\sqcup v$ is $\operatorname{mult}_u v + 1$.

Recall that $\operatorname{mult}_u v$ is the number of terms in the CFL factorization of v which are equal to u (by the definition of $\operatorname{mult}_u v$). In other words,

$$\begin{aligned}\operatorname{mult}_u v &= \left(\text{the number of terms in } \underbrace{\text{the CFL factorization of } v}_{=(b_1,b_2,\ldots,b_q)} \text{ which are equal to } u\right)\\ &= \left(\text{the number of terms in } (b_1, b_2, \ldots, b_q) \text{ which are equal to } u\right)\\ &= \left(\text{the number of } j \in \{1,2,\ldots,q\} \text{ satisfying } b_j = u\right) = \left|\{j \in \{1,2,\ldots,q\} \mid b_j = u\}\right|.\end{aligned}$$

Now, the multiplicity with which the word uv appears in the multiset $u \sqcup\!\sqcup v$ is

$$\begin{aligned}\underbrace{\operatorname{mult}_u v}_{=|\{j\in\{1,2,\ldots,q\} \mid b_j=u\}|} +1 &= \left|\{j \in \{1,2,\ldots,q\} \mid b_j = u\}\right| + 1\\ &= 1 + \left|\{j \in \{1,2,\ldots,q\} \mid b_j = u\}\right| = h.\end{aligned}$$

Hence, we can apply Lemma 6.5.18 to $z = uv$. As a result, we conclude that

$$M_u M_v = h M_{uv} + \left(\text{a sum of terms of the form } M_w \text{ with } w \in \operatorname{Comp}_{n+m} \text{ satisfying } w \underset{\text{wll}}{<} uv\right).$$

This proves Corollary 6.5.19. □

13.181. **Solution to Exercise 6.5.22.** *Solution to Exercise 6.5.22.*

Proof of Corollary 6.5.21. Notice that $|x| = k$ (since $x \in \operatorname{Comp}_k$), thus $|x^s| = s \underbrace{|x|}_{=k} = sk$ and thus $x^s \in \operatorname{Comp}_{sk}$.

Clearly, (x) is the CFL factorization of the word x (since x is Lyndon). On the other hand, $\left(\underbrace{x, x, \ldots, x}_{s \text{ times}}\right)$ is the CFL factorization of the word x^s (since $\left(\underbrace{x, x, \ldots, x}_{s \text{ times}}\right)$ is a tuple of Lyndon words (since x is Lyndon) satisfying $x^s = \underbrace{xx \cdots x}_{s \text{ times}}$ and $x \geq x \geq \cdots \geq x$). Hence, Theorem 6.2.2(c) (applied to $u = x$, $v = x^s$, $p = 1$, $q = s$, $(a_1, a_2, \ldots, a_p) = (x)$ and $(b_1, b_2, \ldots, b_q) = \left(\underbrace{x, x, \ldots, x}_{s \text{ times}}\right)$) yields that the lexicographically highest element of the multiset $x \sqcup\!\sqcup x^s$ is xx^s (since $x \geq x$ for every $i \in \{1, 2, \ldots, 1\}$ and $j \in \{1, 2, \ldots, s\}$). In other words, the lexicographically highest element of the multiset $x \sqcup\!\sqcup x^s$ is x^{s+1} (since $xx^s = x^{s+1}$). This proves Corollary 6.5.21(a).

(c) Let $t \in \operatorname{Comp}_{sk}$ be such that $t \underset{\text{wll}}{<} x^s$. Then, Lemma 6.4.11(c) (applied to $n = k$, $m = sk$, $u = x$, $v = x^s$, $z = x^{s+1}$ and $v' = t$) yields

$$\begin{aligned} M_x M_t &= \left(\text{a sum of terms of the form } M_w \text{ with } w \in \underbrace{\operatorname{Comp}_{k+sk}}_{\substack{= \operatorname{Comp}_{(s+1)k} \\ (\text{since } k+sk=(s+1)k)}} \text{ satisfying } w \underset{\text{wll}}{<} x^{s+1} \right) \\ &= \left(\text{a sum of terms of the form } M_w \text{ with } w \in \operatorname{Comp}_{(s+1)k} \text{ satisfying } w \underset{\text{wll}}{<} x^{s+1} \right). \end{aligned}$$

This proves Corollary 6.5.21(c).

(b) Let $h = 1 + \left| \left\{ j \in \{1, 2, \ldots, s\} \mid x = x \right\} \right|$. Of course, $h = s + 1$ [1173]. Recall that x is a Lyndon word, and that $\left(\underbrace{x, x, \ldots, x}_{s \text{ times}} \right)$ is the CFL factorization of the word x^s. Notice also that $x \geq x$ for every $j \in \{1, 2, \ldots, s\}$. Thus, Corollary 6.5.19 (applied to $n = k$, $m = sk$, $u = x$, $v = x^s$, $q = s$ and $(b_1, b_2, \ldots, b_q) = \left(\underbrace{x, x, \ldots, x}_{s \text{ times}} \right)$) yields that

$$\begin{aligned} M_x M_{x^s} = \underbrace{h}_{=s+1} \underbrace{M_{xx^s}}_{\substack{= M_{x^{s+1}} \\ (\text{since } xx^s = x^{s+1})}} \\ + \left(\text{a sum of terms of the form } M_w \text{ with } w \in \underbrace{\operatorname{Comp}_{k+sk}}_{\substack{= \operatorname{Comp}_{(s+1)k} \\ (\text{since } k+sk=(s+1)k)}} \text{ satisfying } w \underset{\text{wll}}{<} \underbrace{xx^s}_{=x^{s+1}} \right) \\ = (s+1) M_{x^{s+1}} + \left(\text{a sum of terms of the form } M_w \text{ with } w \in \operatorname{Comp}_{(s+1)k} \text{ satisfying } w \underset{\text{wll}}{<} x^{s+1} \right). \end{aligned}$$

This proves Corollary 6.5.21(b). □

13.182. **Solution to Exercise 6.5.24.** *Solution to Exercise 6.5.24.*

Proof of Corollary 6.5.23. We assumed that $a_i > b_j$ for every $i \in \{1, 2, \ldots, p\}$ and $j \in \{1, 2, \ldots, q\}$. Thus, $a_i \geq b_j$ for every $i \in \{1, 2, \ldots, p\}$ and $j \in \{1, 2, \ldots, q\}$. Hence, the lexicographically highest element of the multiset $u \sqcup\!\sqcup v$ is uv (by Theorem 6.2.2(c)). Also, the multiplicity with which the word uv appears in the multiset $u \sqcup\!\sqcup v$ is 1 (by Theorem 6.2.2(d)). Hence, Lemma 6.5.18 (applied to $z = uv$ and $h = 1$) yields

$$\begin{aligned} M_u M_v &= \underbrace{1 M_{uv}}_{= M_{uv}} + \left(\text{a sum of terms of the form } M_w \text{ with } w \in \operatorname{Comp}_{n+m} \text{ satisfying } w \underset{\text{wll}}{<} uv \right) \\ &= M_{uv} + \left(\text{a sum of terms of the form } M_w \text{ with } w \in \operatorname{Comp}_{n+m} \text{ satisfying } w \underset{\text{wll}}{<} uv \right). \end{aligned}$$

This proves Corollary 6.5.23. □

[1173] *Proof.* Every $j \in \{1, 2, \ldots, s\}$ satisfies $x = x$. Hence, the set $\{ j \in \{1, 2, \ldots, s\} \mid x = x \}$ equals the whole $\{1, 2, \ldots, s\}$. In other words, $\{ j \in \{1, 2, \ldots, s\} \mid x = x \} = \{1, 2, \ldots, s\}$. Now, $h = 1 + \left| \underbrace{\{ j \in \{1, 2, \ldots, s\} \mid x = x \}}_{= \{1, 2, \ldots, s\}} \right| = 1 + \underbrace{|\{1, 2, \ldots, s\}|}_{=s} = 1 + s = s + 1$, qed.

13.183. **Solution to Exercise 6.5.26.** *Solution to Exercise 6.5.26.*

Proof of Corollary 6.5.25. The tuple $(a_1, a_2, \ldots, a_p)$ is the CFL factorization of u. Thus, $(a_1, a_2, \ldots, a_p)$ is a tuple of Lyndon words satisfying $u = a_1 a_2 \cdots a_p$ and $a_1 \geq a_2 \geq \cdots \geq a_p$.

Now, $(a_1, a_2, \ldots, a_k)$ is a tuple of Lyndon words (since $(a_1, a_2, \ldots, a_p)$ is a tuple of Lyndon words) satisfying $x = a_1 a_2 \cdots a_k$ and $a_1 \geq a_2 \geq \cdots \geq a_k$ (since $a_1 \geq a_2 \geq \cdots \geq a_p$). In other words, $(a_1, a_2, \ldots, a_k)$ is the CFL factorization of x (by the definition of a CFL factorization).

Also, $(a_{k+1}, a_{k+2}, \ldots, a_p)$ is a tuple of Lyndon words (since $(a_1, a_2, \ldots, a_p)$ is a tuple of Lyndon words) satisfying $y = a_{k+1} a_{k+2} \cdots a_p$ and $a_{k+1} \geq a_{k+2} \geq \cdots \geq a_p$ (since $a_1 \geq a_2 \geq \cdots \geq a_p$). In other words, $(a_{k+1}, a_{k+2}, \ldots, a_p)$ is the CFL factorization of y (by the definition of a CFL factorization).

We have $x \in \operatorname{Comp}_{|x|}$ and $y \in \operatorname{Comp}_{|y|}$. Also, multiplying the equalities $x = a_1 a_2 \cdots a_k$ and $y = a_{k+1} a_{k+2} \cdots a_p$, we obtain $xy = (a_1 a_2 \cdots a_k)(a_{k+1} a_{k+2} \cdots a_p) = a_1 a_2 \cdots a_p = u$. Thus, $|x| + |y| = \left| \underbrace{xy}_{=u} \right| = |u| = n$ (since $u \in \operatorname{Comp}_n$).

We have $a_i > a_{k+j}$ for every $i \in \{1, 2, \ldots, k\}$ and $j \in \{1, 2, \ldots, p-k\}$ [1174]. Hence, Corollary 6.5.23 (applied to x, y, $|x|$, $|y|$, k, $p-k$, $(a_1, a_2, \ldots, a_k)$ and $(a_{k+1}, a_{k+2}, \ldots, a_p)$ instead of u, v, n, m, p, q, $(a_1, a_2, \ldots, a_p)$ and $(b_1, b_2, \ldots, b_q)$) yields

$$M_x M_y = \underbrace{M_{xy}}_{\substack{=M_u \\ (\text{since } xy=u)}} + \left(\text{a sum of terms of the form } M_w \text{ with } w \in \underbrace{\operatorname{Comp}_{|x|+|y|}}_{\substack{=\operatorname{Comp}_n \\ (\text{since } |x|+|y|=n)}} \text{ satisfying } w \underset{\text{wll}}{<} \underbrace{xy}_{=u} \right)$$
$$= M_u + \left(\text{a sum of terms of the form } M_w \text{ with } w \in \operatorname{Comp}_n \text{ satisfying } w \underset{\text{wll}}{<} u \right).$$

Thus,

$$M_u = M_x M_y - \left(\text{a sum of terms of the form } M_w \text{ with } w \in \operatorname{Comp}_n \text{ satisfying } w \underset{\text{wll}}{<} u \right).$$

This proves Corollary 6.5.25. □

13.184. **Solution to Exercise 6.5.28.** *Solution to Exercise 6.5.28.*

Proof of Corollary 6.5.27. We shall prove Corollary 6.5.27 by induction over s:

Induction base: We have $x^0 = \varnothing$ and therefore $M_{x^0} = M_\varnothing = 1$. Now, $\underbrace{M_x^0}_{=1} - \underbrace{0!}_{=1} \underbrace{M_{x^0}}_{=1} = 1 - 1 = 0 \in \sum_{\substack{w \in \operatorname{Comp}_{0k}; \\ w \underset{\text{wll}}{<} x^0}} \mathbf{k} M_w$. In other words, Corollary 6.5.27 holds for $s = 0$. This completes the induction base.

Induction step: Let $\mathbf{S}$ be a nonnegative integer. Assume that Corollary 6.5.27 holds for $s = \mathbf{S}$. We need to prove that Corollary 6.5.27 holds for $s = \mathbf{S} + 1$.

Notice that $x \in \operatorname{Comp}_k$, thus $|x| = k$, and thus $\left| x^{\mathbf{S}} \right| = \mathbf{S} \underbrace{|x|}_{=k} = \mathbf{S} k$, so that $x^{\mathbf{S}} \in \operatorname{Comp}_{\mathbf{S}k}$.

[1174] *Proof.* Recall that $a_1 \geq a_2 \geq \cdots \geq a_p$. Thus,

$$a_s \geq a_t \qquad \text{for any } s \in \{1, 2, \ldots, p\} \text{ and } t \in \{1, 2, \ldots, p\} \text{ satisfying } s \leq t. \tag{13.183.1}$$

Let $i \in \{1, 2, \ldots, k\}$ and $j \in \{1, 2, \ldots, p-k\}$. Then, $i \leq k$ (since $i \in \{1, 2, \ldots, k\}$) and thus $a_i \geq a_k$ (by (13.183.1), applied to $s = i$ and $t = k$). Also, $j \geq 1$ (since $j \in \{1, 2, \ldots, p-k\}$), and thus $k + \underbrace{j}_{\geq 1} \geq k + 1$, so that $k + 1 \leq k + j$. Hence, $a_{k+1} \geq a_{k+j}$ (by (13.183.1), applied to $s = k+1$ and $t = k+j$). Now, $a_i \geq a_k > a_{k+1} \geq a_{k+j}$, qed.

We know that Corollary 6.5.27 holds for $s = \mathbf{S}$. In other words,

$$M_x^{\mathbf{S}} - \mathbf{S}! M_{x^{\mathbf{S}}} \in \sum_{\substack{w \in \operatorname{Comp}_{\mathbf{S}k}; \\ w \underset{\text{wll}}{<} x^{\mathbf{S}}}} \mathbf{k} M_w = \sum_{\substack{t \in \operatorname{Comp}_{\mathbf{S}k}; \\ t \underset{\text{wll}}{<} x^{\mathbf{S}}}} \mathbf{k} M_t \tag{13.184.1}$$

(here, we renamed the summation index w as t).

Using Corollary 6.5.21(c), it is easy to see that

$$M_x M_t \in \sum_{\substack{w \in \operatorname{Comp}_{(\mathbf{S}+1)k}; \\ w \underset{\text{wll}}{<} x^{\mathbf{S}+1}}} \mathbf{k} M_w \qquad \text{for every } t \in \operatorname{Comp}_{\mathbf{S}k} \text{ satisfying } t \underset{\text{wll}}{<} x^{\mathbf{S}}. \tag{13.184.2}$$

[1175] Now,

$$M_x \left(M_x^{\mathbf{S}} - \mathbf{S}! M_{x^{\mathbf{S}}}\right) = \underbrace{M_x M_x^{\mathbf{S}}}_{= M_x^{\mathbf{S}+1}} - \mathbf{S}! M_x M_{x^{\mathbf{S}}} = M_x^{\mathbf{S}+1} - \mathbf{S}! M_x M_{x^{\mathbf{S}}},$$

so that

$$\begin{aligned}
M_x^{\mathbf{S}+1} - \mathbf{S}! M_x M_{x^{\mathbf{S}}} &= M_x \underbrace{\left(M_x^{\mathbf{S}} - \mathbf{S}! M_{x^{\mathbf{S}}}\right)}_{\substack{\in \sum_{\substack{t \in \operatorname{Comp}_{\mathbf{S}k}; \\ t \underset{\text{wll}}{<} x^{\mathbf{S}}}} \mathbf{k} M_t \\ (\text{by } (13.184.1))}} \in M_x \left(\sum_{\substack{t \in \operatorname{Comp}_{\mathbf{S}k}; \\ t \underset{\text{wll}}{<} x^{\mathbf{S}}}} \mathbf{k} M_t \right) \\
&= \sum_{\substack{t \in \operatorname{Comp}_{\mathbf{S}k}; \\ t \underset{\text{wll}}{<} x^{\mathbf{S}}}} \mathbf{k} \underbrace{M_x M_t}_{\substack{\in \sum_{\substack{w \in \operatorname{Comp}_{(\mathbf{S}+1)k}; \\ w \underset{\text{wll}}{<} x^{\mathbf{S}+1}}} \mathbf{k} M_w \\ (\text{by } (13.184.2))}} \subset \sum_{\substack{t \in \operatorname{Comp}_{\mathbf{S}k}; \\ t \underset{\text{wll}}{<} x^{\mathbf{S}}}} \mathbf{k} \left(\sum_{\substack{w \in \operatorname{Comp}_{(\mathbf{S}+1)k}; \\ w \underset{\text{wll}}{<} x^{\mathbf{S}+1}}} \mathbf{k} M_w \right) \\
&\subset \sum_{\substack{w \in \operatorname{Comp}_{(\mathbf{S}+1)k}; \\ w \underset{\text{wll}}{<} x^{\mathbf{S}+1}}} \mathbf{k} M_w \qquad \left(\text{since } \sum_{\substack{w \in \operatorname{Comp}_{(\mathbf{S}+1)k}; \\ w \underset{\text{wll}}{<} x^{\mathbf{S}+1}}} \mathbf{k} M_w \text{ is a } \mathbf{k}\text{-module} \right).
\end{aligned} \tag{13.184.3}$$

Now, Corollary 6.5.21(b) (applied to $s = \mathbf{S}$) yields

$$M_x M_{x^{\mathbf{S}}} = (\mathbf{S}+1) M_{x^{\mathbf{S}+1}} + \left(\text{a sum of terms of the form } M_w \text{ with } w \in \operatorname{Comp}_{(\mathbf{S}+1)k} \text{ satisfying } w \underset{\text{wll}}{<} x^{\mathbf{S}+1} \right).$$

[1175] *Proof of (13.184.2):* Let $t \in \operatorname{Comp}_{\mathbf{S}k}$ be such that $t \underset{\text{wll}}{<} x^{\mathbf{S}}$. Then, Corollary 6.5.21(c) (applied to $s = \mathbf{S}$) yields

$$\begin{aligned}
M_x M_t &= \left(\text{a sum of terms of the form } M_w \text{ with } w \in \operatorname{Comp}_{(\mathbf{S}+1)k} \text{ satisfying } w \underset{\text{wll}}{<} x^{\mathbf{S}+1} \right) \\
&\in \sum_{\substack{w \in \operatorname{Comp}_{(\mathbf{S}+1)k}; \\ w \underset{\text{wll}}{<} x^{\mathbf{S}+1}}} \mathbf{k} M_w .
\end{aligned}$$

This proves (13.184.2).

Thus,

$$\begin{aligned}
&M_x M_{x^{\mathbf{S}}} - (\mathbf{S}+1)\, M_{x^{\mathbf{S}+1}} \\
&= \left(\text{a sum of terms of the form } M_w \text{ with } w \in \operatorname{Comp}_{(\mathbf{S}+1)k} \text{ satisfying } w \underset{\text{wll}}{<} x^{\mathbf{S}+1}\right) \\
&\in \sum_{\substack{w \in \operatorname{Comp}_{(\mathbf{S}+1)k};\\ w \underset{\text{wll}}{<} x^{\mathbf{S}+1}}} \mathbf{k} M_w . \qquad (13.184.4)
\end{aligned}$$

Now,

$$\begin{aligned}
M_x^{\mathbf{S}+1} - (\mathbf{S}+1)! M_{x^{\mathbf{S}+1}} &= \left(M_x^{\mathbf{S}+1} - \mathbf{S}! M_x M_{x^{\mathbf{S}}}\right) + \left(\mathbf{S}! M_x M_{x^{\mathbf{S}}} - \underbrace{(\mathbf{S}+1)!}_{=\mathbf{S}! \cdot (\mathbf{S}+1)} M_{x^{\mathbf{S}+1}}\right) \\
&= \left(M_x^{\mathbf{S}+1} - \mathbf{S}! M_x M_{x^{\mathbf{S}}}\right) + \underbrace{\left(\mathbf{S}! M_x M_{x^{\mathbf{S}}} - \mathbf{S}! \cdot (\mathbf{S}+1)\, M_{x^{\mathbf{S}+1}}\right)}_{=\mathbf{S}! \cdot \left(M_x M_{x^{\mathbf{S}}} - (\mathbf{S}+1) M_{x^{\mathbf{S}+1}}\right)} \\
&= \underbrace{\left(M_x^{\mathbf{S}+1} - \mathbf{S}! M_x M_{x^{\mathbf{S}}}\right)}_{\substack{\in \sum_{\substack{w \in \operatorname{Comp}_{(\mathbf{S}+1)k};\\ w \underset{\text{wll}}{<} x^{\mathbf{S}+1}}} \mathbf{k} M_w \\ \text{(by (13.184.3))}}} + \mathbf{S}! \cdot \underbrace{\left(M_x M_{x^{\mathbf{S}}} - (\mathbf{S}+1)\, M_{x^{\mathbf{S}+1}}\right)}_{\substack{\in \sum_{\substack{w \in \operatorname{Comp}_{(\mathbf{S}+1)k};\\ w \underset{\text{wll}}{<} x^{\mathbf{S}+1}}} \mathbf{k} M_w \\ \text{(by (13.184.4))}}} \\
&\in \sum_{\substack{w \in \operatorname{Comp}_{(\mathbf{S}+1)k};\\ w \underset{\text{wll}}{<} x^{\mathbf{S}+1}}} \mathbf{k} M_w + \mathbf{S}! \cdot \sum_{\substack{w \in \operatorname{Comp}_{(\mathbf{S}+1)k};\\ w \underset{\text{wll}}{<} x^{\mathbf{S}+1}}} \mathbf{k} M_w \subset \sum_{\substack{w \in \operatorname{Comp}_{(\mathbf{S}+1)k};\\ w \underset{\text{wll}}{<} x^{\mathbf{S}+1}}} \mathbf{k} M_w
\end{aligned}$$

(since $\sum_{\substack{w \in \operatorname{Comp}_{(\mathbf{S}+1)k};\\ w \underset{\text{wll}}{<} x^{\mathbf{S}+1}}} \mathbf{k} M_w$ is a $\mathbf{k}$-module). In other words, Corollary 6.5.27 holds for $s = \mathbf{S}+1$. This completes the induction step. Thus, Corollary 6.5.27 is proven by induction. □

13.185. **Solution to Exercise 6.5.30.** *Solution to Exercise 6.5.30.*

Proof of Corollary 6.5.29. We first assume that $\mathbf{k} = \mathbb{Z}$.

Notice that $|x| = k$ (since $x \in \operatorname{Comp}_k$). Also, $M_x^{\langle s \rangle} \in \operatorname{QSym}$ (by Corollary 6.5.8(a), applied to $\alpha = x$) and thus $\underbrace{M_x^{\langle s \rangle}}_{\in \operatorname{QSym}} - \underbrace{M_{x^s}}_{\in \operatorname{QSym}} \in \operatorname{QSym} - \operatorname{QSym} \subset \operatorname{QSym}$.

Corollary 6.5.27 yields

$$M_x^s - s! M_{x^s} \in \sum_{\substack{w \in \operatorname{Comp}_{sk};\\ w \underset{\text{wll}}{<} x^s}} \mathbf{k} M_w . \qquad (13.185.1)$$

On the other hand, the composition x is Lyndon and therefore nonempty. Hence, (6.5.2) (applied to $\alpha = x$) yields

$$s! M_x^{\langle s \rangle} - M_x^s \in \underbrace{\sum_{\substack{\beta \in \operatorname{Comp}_{s|x|};\\ \ell(\beta) \leq (s-1)\ell(x)}} \mathbf{k} M_\beta}_{\substack{= \sum_{\substack{\beta \in \operatorname{Comp}_{sk};\\ \ell(\beta) \leq (s-1)\ell(x)}} \\ \text{(since } |x| = k)}} = \sum_{\substack{\beta \in \operatorname{Comp}_{sk};\\ \ell(\beta) \leq (s-1)\ell(x)}} \mathbf{k} M_\beta = \sum_{\substack{w \in \operatorname{Comp}_{sk};\\ \ell(w) \leq (s-1)\ell(x)}} \mathbf{k} M_w$$

(here, we renamed the summation index β as w). But every $w \in \operatorname{Comp}_{sk}$ satisfying $\ell(w) \leq (s-1)\ell(x)$ must also satisfy $w \underset{\text{wll}}{<} x^s$ [1176]. Hence, the sum $\sum_{\substack{w \in \operatorname{Comp}_{sk};\\ \ell(w) \leq (s-1)\ell(x)}} \mathbf{k} M_w$ is a subsum of the sum $\sum_{\substack{w \in \operatorname{Comp}_{sk};\\ w \underset{\text{wll}}{<} x^s}} \mathbf{k} M_w$. Thus, $\sum_{\substack{w \in \operatorname{Comp}_{sk};\\ \ell(w) \leq (s-1)\ell(x)}} \mathbf{k} M_w \subset \sum_{\substack{w \in \operatorname{Comp}_{sk};\\ w \underset{\text{wll}}{<} x^s}} \mathbf{k} M_w$. Thus,

$$s! M_x^{\langle s\rangle} - M_x^s \in \sum_{\substack{w \in \operatorname{Comp}_{sk};\\ \ell(w) \leq (s-1)\ell(x)}} \mathbf{k} M_w \subset \sum_{\substack{w \in \operatorname{Comp}_{sk};\\ w \underset{\text{wll}}{<} x^s}} \mathbf{k} M_w. \tag{13.185.2}$$

Now,

$$s!\left(M_x^{\langle s\rangle} - M_{x^s}\right) = s! M_x^{\langle s\rangle} - s! M_{x^s} = \underbrace{\left(s! M_x^{\langle s\rangle} - M_x^s\right)}_{\substack{\in \sum_{\substack{w \in \operatorname{Comp}_{sk};\\ w \underset{\text{wll}}{<} x^s}} \mathbf{k} M_w \\ \text{(by (13.185.2))}}} + \underbrace{\left(M_x^s - s! M_{x^s}\right)}_{\substack{\in \sum_{\substack{w \in \operatorname{Comp}_{sk};\\ w \underset{\text{wll}}{<} x^s}} \mathbf{k} M_w \\ \text{(by (13.185.1))}}}$$

$$\in \sum_{\substack{w \in \operatorname{Comp}_{sk};\\ w \underset{\text{wll}}{<} x^s}} \mathbf{k} M_w + \sum_{\substack{w \in \operatorname{Comp}_{sk};\\ w \underset{\text{wll}}{<} x^s}} \mathbf{k} M_w \subset \sum_{\substack{w \in \operatorname{Comp}_{sk};\\ w \underset{\text{wll}}{<} x^s}} \mathbf{k} M_w \tag{13.185.3}$$

(since $\sum_{\substack{w \in \operatorname{Comp}_{sk};\\ w \underset{\text{wll}}{<} x^s}} \mathbf{k} M_w$ is a $\mathbf{k}$-module).

But recall that we assumed that $\mathbf{k} = \mathbb{Z}$. Hence, if N is a positive integer and if f is an element of QSym satisfying $Nf \in \sum_{\substack{w \in \operatorname{Comp}_{sk};\\ w \underset{\text{wll}}{<} x^s}} \mathbf{k} M_w$, then $f \in \sum_{\substack{w \in \operatorname{Comp}_{sk};\\ w \underset{\text{wll}}{<} x^s}} \mathbf{k} M_w$ [1177]. Applying this to $N = s!$ and

[1176] *Proof.* Let $w \in \operatorname{Comp}_{sk}$. Notice that $|x^s| = s \underbrace{|x|}_{=k} = sk$, so that $x^s \in \operatorname{Comp}_{sk}$. Now, we have $\ell(x) > 0$ (since x is nonempty), and

$$\ell(x^s) = s\ell(x) = (s-1)\ell(x) + \underbrace{\ell(x)}_{>0} > (s-1)\ell(x).$$

Thus, $(s-1)\ell(x) < \ell(x^s)$, so that $\ell(w) \leq (s-1)\ell(x) < \ell(x^s)$.

But the definition of the wll-order shows that if $n \in \mathbb{N}$, and if α and β are two elements of Comp_n satisfying $\ell(\alpha) < \ell(\beta)$, then $\alpha \underset{\text{wll}}{<} \beta$. Applying this to $\alpha = w$ and $\beta = x^s$, we obtain $w \underset{\text{wll}}{<} x^s$ (since $\ell(w) < \ell(x^s)$), qed.

[1177] *Proof.* Let N be a positive integer. Let f be an element of QSym satisfying $Nf \in \sum_{\substack{w \in \operatorname{Comp}_{sk};\\ w \underset{\text{wll}}{<} x^s}} \mathbf{k} M_w$.

We know that $(M_\alpha)_{\alpha \in \operatorname{Comp}}$ is a basis of the $\mathbf{k}$-module QSym. In the following, whenever $g \in \operatorname{QSym}$ and $\beta \in \operatorname{Comp}$, we will let $\operatorname{coord}_{M_\beta} g$ denote the M_β-coordinate of g with respect to the basis $(M_\alpha)_{\alpha \in \operatorname{Comp}}$ of QSym. Then, every $g \in \operatorname{QSym}$ satisfies

$$g = \sum_{\beta \in \operatorname{Comp}} \left(\operatorname{coord}_{M_\beta} g\right) M_\beta \tag{13.185.4}$$

(by the definition of coordinates). Notice also that

$$\operatorname{coord}_{M_\beta}(M_\gamma) = \delta_{\beta,\gamma} \qquad \text{for every compositions } \beta \text{ and } \gamma. \tag{13.185.5}$$

Now, $Nf \in \sum_{\substack{w \in \operatorname{Comp}_{sk};\\ w \underset{\text{wll}}{<} x^s}} \mathbf{k} M_w$. Thus, there exists a family $(\lambda_w)_{\substack{w \in \operatorname{Comp}_{sk};\\ w \underset{\text{wll}}{<} x^s}} \in \mathbf{k}^{\left\{w \in \operatorname{Comp}_{sk} \ \mid \ w \underset{\text{wll}}{<} x^s\right\}}$ of elements of $\mathbf{k}$ satisfying $Nf = \sum_{\substack{w \in \operatorname{Comp}_{sk};\\ w \underset{\text{wll}}{<} x^s}} \lambda_w M_w$. Consider this family $(\lambda_w)_{\substack{w \in \operatorname{Comp}_{sk};\\ w \underset{\text{wll}}{<} x^s}}$.

Let β be a composition such that we don't have $\left(\beta \in \operatorname{Comp}_{sk} \text{ and } \beta \underset{\text{wll}}{<} x^s\right)$. Then,

$$\text{every composition } w \in \operatorname{Comp}_{sk} \text{ satisfying } w \underset{\text{wll}}{<} x^s \text{ satisfies } \beta \neq w \tag{13.185.6}$$

$f = M_x^{\langle s \rangle} - M_{x^s}$, we obtain $M_x^{\langle s \rangle} - M_{x^s} \in \sum\limits_{\substack{w \in \operatorname{Comp}_{sk}; \\ w \underset{\mathrm{wll}}{<} x^s}} \mathbf{k} M_w$ (because of (13.185.3)). In other words, Corollary 6.5.29 holds under the assumption that $\mathbf{k} = \mathbb{Z}$.

Now, let us forget that we have assumed that $\mathbf{k} = \mathbb{Z}$. We thus have shown that Corollary 6.5.29 holds under the assumption that $\mathbf{k} = \mathbb{Z}$. In other words, we have shown that the relation

$$M_x^{\langle s \rangle} - M_{x^s} \in \sum_{\substack{w \in \operatorname{Comp}_{sk}; \\ w \underset{\mathrm{wll}}{<} x^s}} \mathbb{Z} M_w \tag{13.185.8}$$

holds in $\mathbb{Z}[[\mathbf{x}]]$.

Now, recall that there is a canonical ring homomorphism $\varphi : \mathbb{Z} \to \mathbf{k}$. This homomorphism gives rise to a ring homomorphism $\varphi[[\mathbf{x}]] : \mathbb{Z}[[\mathbf{x}]] \to \mathbf{k}[[\mathbf{x}]]$, and this latter homomorphism $\varphi[[\mathbf{x}]]$ has the properties that:

- it is $\mathbb{Z}$-linear;
- it sends the element M_α of $\mathbb{Z}[[\mathbf{x}]]$ to the element M_α of $\mathbf{k}[[\mathbf{x}]]$ for every composition α;

(since every such w satisfies $\left(w \in \operatorname{Comp}_{sk} \text{ and } w \underset{\mathrm{wll}}{<} x^s \right)$, whereas β does not satisfy $\left(\beta \in \operatorname{Comp}_{sk} \text{ and } \beta \underset{\mathrm{wll}}{<} x^s \right)$). Now,

$$\operatorname{coord}_{M_\beta} \left(\underbrace{Nf}_{= \sum\limits_{\substack{w \in \operatorname{Comp}_{sk}; \\ w \underset{\mathrm{wll}}{<} x^s}} \lambda_w M_w} \right) = \operatorname{coord}_{M_\beta} \left(\sum_{\substack{w \in \operatorname{Comp}_{sk}; \\ w \underset{\mathrm{wll}}{<} x^s}} \lambda_w M_w \right) = \sum_{\substack{w \in \operatorname{Comp}_{sk}; \\ w \underset{\mathrm{wll}}{<} x^s}} \lambda_w \underbrace{\operatorname{coord}_{M_\beta}(M_w)}_{\substack{= \delta_{\beta, w} \\ \text{(by (13.185.5), applied} \\ \text{to } \gamma = w)}}$$
$$= \sum_{\substack{w \in \operatorname{Comp}_{sk}; \\ w \underset{\mathrm{wll}}{<} x^s}} \lambda_w \underbrace{\delta_{\beta, w}}_{\substack{=0 \\ \text{(since } \beta \neq w \\ \text{(by (13.185.6)))}}} = \sum_{\substack{w \in \operatorname{Comp}_{sk}; \\ w \underset{\mathrm{wll}}{<} x^s}} \lambda_w 0 = 0.$$

Compared with $\operatorname{coord}_{M_\beta}(Nf) = N \operatorname{coord}_{M_\beta} f$, this yields $N \operatorname{coord}_{M_\beta} f = 0$.

But N is a positive integer. Thus, if an element ρ of $\mathbb{Z}$ satisfies $N\rho = 0$, then $\rho = 0$. Applying this to $\rho = \operatorname{coord}_{M_\beta} f$, we obtain $\operatorname{coord}_{M_\beta} f = 0$ (since $\operatorname{coord}_{M_\beta} f \in \mathbf{k} = \mathbb{Z}$ and $N \operatorname{coord}_{M_\beta} f = 0$).

Now, let us forget that we fixed β. We thus have shown that if β is a composition such that we don't have $\left(\beta \in \operatorname{Comp}_{sk} \text{ and } \beta \underset{\mathrm{wll}}{<} x^s \right)$, then

$$\operatorname{coord}_{M_\beta} f = 0. \tag{13.185.7}$$

Now, (13.185.4) (applied to $g = f$) yields

$$f = \sum_{\beta \in \operatorname{Comp}} \left(\operatorname{coord}_{M_\beta} f \right) M_\beta = \underbrace{\sum_{\substack{\beta \in \operatorname{Comp}; \\ \beta \in \operatorname{Comp}_{sk} \text{ and } \beta \underset{\mathrm{wll}}{<} x^s}}}_{\substack{= \sum\limits_{\substack{\beta \in \operatorname{Comp}_{sk}; \\ \beta \underset{\mathrm{wll}}{<} x^s}} \\ \text{(since every } \beta \in \operatorname{Comp}_{sk} \\ \text{satisfies } \beta \in \operatorname{Comp})}} \left(\operatorname{coord}_{M_\beta} f \right) M_\beta + \sum_{\substack{\beta \in \operatorname{Comp}; \\ \text{we don't have } \left(\beta \in \operatorname{Comp}_{sk} \text{ and } \beta \underset{\mathrm{wll}}{<} x^s \right)}} \underbrace{\left(\operatorname{coord}_{M_\beta} f \right)}_{\substack{=0 \\ \text{(by (13.185.7))}}} M_\beta$$
$$= \sum_{\substack{\beta \in \operatorname{Comp}_{sk}; \\ \beta \underset{\mathrm{wll}}{<} x^s}} \left(\operatorname{coord}_{M_\beta} f \right) M_\beta + \underbrace{\sum_{\substack{\beta \in \operatorname{Comp}; \\ \text{we don't have } \left(\beta \in \operatorname{Comp}_{sk} \text{ and } \beta \underset{\mathrm{wll}}{<} x^s \right)}} 0 M_\beta}_{=0} = \sum_{\substack{\beta \in \operatorname{Comp}_{sk}; \\ \beta \underset{\mathrm{wll}}{<} x^s}} \left(\operatorname{coord}_{M_\beta} f \right) M_\beta$$
$$= \sum_{\substack{w \in \operatorname{Comp}_{sk}; \\ w \underset{\mathrm{wll}}{<} x^s}} \underbrace{\left(\operatorname{coord}_{M_w} f \right)}_{\in \mathbf{k}} M_w \qquad \text{(here, we renamed the summation index } \beta \text{ as } w\text{)}$$
$$\in \sum_{\substack{w \in \operatorname{Comp}_{sk}; \\ w \underset{\mathrm{wll}}{<} x^s}} \mathbf{k} M_w,$$

qed.

- it sends the element $M_\alpha^{\langle s\rangle}$ of $\mathbb{Z}[[\mathbf{x}]]$ to the element $M_\alpha^{\langle s\rangle}$ of $\mathbf{k}[[\mathbf{x}]]$ for every composition α and every nonnegative integer s.

Hence, by applying the homomorphism $\varphi[[\mathbf{x}]]$ to both sides of (13.185.8), we obtain

$$M_x^{\langle s\rangle} - M_{x^s} \in \sum_{\substack{w\in\operatorname{Comp}_{sk};\\ w \underset{\text{wll}}{<} x^s}} \underbrace{\mathbb{Z}M_w}_{\substack{\subset \mathbf{k}M_w\\ \text{(due to the canonical}\\ \text{ring homomorphism } \mathbb{Z}\to\mathbf{k})}} \subset \sum_{\substack{w\in\operatorname{Comp}_{sk};\\ w \underset{\text{wll}}{<} x^s}} \mathbf{k}M_w .$$

This proves Corollary 6.5.29. □

13.186. **Solution to Exercise 6.5.31.** *Solution to Exercise 6.5.31.*

Proof of Theorem 6.5.13. The family $\left(M_w^{\langle s\rangle}\right)_{(w,s)\in\mathfrak{RL}\times\{1,2,3,\ldots\}}$ is a reindexing of the family $\left(M_{\operatorname{red}\alpha}^{\langle\gcd\alpha\rangle}\right)_{\alpha\in\mathfrak{L}}$ (according to Lemma 6.5.14). In other words, the family $\left(M_w^{\langle s\rangle}\right)_{(w,s)\in\mathfrak{RL}\times\{1,2,3,\ldots\}}$ is a reindexing of the family $\left(M_{\operatorname{red}w}^{\langle\gcd w\rangle}\right)_{w\in\mathfrak{L}}$ (here, we renamed the index α as w).

We need to show that the family $\left(M_w^{\langle s\rangle}\right)_{(w,s)\in\mathfrak{RL}\times\{1,2,3,\ldots\}}$ is an algebraically independent generating set of the $\mathbf{k}$-algebra QSym. It is enough to prove that the family $\left(M_{\operatorname{red}w}^{\langle\gcd w\rangle}\right)_{w\in\mathfrak{L}}$ is an algebraically independent generating set of the $\mathbf{k}$-algebra QSym (since the family $\left(M_w^{\langle s\rangle}\right)_{(w,s)\in\mathfrak{RL}\times\{1,2,3,\ldots\}}$ is a reindexing of the family $\left(M_{\operatorname{red}w}^{\langle\gcd w\rangle}\right)_{w\in\mathfrak{L}}$). We shall prove the latter claim.

Indeed, the main difficulty is to show that

$$\text{the family } \left(M_{\operatorname{red}w}^{\langle\gcd w\rangle}\right)_{w\in\mathfrak{L}} \text{ generates the } \mathbf{k}\text{-algebra QSym}. \tag{13.186.1}$$

Once (13.186.1) is proven, we will be able to complete the proof of Theorem 6.5.13 as follows:

For every $w\in\mathfrak{L}$, we have $M_{\operatorname{red}w}^{\langle\gcd w\rangle}\in\operatorname{QSym}$ [1178].

Let $\operatorname{wt}:\mathfrak{A}\to\{1,2,3,\ldots\}$ be the identity map (this is well-defined since $\mathfrak{A}=\{1,2,3,\ldots\}$). Obviously, for every $N\in\{1,2,3,\ldots\}$, the set $\operatorname{wt}^{-1}(N)$ is finite.

For every word $w\in\mathfrak{A}^*$, define an element $\operatorname{Wt}(w)\in\mathbb{N}$ by $\operatorname{Wt}(w)=\operatorname{wt}(w_1)+\operatorname{wt}(w_2)+\cdots+\operatorname{wt}(w_k)$, where k is the length of w. Then,

$$\text{every } w\in\mathfrak{A}^* \text{ satisfies } \operatorname{Wt}(w)=|w| \tag{13.186.2}$$

[1179].

For every $w\in\mathfrak{L}$, the element $M_{\operatorname{red}w}^{\langle\gcd w\rangle}$ of QSym is homogeneous of degree $\operatorname{Wt}(w)$ [1180].

The $\mathbf{k}$-module QSym has a basis $(g_u)_{u\in\mathfrak{A}^*}$ having the property that for every $u\in\mathfrak{A}^*$, the element g_u of QSym is homogeneous of degree $\operatorname{Wt}(u)$ [1181].

[1178] *Proof.* Let $w\in\mathfrak{L}$. Then, Corollary 6.5.8(a) (applied to $\gcd w$ and $\operatorname{red}w$ instead of s and w) yields $M_{\operatorname{red}w}^{\langle\gcd w\rangle}\in\operatorname{QSym}$, qed.

[1179] *Proof of (13.186.2):* Let $w\in\mathfrak{A}^*$. Let k be the length of w. Then, $w=(w_1,w_2,\ldots,w_k)$, so that $|w|=w_1+w_2+\cdots+w_k$. Recall that wt is the identity map. In other words, $\operatorname{wt}=\operatorname{id}$. Now, the definition of $\operatorname{Wt}(w)$ yields

$$\begin{aligned}\operatorname{Wt}(w)&=\operatorname{wt}(w_1)+\operatorname{wt}(w_2)+\cdots+\operatorname{wt}(w_k)=\operatorname{id}(w_1)+\operatorname{id}(w_2)+\cdots+\operatorname{id}(w_k)\qquad(\text{since }\operatorname{wt}=\operatorname{id})\\&=w_1+w_2+\cdots+w_k=|w|,\end{aligned}$$

and thus (13.186.2) is proven.

[1180] *Proof.* Let $w\in\mathfrak{L}$. Then, w is a Lyndon word (since $\mathfrak{L}$ is the set of all Lyndon words), hence nonempty. Remark 6.5.11(f) (applied to $\alpha=w$) now yields $(\gcd w)|\operatorname{red}w|=|w|=\operatorname{Wt}(w)$ (by (13.186.2)). But Corollary 6.5.8(b) (applied to $\gcd w$ and $\operatorname{red}w$ instead of s and w) yields $M_{\operatorname{red}w}^{\langle\gcd w\rangle}\in\operatorname{QSym}_{(\gcd w)|\operatorname{red}w|}=\operatorname{QSym}_{\operatorname{Wt}(w)}$ (since $(\gcd w)|\operatorname{red}w|=\operatorname{Wt}(w)$). In other words, the element $M_{\operatorname{red}w}^{\langle\gcd w\rangle}$ of QSym is homogeneous of degree $\operatorname{Wt}(w)$, qed.

[1181] *Proof.* We shall show that $(M_u)_{u\in\mathfrak{A}^*}$ is such a basis.

Once (13.186.1) is proven, it thus follows that we can apply Lemma 6.3.7(c) to $A = \operatorname{QSym}$ and $b_w = M_{\operatorname{red} w}^{\langle \gcd w \rangle}$. From this, we can conclude that the family $\left(M_{\operatorname{red} w}^{\langle \gcd w \rangle}\right)_{w \in \mathfrak{L}}$ is an algebraically independent generating set of the $\mathbf{k}$-algebra QSym (provided that (13.186.1) is proven); this is precisely what we need to prove.

Hence, in order to complete the proof of Theorem 6.5.13, it is sufficient to prove (13.186.1). So we shall now prove (13.186.1).

Let U denote the $\mathbf{k}$-subalgebra of QSym generated by the family $\left(M_{\operatorname{red} w}^{\langle \gcd w \rangle}\right)_{w \in \mathfrak{L}}$. Then, U is a $\mathbf{k}$-submodule of QSym. It is clear that

$$M_\beta^{\langle s \rangle} \in U \qquad \text{for every reduced Lyndon composition } \beta \text{ and every } s \in \{1, 2, 3, \ldots\}. \tag{13.186.3}$$

[1182] Using this and using Exercise 6.5.4(d), it is now easy to see that

$$M_\beta^{\langle s \rangle} \in U \qquad \text{for every Lyndon composition } \beta \text{ and every } s \in \{1, 2, 3, \ldots\}. \tag{13.186.4}$$

[1183]

We shall now prove that

$$M_\beta \in U \qquad \text{for every composition } \beta. \tag{13.186.5}$$

Proof of (13.186.5): We will prove (13.186.5) by strong induction over $|\beta|$:

Indeed, $\mathfrak{A}^* = \operatorname{Comp}$, so that $(M_u)_{u \in \mathfrak{A}^*} = (M_u)_{u \in \operatorname{Comp}}$ is clearly a basis of the $\mathbf{k}$-module QSym. Furthermore, for every $u \in \mathfrak{A}^*$, the element M_u of QSym is homogeneous of degree $|u|$. Since $\operatorname{Wt}(u) = |u|$ for every $u \in \mathfrak{A}^*$ (by (13.186.2), applied to $w = u$), this rewrites as follows: For every $u \in \mathfrak{A}^*$, the element M_u of QSym is homogeneous of degree $\operatorname{Wt}(u)$.

Thus, $(M_u)_{u \in \mathfrak{A}^*}$ is a basis of the $\mathbf{k}$-module QSym having the property that for every $u \in \mathfrak{A}^*$, the element M_u of QSym is homogeneous of degree $\operatorname{Wt}(u)$. Hence, the $\mathbf{k}$-module QSym has a basis $(g_u)_{u \in \mathfrak{A}^*}$ having the property that for every $u \in \mathfrak{A}^*$, the element g_u of QSym is homogeneous of degree $\operatorname{Wt}(u)$ (namely, $(M_u)_{u \in \mathfrak{A}^*}$ is such a basis), qed.

[1182] *Proof of (13.186.3):* Let β be a reduced Lyndon composition. Let $s \in \{1, 2, 3, \ldots\}$.

The composition β is Lyndon and thus nonempty. Remark 6.5.11(e) (applied to $\alpha = \beta$) yields that the composition $\beta\{s\}$ is nonempty and satisfies $\operatorname{red}(\beta\{s\}) = \operatorname{red}\beta$ and $\gcd(\beta\{s\}) = s \gcd \beta$. But β is reduced; thus, $\gcd \beta = 1$ (by the definition of "reduced"). Hence, $\gcd(\beta\{s\}) = s \underbrace{\gcd \beta}_{=1} = s$. Also, Remark 6.5.11(d) (applied to $\alpha = \beta$) yields $\operatorname{red}\beta = \beta$, so that $\operatorname{red}(\beta\{s\}) = \operatorname{red}\beta = \beta$.

Remark 6.5.11(b) (applied to $\alpha = \beta\{s\}$) shows that the composition $\beta\{s\}$ is Lyndon if and only if the composition $\operatorname{red}(\beta\{s\})$ is Lyndon. Since the composition $\operatorname{red}(\beta\{s\}) = \beta$ is Lyndon, this yields that the composition $\beta\{s\}$ is Lyndon. In other words, $\beta\{s\} \in \mathfrak{L}$ (since $\mathfrak{L}$ is the set of all Lyndon words). Hence, $M_{\operatorname{red}(\beta\{s\})}^{\langle \gcd(\beta\{s\}) \rangle}$ is an element of the family $\left(M_{\operatorname{red} w}^{\langle \gcd w \rangle}\right)_{w \in \mathfrak{L}}$ (namely, the element for $w = \beta\{s\}$). In other words, $M_\beta^{\langle s \rangle}$ is an element of the family $\left(M_{\operatorname{red} w}^{\langle \gcd w \rangle}\right)_{w \in \mathfrak{L}}$ (since $\operatorname{red}(\beta\{s\}) = \beta$ and $\gcd(\beta\{s\}) = s$). Hence, $M_\beta^{\langle s \rangle}$ belongs to the $\mathbf{k}$-subalgebra of QSym generated by this family $\left(M_{\operatorname{red} w}^{\langle \gcd w \rangle}\right)_{w \in \mathfrak{L}}$. Since the $\mathbf{k}$-subalgebra of QSym generated by the family $\left(M_{\operatorname{red} w}^{\langle \gcd w \rangle}\right)_{w \in \mathfrak{L}}$ is U, this rewrites as follows: $M_\beta^{\langle s \rangle}$ belongs to U. In other words, $M_\beta^{\langle s \rangle} \in U$. This proves (13.186.3).

[1183] *Proof.* Let β be a Lyndon composition, and let $s \in \{1, 2, 3, \ldots\}$. The composition β is Lyndon and thus nonempty; hence, $\gcd \beta$ and $\operatorname{red}\beta$ are well-defined.

Remark 6.5.11(b) (applied to $\alpha = \beta$) yields that the composition β is Lyndon if and only if the composition $\operatorname{red}\beta$ is Lyndon. Since the composition β is Lyndon, we therefore conclude that the composition $\operatorname{red}\beta$ is Lyndon. Also, the composition $\operatorname{red}\beta$ is reduced (by Remark 6.5.11(c)). Thus, $\operatorname{red}\beta$ is a reduced Lyndon composition. Remark 6.5.11(a) (applied to $\alpha = \beta$) yields $\beta = (\operatorname{red}\beta)\{\gcd\beta\}$, so that $(\operatorname{red}\beta)\{\gcd\beta\} = \beta$.

Let $\alpha = \operatorname{red}\beta$. Recall that $\operatorname{red}\beta$ is a reduced Lyndon composition. In other words, α is a reduced Lyndon composition (since $\alpha = \operatorname{red}\beta$).

Exercise 6.5.4(d) (applied to $n = \gcd\beta$) yields that there exists a polynomial $P \in \mathbf{k}[z_1, z_2, z_3, \ldots]$ such that $M_{\alpha\{\gcd\beta\}}^{\langle s \rangle} = P\left(M_\alpha^{\langle 1 \rangle}, M_\alpha^{\langle 2 \rangle}, M_\alpha^{\langle 3 \rangle}, \ldots\right)$. Consider this P. We have $\underbrace{\alpha}_{=\operatorname{red}\beta}\{\gcd\beta\} = (\operatorname{red}\beta)\{\gcd\beta\} = \beta$. Hence, $M_{\alpha\{\gcd\beta\}}^{\langle s \rangle} = P\left(M_\alpha^{\langle 1 \rangle}, M_\alpha^{\langle 2 \rangle}, M_\alpha^{\langle 3 \rangle}, \ldots\right)$ rewrites as $M_\beta^{\langle s \rangle} = P\left(M_\alpha^{\langle 1 \rangle}, M_\alpha^{\langle 2 \rangle}, M_\alpha^{\langle 3 \rangle}, \ldots\right)$.

However, for every $j \in \{1, 2, 3, \ldots\}$, we have $M_\alpha^{\langle j \rangle} \in U$ (by (13.186.3), applied to α and j instead of β and s). In other words, $M_\alpha^{\langle 1 \rangle}$, $M_\alpha^{\langle 2 \rangle}$, $M_\alpha^{\langle 3 \rangle}$, $\ldots$ are elements of U. Therefore, $Q\left(M_\alpha^{\langle 1 \rangle}, M_\alpha^{\langle 2 \rangle}, M_\alpha^{\langle 3 \rangle}, \ldots\right) \in U$ for every $Q \in \mathbf{k}[z_1, z_2, z_3, \ldots]$ (since U is a $\mathbf{k}$-subalgebra of QSym). Applied to $Q = P$, this yields $P\left(M_\alpha^{\langle 1 \rangle}, M_\alpha^{\langle 2 \rangle}, M_\alpha^{\langle 3 \rangle}, \ldots\right) \in U$. Thus, $M_\beta^{\langle s \rangle} = P\left(M_\alpha^{\langle 1 \rangle}, M_\alpha^{\langle 2 \rangle}, M_\alpha^{\langle 3 \rangle}, \ldots\right) \in U$. This proves (13.186.4).

Induction step: Let $N \in \mathbb{N}$. We assume that (13.186.5) holds for all compositions β satisfying $|\beta| < N$. We need to show that (13.186.5) also holds for all compositions β satisfying $|\beta| = N$. In other words, we need to prove that $M_\beta \in U$ for every composition β satisfying $|\beta| = N$. In other words, we need to prove that

$$M_\beta \in U \qquad \text{for every } \beta \in \operatorname{Comp}_N . \tag{13.186.6}$$

Proof of (13.186.6): We will prove (13.186.6) by strong induction over β with respect to the wll-order on Comp_N. In other words, we fix some $\alpha \in \operatorname{Comp}_N$, and we assume that (13.186.6) holds for all $\beta \in \operatorname{Comp}_N$ satisfying $\beta \underset{\text{wll}}{<} \alpha$. We now need to prove that (13.186.6) holds for $\beta = \alpha$. In other words, we need to prove that $M_\alpha \in U$.

If $\alpha = \varnothing$, then $M_\alpha \in U$ is obvious[1184]. Hence, for the rest of this proof of $M_\alpha \in U$, we can WLOG assume that $\alpha \neq \varnothing$. Assume this. The composition α is nonempty (since $\alpha \neq \varnothing$), so that it satisfies $|\alpha| \neq 0$. Since $|\alpha| = N$ (because $\alpha \in \operatorname{Comp}_N$), we have $N = |\alpha| \neq 0$. Thus, N is a positive integer.

We have assumed that (13.186.6) holds for all $\beta \in \operatorname{Comp}_N$ satisfying $\beta \underset{\text{wll}}{<} \alpha$. In other words,

$$M_\beta \in U \qquad \text{for all } \beta \in \operatorname{Comp}_N \text{ satisfying } \beta \underset{\text{wll}}{<} \alpha. \tag{13.186.7}$$

Also, we have assumed that (13.186.5) holds for all compositions β satisfying $|\beta| < N$. In other words,

$$M_\beta \in U \qquad \text{for all compositions } \beta \text{ satisfying } |\beta| < N. \tag{13.186.8}$$

Let $(a_1, a_2, \ldots, a_p)$ be the CFL factorization of the word α. Then, $(a_1, a_2, \ldots, a_p)$ is a tuple of Lyndon words satisfying $\alpha = a_1 a_2 \cdots a_p$ and $a_1 \geq a_2 \geq \cdots \geq a_p$ (according to the definition of a CFL factorization). We have $p \neq 0$ (since otherwise, we would have $p = 0$ and thus $\alpha = a_1 a_2 \cdots a_p = (\text{empty product}) = \varnothing$, contradicting $\alpha \neq \varnothing$). Thus, $p \in \{1, 2, 3, \ldots\}$. Hence, the word a_1 is well-defined. Clearly, a_1 is a Lyndon word (since $(a_1, a_2, \ldots, a_p)$ is a tuple of Lyndon words).

We distinguish between two cases:

Case 1: All of the words a_1, a_2, $\ldots$, a_p are equal.

Case 2: Not all of the words a_1, a_2, $\ldots$, a_p are equal.

Let us consider Case 1 first. In this case, all of the words a_1, a_2, $\ldots$, a_p are equal. In other words, $a_1 = a_2 = \cdots = a_p$. Thus, $a_1 = a_i$ for every $i \in \{1, 2, \ldots, p\}$.

Let $x = a_1$. Then, $x = a_1 = a_i$ for every $i \in \{1, 2, \ldots, p\}$. Multiplying these identities for all $i \in \{1, 2, \ldots, p\}$, we obtain $\underbrace{xx \cdots x}_{p \text{ times}} = a_1 a_2 \cdots a_p = \alpha$, so that $\alpha = \underbrace{xx \cdots x}_{p \text{ times}} = x^p$, thus $x^p = \alpha$. Also, $x = a_1$ is a Lyndon word. Let $k = |x|$. Then, $x \in \operatorname{Comp}_k$. Also, $N = \left| \underbrace{\alpha}_{= x^p} \right| = |x^p| = p\,|x|$, so that $p\,|x| = N$. Thus, $p \underbrace{k}_{= |x|} = p\,|x| = N$.

Now, Corollary 6.5.29 (applied to $s = p$) yields

$$M_x^{\langle p \rangle} - M_{x^p} \in \sum_{\substack{w \in \operatorname{Comp}_{pk};\\ w \underset{\text{wll}}{<} x^p}} \mathbf{k} M_w = \sum_{\substack{w \in \operatorname{Comp}_N;\\ w \underset{\text{wll}}{<} \alpha}} \mathbf{k} \underbrace{M_w}_{\substack{\in U\\ \text{(by (13.186.7), applied}\\ \text{to } \beta = w)}} \qquad (\text{since } pk = N \text{ and } x^p = \alpha)$$

$$\subset \sum_{\substack{w \in \operatorname{Comp}_N;\\ w \underset{\text{wll}}{<} \alpha}} \mathbf{k} U \subset U \qquad (\text{since } U \text{ is a } \mathbf{k}\text{-submodule of } \operatorname{QSym}) .$$

Hence,

$$M_{x^p} - M_x^{\langle p \rangle} = - \underbrace{\left(M_x^{\langle p \rangle} - M_{x^p} \right)}_{\in U} \in -U \subset U \qquad (\text{since } U \text{ is a } \mathbf{k}\text{-submodule of } \operatorname{QSym}) ,$$

[1184] *Proof.* Assume that $\alpha = \varnothing$. Then, $M_\alpha = M_\varnothing = 1 \in U$ (since U is a $\mathbf{k}$-subalgebra of QSym), qed.

so that

$$M_{x^p} \in \underbrace{M_x^{\langle p \rangle}}_{\substack{\in U \\ \text{(by (13.186.4), applied} \\ \text{to } x \text{ and } p \text{ instead of } \beta \text{ and } s)}} + U \in U + U \subset U$$

(since U is a $\mathbf{k}$-submodule of QSym). Since $x^p = \alpha$, this rewrites as $M_\alpha \in U$. Thus, we have proven $M_\alpha \in U$ in Case 1.

Let us now consider Case 2. In this case, not all of the words a_1, a_2, ..., a_p are equal. Hence, there exists some $k \in \{1, 2, \ldots, p-1\}$ such that $a_k \neq a_{k+1}$. Consider this k.

We have $a_k \geq a_{k+1}$ (since $a_1 \geq a_2 \geq \cdots \geq a_p$). Combined with $a_k \neq a_{k+1}$, this yields $a_k > a_{k+1}$. Let x be the word $a_1 a_2 \cdots a_k$, and let y be the word $a_{k+1} a_{k+2} \cdots a_p$. Then, Corollary 6.5.25 (applied to $u = \alpha$ and $n = N$) yields

$$M_\alpha = M_x M_y - \left(\text{a sum of terms of the form } M_w \text{ with } w \in \operatorname{Comp}_N \text{ satisfying } w \underset{\text{wll}}{<} \alpha \right).$$

Thus,

$$\begin{aligned} M_x M_y - M_\alpha &= \left(\text{a sum of terms of the form } M_w \text{ with } w \in \operatorname{Comp}_N \text{ satisfying } w \underset{\text{wll}}{<} \alpha \right) \\ &\in \sum_{\substack{w \in \operatorname{Comp}_N; \\ w \underset{\text{wll}}{<} \alpha}} \mathbf{k} \underbrace{M_w}_{\substack{\in U \\ \text{(by (13.186.7), applied} \\ \text{to } \beta = w)}} \subset \sum_{\substack{w \in \operatorname{Comp}_N; \\ w \underset{\text{wll}}{<} \alpha}} \mathbf{k} U \subset U \end{aligned}$$

(since U is a $\mathbf{k}$-submodule of QSym). Hence,

$$M_\alpha \in M_x M_y - U. \tag{13.186.9}$$

Now, it is easy to see that $N = |x| + |y|$ [1185]. But $|x| > 0$ [1186] and $|y| > 0$ [1187]. Hence, $N = |x| + \underbrace{|y|}_{>0} > |x|$, which yields $|x| < N$. Hence, (13.186.8) (applied to $\beta = x$) yields $M_x \in U$. Also, $N = \underbrace{|x|}_{>0} + |y| > |y|$, and thus $|y| < N$, so that (13.186.8) (applied to $\beta = y$) yields $M_y \in U$. Now, (13.186.9) becomes $M_\alpha \in \underbrace{M_x}_{\in U} \underbrace{M_y}_{\in U} - U \subset UU - U \subset U$ (since U is a $\mathbf{k}$-subalgebra of QSym). Thus, we have proven $M_\alpha \in U$ in Case 2.

Now, we have proven $M_\alpha \in U$ in each of the two Cases 1 and 2. Since these two Cases cover all possibilities, this yields that we always have $M_\alpha \in U$. In other words, (13.186.6) holds for $\beta = \alpha$.

Thus, we have completed the induction step of our induction over β. Therefore, we have proven (13.186.6) by induction. So we now know that $M_\beta \in U$ for every $\beta \in \operatorname{Comp}_N$. In other words, $M_\beta \in U$ for every composition β satisfying $|\beta| = N$. In other words, (13.186.5) holds for all compositions β satisfying $|\beta| = N$.

Thus, we have completed the induction step of our induction over N. Hence, (13.186.5) is proven by induction over N.

[1185] *Proof.* Multiplying the equalities $x = a_1 a_2 \cdots a_k$ and $y = a_{k+1} a_{k+2} \cdots a_p$, we obtain

$$xy = (a_1 a_2 \cdots a_k)(a_{k+1} a_{k+2} \cdots a_p) = a_1 a_2 \cdots a_p = \alpha,$$

so that $\alpha = xy$ and thus $|\alpha| = |xy| = |x| + |y|$, so that $|x| + |y| = |\alpha| = N$, qed.

[1186] *Proof.* Notice that a_k is a Lyndon word (since $(a_1, a_2, \ldots, a_p)$ is a tuple of Lyndon words), and thus nonempty.

But a_k is a suffix of the word $a_1 a_2 \cdots a_k$. Hence, $\ell(a_k) \leq \ell(a_1 a_2 \cdots a_k)$, so that $\ell(a_1 a_2 \cdots a_k) \geq \ell(a_k) > 0$ (since the word a_k is nonempty). Now, $\ell\left(\underbrace{x}_{= a_1 a_2 \cdots a_k}\right) = \ell(a_1 a_2 \cdots a_k) > 0$, so that the word x is nonempty, and therefore $|x| > 0$, qed.

[1187] *Proof.* Notice that a_{k+1} is a Lyndon word (since $(a_1, a_2, \ldots, a_p)$ is a tuple of Lyndon words), and thus nonempty.

But a_{k+1} is a prefix of the word $a_{k+1} a_{k+2} \cdots a_p$. Hence, $\ell(a_{k+1}) \leq \ell(a_{k+1} a_{k+2} \cdots a_p)$, so that $\ell(a_{k+1} a_{k+2} \cdots a_p) \geq \ell(a_{k+1}) > 0$ (since the word a_{k+1} is nonempty). Now, $\ell\left(\underbrace{y}_{= a_{k+1} a_{k+2} \cdots a_p}\right) = \ell(a_{k+1} a_{k+2} \cdots a_p) > 0$, so that the word y is nonempty, and therefore $|y| > 0$, qed.

Now, recall that the family $(M_\beta)_{\beta\in\mathrm{Comp}}$ is a basis of the $\mathbf{k}$-module QSym, and thus generates this $\mathbf{k}$-module. Hence,

$$\mathrm{QSym} = \sum_{\beta\in\mathrm{Comp}} \mathbf{k} \underbrace{M_\beta}_{\substack{\in U \\ (\text{by } (13.186.5))}} \subset \sum_{\beta\in\mathrm{Comp}} \mathbf{k}U \subset U$$

(since U is a $\mathbf{k}$-submodule of QSym). Combined with $U \subset \mathrm{QSym}$, this yields $U = \mathrm{QSym}$. Since U is the $\mathbf{k}$-subalgebra of QSym generated by the family $\left(M_{\operatorname{red} w}^{\langle \gcd w\rangle}\right)_{w\in\mathfrak{L}}$, this shows that the $\mathbf{k}$-subalgebra of QSym generated by the family $\left(M_{\operatorname{red} w}^{\langle \gcd w\rangle}\right)_{w\in\mathfrak{L}}$ is QSym itself. In other words, the family $\left(M_{\operatorname{red} w}^{\langle \gcd w\rangle}\right)_{w\in\mathfrak{L}}$ generates the $\mathbf{k}$-algebra QSym. Thus, (13.186.1) is proven. As we know, this completes the proof of Theorem 6.5.13. □

13.187. **Solution to Exercise 6.5.32.** *Solution to Exercise 6.5.32.*

Proof of Theorem 6.4.3. We know (from the proof of Theorem 6.5.13) that the family $\left(M_{\operatorname{red} w}^{\langle \gcd w\rangle}\right)_{w\in\mathfrak{L}}$ is an algebraically independent generating set of the $\mathbf{k}$-algebra QSym.

Now, define a grading on the $\mathbf{k}$-algebra $\mathbf{k}[x_w \mid w\in\mathfrak{L}]$ by setting $\deg(x_w) = \sum_{i=1}^{\ell(w)} w_i$ for every $w\in\mathfrak{L}$. By the universal property of the polynomial algebra $\mathbf{k}[x_w \mid w\in\mathfrak{L}]$, we can define a $\mathbf{k}$-algebra homomorphism $\Phi : \mathbf{k}[x_w \mid w\in\mathfrak{L}] \to \mathrm{QSym}$ by setting

$$\Phi(x_w) = M_{\operatorname{red} w}^{\langle \gcd w\rangle} \qquad \text{for every } w\in\mathfrak{L}.$$

[1188] This homomorphism Φ is a $\mathbf{k}$-algebra isomorphism (since $\left(M_{\operatorname{red} w}^{\langle \gcd w\rangle}\right)_{w\in\mathfrak{L}}$ is an algebraically independent generating set of the $\mathbf{k}$-algebra QSym) and is graded (because for every $w\in\mathfrak{L}$, the element $M_{\operatorname{red} w}^{\langle \gcd w\rangle}$ of QSym is homogeneous of degree $\deg(x_w)$ [1189]). Thus, Φ is an isomorphism of graded $\mathbf{k}$-algebras. Hence, $\mathrm{QSym} \cong \mathbf{k}[x_w \mid w\in\mathfrak{L}]$ as graded $\mathbf{k}$-algebras. Thus, QSym is a polynomial algebra. This proves Theorem 6.4.3. □

13.188. **Solution to Exercise 6.5.34.** *Solution to Exercise 6.5.34.*

Proof of Corollary 6.5.33. Theorem 6.5.13 yields that the family $\left(M_w^{\langle s\rangle}\right)_{(w,s)\in\mathfrak{RL}\times\{1,2,3,\ldots\}}$ is an algebraically independent generating set of the $\mathbf{k}$-algebra QSym.

Notice that (1) is a reduced Lyndon composition; that is, $(1)\in\mathfrak{RL}$ (since $\mathfrak{RL}$ is the set of all reduced Lyndon compositions). Hence, $\{(1)\}\subset\mathfrak{RL}$, whence $\{(1)\}\times\{1,2,3,\ldots\}\subset\mathfrak{RL}\times\{1,2,3,\ldots\}$.

The following fact is straightforward to check: If A is a commutative $\mathbf{k}$-algebra, and if $(a_i)_{i\in I}$ is an algebraically independent generating set of the $\mathbf{k}$-algebra A, and if J is a subset of I, then $(a_i)_{i\in I\setminus J}$ is an algebraically independent generating set of the $\mathbf{k}[a_i \mid i\in J]$-algebra A. [1190] We can apply this fact to $A = \mathrm{QSym}$, $I = \mathfrak{RL}\times\{1,2,3,\ldots\}$, $(a_i)_{i\in I} = \left(M_w^{\langle s\rangle}\right)_{(w,s)\in\mathfrak{RL}\times\{1,2,3,\ldots\}}$ and $J = \{(1)\}\times\{1,2,3,\ldots\}$. As a result, we conclude that $\left(M_w^{\langle s\rangle}\right)_{(w,s)\in(\mathfrak{RL}\times\{1,2,3,\ldots\})\setminus(\{(1)\}\times\{1,2,3,\ldots\})}$ is an algebraically independent generating set of

[1188] This is well-defined since $M_{\operatorname{red} w}^{\langle \gcd w\rangle}\in\mathrm{QSym}$ (by Corollary 6.5.8(a), applied to $\alpha = \operatorname{red} w$ and $s = \gcd w$).

[1189] *Proof.* Let $w\in\mathfrak{L}$. Then, w is a Lyndon word (since $\mathfrak{L}$ is the set of all Lyndon words), and thus nonempty. Hence, Remark 6.5.11(f) (applied to $\alpha = w$) yields $(\gcd w)\left|\operatorname{red} w\right| = |w| = \sum_{i=1}^{\ell(w)} w_i = \deg(x_w)$. Now, Corollary 6.5.8(b) (applied to $\alpha = \operatorname{red} w$ and $s = \gcd w$) yields $M_{\operatorname{red} w}^{\langle \gcd w\rangle}\in\mathrm{QSym}_{(\gcd w)\left|\operatorname{red} w\right|} = \mathrm{QSym}_{\deg(x_w)}$ (since $(\gcd w)\left|\operatorname{red} w\right| = \deg(x_w)$). In other words, the element $M_{\operatorname{red} w}^{\langle \gcd w\rangle}$ of QSym is homogeneous of degree $\deg(x_w)$, qed.

[1190] The idea behind this fact is that a polynomial ring in the indeterminates $(x_i)_{i\in I}$ can be regarded as a polynomial ring in the indeterminates $(x_i)_{i\in I\setminus J}$ over the polynomial ring in the indeterminates $(x_i)_{i\in J}$.

the $\mathbf{k}\left[M_w^{\langle s\rangle} \mid (w,s)\in\{(1)\}\times\{1,2,3,\ldots\}\right]$-algebra QSym. Hence, the $\mathbf{k}\left[M_w^{\langle s\rangle} \mid (w,s)\in\{(1)\}\times\{1,2,3,\ldots\}\right]$-algebra QSym has an algebraically independent generating set; it is therefore a polynomial algebra over $\mathbf{k}\left[M_w^{\langle s\rangle} \mid (w,s)\in\{(1)\}\times\{1,2,3,\ldots\}\right]$. Since $\mathbf{k}\left[M_w^{\langle s\rangle} \mid (w,s)\in\{(1)\}\times\{1,2,3,\ldots\}\right] = \Lambda$ [1191], this rewrites as follows: The Λ-algebra QSym is a polynomial algebra over Λ. This proves Corollary 6.5.33. □

13.189. **Solution to Exercise 6.6.8.** *Solution to Exercise 6.6.8.* Before we step to the proofs of Proposition 6.6.5, Lemma 6.6.6 and Proposition 6.6.7, let us state a basic fact about cycles of permutations:

Lemma 13.189.1. *Let $n\in\mathbb{N}$. Let $\tau\in\mathfrak{S}_n$ be a permutation. Let z be a cycle of τ. Let k be the size of z. Let $p\in z$. For each $u\in\mathbb{N}$, we set*

$$p_u := \tau^u(p).$$

Then:

(a) *We have $z=\{p_0,p_1,p_2,\ldots\}$.*
(b) *We have $\tau^k(p)=p$.*
(c) *We have $z=\{p_0,p_1,\ldots,p_{k-1}\}$.*
(d) *We have $\tau^i(p_u)=p_{u+i}$ for each $u\in\mathbb{N}$ and $i\in\mathbb{N}$.*
(e) *We have $p_{u+k}=p_u$ for each $u\in\mathbb{N}$.*
(f) *We have $p_u\in z$ for each $u\in\mathbb{N}$.*
(g) *The k elements $p_0,p_1,\ldots,p_{k-1}$ are distinct.*
(h) *We have $z=\{p_u,p_{u+1},\ldots,p_{u+k-1}\}$ for each $u\in\mathbb{N}$.*
(i) *We have $\operatorname{ord}_\tau(p)=k$.*

Proof of Lemma 13.189.1. All of these are well-known properties of cycles in permutations. □

We record a simple corollary of Lemma 13.189.1:

Lemma 13.189.2. *Let $n\in\mathbb{N}$. Let $\tau\in\mathfrak{S}_n$ be a permutation. Let $p\in\{1,2,\ldots,n\}$. Let $j=\operatorname{ord}_\tau(p)$. Then:*

(a) *We have $\tau^j(p)=p$.*
(b) *We have $j=\operatorname{ord}_\tau(\tau(p))$.*

Proof of Lemma 13.189.2. Let z be the cycle of τ that contains p. Thus, $p\in z$. Let k be the size of z. For each $u\in\mathbb{N}$, we set $p_u:=\tau^u(p)$.

Lemma 13.189.1(i) yields $\operatorname{ord}_\tau(p)=k$. Hence, $j=\operatorname{ord}_\tau(p)=k$. But Lemma 13.189.1(b) yields $\tau^k(p)=p$. But from $j=k$, we obtain $\tau^j(p)=\tau^k(p)=p$. This proves Lemma 13.189.2(a).

(b) Lemma 13.189.1(f) (applied to $u=1$) yields $p_1\in z$. But the definition of p_1 yields $p_1=\underbrace{\tau^1}_{=\tau}(p)=\tau(p)$. Hence, $\tau(p)=p_1\in z$.

Set $q=\tau(p)$. Thus, $q=\tau(p)\in z$. For each $u\in\mathbb{N}$, we set $q_u:=\tau^u(q)$. Then, Lemma 13.189.1(i) (applied to q and q_u instead of p and p_u) yields $\operatorname{ord}_\tau(q)=k$. Comparing this with $j=k$, we obtain $j=\operatorname{ord}_\tau\left(\underbrace{q}_{=\tau(p)}\right)=\operatorname{ord}_\tau(\tau(p))$. This proves Lemma 13.189.2(b). □

[1191] *Proof.* The first sentence of Proposition 2.4.1 yields that the family $(e_1,e_2,e_3,\ldots)$ generates the $\mathbf{k}$-algebra Λ. In other words, $\Lambda=\mathbf{k}[e_1,e_2,e_3,\ldots]=\mathbf{k}[e_s \mid s\in\{1,2,3,\ldots\}]$.

But

$$\begin{aligned}
&\mathbf{k}\left[M_w^{\langle s\rangle} \mid (w,s)\in\{(1)\}\times\{1,2,3,\ldots\}\right]\\
&=\mathbf{k}\left[\underbrace{M_{(1)}^{\langle s\rangle}}_{\substack{=e_s\\ \text{(by Exercise 6.5.5)}}} \mid s\in\{1,2,3,\ldots\}\right] \qquad \left(\begin{array}{c}\text{since the elements of } \{(1)\}\times\{1,2,3,\ldots\} \text{ are precisely}\\ \text{the pairs of the form } ((1),s) \text{ for } s\in\{1,2,3,\ldots\}\end{array}\right)\\
&=\mathbf{k}[e_s \mid s\in\{1,2,3,\ldots\}]=\Lambda
\end{aligned}$$

(since $\Lambda=\mathbf{k}[e_s \mid s\in\{1,2,3,\ldots\}]$), qed.

Proof of Proposition 6.6.5. Let $k = \operatorname{ord}_\tau(h)$. Thus, k is the smallest positive integer i such that $\tau^i(h) = h$ (by the definition of $\operatorname{ord}_\tau(h)$). Hence, k is a positive integer, and we have $\tau^k(h) = h$. Furthermore, $\tau^{k+1}(h) = \underbrace{\tau}_{=\tau^1}\left(\underbrace{\tau^k(h)}_{=h}\right) = \tau^1(h)$, so that $\tau^1(h) = \tau^{k+1}(h)$.

The definition of $w_{\tau,h}$ yields

$$\begin{aligned} w_{\tau,h} &= w_{\tau^1(h)} w_{\tau^2(h)} \cdots w_{\tau^k(h)} \qquad (\text{since } k = \operatorname{ord}_\tau(h)) \\ &= \left(w_{\tau^1(h)}, w_{\tau^2(h)}, \ldots, w_{\tau^k(h)}\right). \end{aligned} \tag{13.189.1}$$

Hence, the word $w_{\tau,h}$ has length k. In other words, the word $w_{\tau,h}$ has length $\operatorname{ord}_\tau(h)$ (since $k = \operatorname{ord}_\tau(h)$). Moreover, the word $w_{\tau,h}$ is nonempty (since it has length k, but k is a positive integer). This proves Proposition 6.6.5(a).

(b) The first letter of the word $w_{\tau,h}$ exists (since the word $w_{\tau,h}$ is nonempty). This first letter is $w_{\tau^1(h)}$ (since $w_{\tau,h} = \left(w_{\tau^1(h)}, w_{\tau^2(h)}, \ldots, w_{\tau^k(h)}\right)$). In other words, this first letter is $w_{\tau(h)}$ (since $\tau^1 = \tau$). This proves Proposition 6.6.5(b).

(c) The last letter of the word $w_{\tau,h}$ exists (since the word $w_{\tau,h}$ is nonempty). This last letter is $w_{\tau^k(h)}$ (since $w_{\tau,h} = \left(w_{\tau^1(h)}, w_{\tau^2(h)}, \ldots, w_{\tau^k(h)}\right)$). In other words, this last letter is w_h (since $\tau^k(h) = h$). This proves Proposition 6.6.5(c).

(d) From (13.189.1), we obtain

$$\begin{aligned} c \cdot w_{\tau,h} &= c \cdot \left(w_{\tau^1(h)}, w_{\tau^2(h)}, \ldots, w_{\tau^k(h)}\right) = \left(w_{\tau^2(h)}, w_{\tau^3(h)}, \ldots, w_{\tau^k(h)}, w_{\tau^1(h)}\right) \\ &\qquad \left(\text{by the definition of the action of } C \text{ on } \mathfrak{A}^k\right) \\ &= \left(w_{\tau^2(h)}, w_{\tau^3(h)}, \ldots, w_{\tau^k(h)}, w_{\tau^{k+1}(h)}\right) \qquad \left(\text{since } \tau^1(h) = \tau^{k+1}(h)\right) \\ &= \left(w_{\tau^2(h)}, w_{\tau^3(h)}, \ldots, w_{\tau^{k+1}(h)}\right). \end{aligned} \tag{13.189.2}$$

On the other hand, we have $k = \operatorname{ord}_\tau(h)$ and therefore $k = \operatorname{ord}_\tau(\tau(h))$ (by Lemma 13.189.2(b), applied to $j = k$ and $p = h$). Hence, the definition of $w_{\tau,\tau(h)}$ yields

$$\begin{aligned} w_{\tau,\tau(h)} &= w_{\tau^1(\tau(h))} w_{\tau^2(\tau(h))} \cdots w_{\tau^k(\tau(h))} = \left(w_{\tau^1(\tau(h))}, w_{\tau^2(\tau(h))}, \ldots, w_{\tau^k(\tau(h))}\right) \\ &\qquad \left(\text{since } w_{\tau^1(\tau(h))}, w_{\tau^2(\tau(h))}, \ldots, w_{\tau^k(\tau(h))} \text{ are single letters}\right) \\ &= \left(w_{\tau^2(h)}, w_{\tau^3(h)}, \ldots, w_{\tau^{k+1}(h)}\right) \\ &\qquad \left(\begin{array}{c} \text{since each } i \in \{1, 2, \ldots, k\} \text{ satisfies } w_{\tau^i(\tau(h))} = w_{\tau^{i+1}(h)} \\ \left(\text{because } \tau^i(\tau(h)) = \tau^{i+1}(h)\right) \end{array}\right). \end{aligned}$$

Comparing this with (13.189.2), we obtain $w_{\tau,\tau(h)} = c \cdot w_{\tau,h}$. This proves Proposition 6.6.5(d).

(e) Let us first show that

$$w_{\tau,\tau^i(h)} = c^i \cdot w_{\tau,h} \qquad \text{for each } i \in \mathbb{N}. \tag{13.189.3}$$

[*Proof of (13.189.3):* We shall prove (13.189.3) by induction on i:

Induction base: We have $w_{\tau,\tau^0(h)} = w_{\tau,h}$ (since $\underbrace{\tau^0}_{=\operatorname{id}}(h) = h$). Comparing this with $\underbrace{c^0}_{=\operatorname{id}} \cdot w_{\tau,h} = w_{\tau,h}$, we obtain $w_{\tau,\tau^0(h)} = c^0 \cdot w_{\tau,h}$. In other words, (13.189.3) holds for $i = 0$. This completes the induction base.

Induction step: Let $j \in \mathbb{N}$. Assume that (13.189.3) holds for $i = j$. We must prove that (13.189.3) holds for $i = j + 1$.

We have assumed that (13.189.3) holds for $i = j$. In other words, we have $w_{\tau,\tau^j(h)} = c^j \cdot w_{\tau,h}$. Now,

$$\begin{aligned} w_{\tau,\tau^{j+1}(h)} &= w_{\tau,\tau(\tau^j(h))} \qquad \left(\text{since } \tau^{j+1}(h) = \tau\left(\tau^j(h)\right)\right) \\ &= c \cdot \underbrace{w_{\tau,\tau^j(h)}}_{=c^j \cdot w_{\tau,h}} \qquad \left(\text{by Proposition 6.6.5(d), applied to } \tau^j(h) \text{ instead of } h\right) \\ &= \underbrace{c \cdot c^j}_{=c^{j+1}} \cdot w_{\tau,h} = c^{j+1} \cdot w_{\tau,h}. \end{aligned}$$

In other words, (13.189.3) holds for $i = j + 1$. This completes the induction step. Thus, (13.189.3) is proved by induction.]

Now, forget that we fixed h. We thus have proved (13.189.3) for each $h \in \{1, 2, \ldots, n\}$.

Now, let $h \in \{1, 2, \ldots, n\}$. We want to prove Proposition 6.6.5(e); in other words, we must prove that $w_{\tau,\tau^i(h)} = c^i \cdot w_{\tau,h}$ for each $i \in \mathbb{Z}$. So let us fix $i \in \mathbb{Z}$. We must then prove that $w_{\tau,\tau^i(h)} = c^i \cdot w_{\tau,h}$. If $i \in \mathbb{N}$, then this follows immediately from (13.189.3). Thus, we WLOG assume that $i \notin \mathbb{N}$. Hence, i is a negative integer (since $i \in \mathbb{Z}$ but $i \notin \mathbb{N}$), so that $-i \in \{1, 2, 3, \ldots\} \subset \mathbb{N}$. Thus, we can apply (13.189.3) to $\tau^i(h)$ and $-i$ instead of h and i (since we have proved (13.189.3) for every value of h, not just for the h that we are currently considering). We thus obtain

$$w_{\tau,\tau^{-i}(\tau^i(h))} = c^{-i} \cdot w_{\tau,\tau^i(h)}.$$

In view of $\tau^{-i}\left(\tau^i(h)\right) = \underbrace{\left(\tau^{-i} \circ \tau^i\right)}_{=\mathrm{id}}(h) = h$, this rewrites as

$$w_{\tau,h} = c^{-i} \cdot w_{\tau,\tau^i(h)}.$$

Hence,

$$c^i \cdot \underbrace{w_{\tau,h}}_{=c^{-i}\cdot w_{\tau,\tau^i(h)}} = \underbrace{c^i \cdot c^{-i}}_{=\mathrm{id}} \cdot w_{\tau,\tau^i(h)} = w_{\tau,\tau^i(h)}.$$

In other words, $w_{\tau,\tau^i(h)} = c^i \cdot w_{\tau,h}$. This proves Proposition 6.6.5(e). □

Proof of Lemma 6.6.6. Definition 5.3.3 says that $\operatorname{std} w$ is the unique permutation $\sigma \in \mathfrak{S}_n$ defined in Proposition 5.3.2. In other words, $\operatorname{std} w$ is the unique permutation $\sigma \in \mathfrak{S}_n$ such that for every two elements a and b of $\{1, 2, \ldots, n\}$ satisfying $a < b$, we have ($\sigma(a) < \sigma(b)$ if and only if $w_a \le w_b$). Hence, $\operatorname{std} w$ is such a permutation σ. In other words, $\operatorname{std} w$ is a permutation in $\mathfrak{S}_n$ and has the property that for every two elements a and b of $\{1, 2, \ldots, n\}$ satisfying $a < b$, we have

$$\left((\operatorname{std} w)(a) < (\operatorname{std} w)(b) \text{ if and only if } w_a \le w_b\right). \qquad (13.189.4)$$

But $\tau = (\operatorname{std} w)^{-1}$, so that $\tau^{-1} = \operatorname{std} w$. Note that τ is a permutation (since $\tau \in \mathfrak{S}_n$). Hence, τ is a bijective map. Thus, in particular, τ is an injective map.

(a) Assume that $\tau^{-1}(\alpha) < \tau^{-1}(\beta)$. In view of $\tau^{-1} = \operatorname{std} w$, this rewrites as $(\operatorname{std} w)(\alpha) < (\operatorname{std} w)(\beta)$. But (13.189.4) (applied to $a = \alpha$ and $b = \beta$) shows that we have

$$\left((\operatorname{std} w)(\alpha) < (\operatorname{std} w)(\beta) \text{ if and only if } w_\alpha \le w_\beta\right).$$

Hence, we have $w_\alpha \le w_\beta$ (since we have $(\operatorname{std} w)(\alpha) < (\operatorname{std} w)(\beta)$). This proves Lemma 6.6.6(a).

(b) Assume that $\tau^{-1}(\alpha) \ge \tau^{-1}(\beta)$. In view of $\tau^{-1} = \operatorname{std} w$, this rewrites as $(\operatorname{std} w)(\alpha) \ge (\operatorname{std} w)(\beta)$. Hence, $(\operatorname{std} w)(\alpha) < (\operatorname{std} w)(\beta)$ does not hold. But (13.189.4) (applied to $a = \alpha$ and $b = \beta$) shows that we have

$$\left((\operatorname{std} w)(\alpha) < (\operatorname{std} w)(\beta) \text{ if and only if } w_\alpha \le w_\beta\right).$$

Hence, $w_\alpha \le w_\beta$ does not hold (since $(\operatorname{std} w)(\alpha) < (\operatorname{std} w)(\beta)$ does not hold). In other words, we have $w_\alpha > w_\beta$. This proves Lemma 6.6.6(b).

(d) Assume that $\tau(\alpha) \ge \tau(\beta)$. But $\alpha \ne \beta$ (since $\alpha < \beta$) and thus $\tau(\alpha) \ne \tau(\beta)$ (since τ is injective). Combining this with $\tau(\alpha) \ge \tau(\beta)$, we obtain $\tau(\alpha) > \tau(\beta)$. In other words, $\tau(\beta) < \tau(\alpha)$. Also, $\beta \ge \alpha$ (since $\alpha < \beta$), so that $\tau^{-1}(\tau(\beta)) = \beta \ge \alpha = \tau^{-1}(\tau(\alpha))$. Hence, Lemma 6.6.6(b) (applied to $\tau(\beta)$ and $\tau(\alpha)$ instead of α and β) yields $w_{\tau(\beta)} > w_{\tau(\alpha)}$. In other words, $w_{\tau(\alpha)} < w_{\tau(\beta)}$. This proves Lemma 6.6.6(d).

(c) We are in one of the following two cases:

Case 1: We have $\tau(\alpha) < \tau(\beta)$.

Case 2: We have $\tau(\alpha) \ge \tau(\beta)$.

Let us first consider Case 1. In this case, we have $\tau(\alpha) < \tau(\beta)$. Also, $\tau^{-1}(\tau(\alpha)) = \alpha < \beta = \tau^{-1}(\tau(\beta))$. Hence, Lemma 6.6.6(a) (applied to $\tau(\alpha)$ and $\tau(\beta)$ instead of α and β) yields $w_{\tau(\alpha)} \le w_{\tau(\beta)}$. Thus, Lemma 6.6.6(c) is proven in Case 1.

Let us next consider Case 2. In this case, we have $\tau(\alpha) \ge \tau(\beta)$. Hence, Lemma 6.6.6(d) yields $w_{\tau(\alpha)} < w_{\tau(\beta)}$. Hence, $w_{\tau(\alpha)} \le w_{\tau(\beta)}$. Thus, Lemma 6.6.6(c) is proven in Case 2.

We have thus proved Lemma 6.6.6(c) in each of the two Cases 1 and 2. Hence, Lemma 6.6.6(c) always holds.

(e) Assume that $w_{\tau(\alpha)} = w_{\tau(\beta)}$. We must prove that $\tau(\alpha) < \tau(\beta)$.

Assume the contrary. Thus, $\tau(\alpha) \geq \tau(\beta)$. Hence, Lemma 6.6.6(d) yields $w_{\tau(\alpha)} < w_{\tau(\beta)}$. This contradicts $w_{\tau(\alpha)} = w_{\tau(\beta)}$. This contradiction shows that our assumption was false. Hence, Lemma 6.6.6(e) is proven.

(f) Assume that $w_{\tau,\alpha} = w_{\tau,\beta}$.

Now, Proposition 6.6.5(b) (applied to $h = \alpha$) yields that the first letter of the word $w_{\tau,\alpha}$ is $w_{\tau(\alpha)}$. Likewise, the first letter of the word $w_{\tau,\beta}$ is $w_{\tau(\beta)}$. Hence, the first letters of the two words $w_{\tau,\alpha}$ and $w_{\tau,\beta}$ are $w_{\tau(\alpha)}$ and $w_{\tau(\beta)}$. Thus, from $w_{\tau,\alpha} = w_{\tau,\beta}$, we obtain $w_{\tau(\alpha)} = w_{\tau(\beta)}$. Hence, Lemma 6.6.6(e) yields $\tau(\alpha) < \tau(\beta)$. It remains to show that $w_{\tau,\tau(\alpha)} = w_{\tau,\tau(\beta)}$.

Proposition 6.6.5(d) (applied to $h = \alpha$) yields $w_{\tau,\tau(\alpha)} = c \cdot w_{\tau,\alpha}$. Proposition 6.6.5(d) (applied to $h = \beta$) yields $w_{\tau,\tau(\beta)} = c \cdot w_{\tau,\beta}$. Thus,

$$w_{\tau,\tau(\alpha)} = c \cdot \underbrace{w_{\tau,\alpha}}_{=w_{\tau,\beta}} = c \cdot w_{\tau,\beta} = w_{\tau,\tau(\beta)}.$$

Thus, the proof of Lemma 6.6.6(f) is complete (since we already have shown that $\tau(\alpha) < \tau(\beta)$).

(g) Assume that $w_{\tau,\alpha} = w_{\tau,\beta}$. We must show that

$$\tau^i(\alpha) < \tau^i(\beta) \qquad \text{for each } i \in \mathbb{N}. \tag{13.189.5}$$

[*Proof of (13.189.5):* We shall prove (13.189.5) by induction on i:

Induction base: We have $\alpha < \beta$. In view of $\underbrace{\tau^0}_{=\mathrm{id}}(\alpha) = \alpha$ and $\underbrace{\tau^0}_{=\mathrm{id}}(\beta) = \beta$, this rewrites as $\tau^0(\alpha) < \tau^0(\beta)$. In other words, (13.189.5) holds for $i = 0$. This completes the induction base.

Induction step: Let $j \in \mathbb{N}$. Assume that (13.189.5) holds for $i = j$. We must prove that (13.189.5) holds for $i = j + 1$.

We have assumed that (13.189.5) holds for $i = j$. In other words, we have $\tau^j(\alpha) < \tau^j(\beta)$.

Proposition 6.6.5(e) (applied to $i = j$ and $h = \alpha$) yields $w_{\tau,\tau^j(\alpha)} = c^j \cdot w_{\tau,\alpha}$. The same argument (applied to β instead of α) yields $w_{\tau,\tau^j(\beta)} = c^j \cdot w_{\tau,\beta}$. Thus,

$$w_{\tau,\tau^j(\alpha)} = c^j \cdot \underbrace{w_{\tau,\alpha}}_{=w_{\tau,\beta}} = c^j \cdot w_{\tau,\beta} = w_{\tau,\tau^j(\beta)}.$$

Hence, Lemma 6.6.6(f) (applied to $\tau^j(\alpha)$ and $\tau^j(\beta)$ instead of α and β) yields $\tau\left(\tau^j(\alpha)\right) < \tau\left(\tau^j(\beta)\right)$ and $w_{\tau,\tau(\tau^j(\alpha))} = w_{\tau,\tau(\tau^j(\beta))}$. Now,

$$\underbrace{\tau^{j+1}}_{=\tau\circ\tau^j}(\alpha) = \left(\tau \circ \tau^j\right)(\alpha) = \tau\left(\tau^j(\alpha)\right) < \tau\left(\tau^j(\beta)\right) = \underbrace{\left(\tau \circ \tau^j\right)}_{=\tau^{j+1}}(\beta) = \tau^{j+1}(\beta).$$

In other words, (13.189.5) holds for $i = j + 1$. This completes the induction step. Thus, (13.189.5) is proved by induction.]

Hence, Lemma 6.6.6(g) is proved.

(h) We have assumed that

$$\text{every } i \in \{0, 1, \ldots, j-1\} \text{ satisfies } w_{\tau^{i+1}(\alpha)} = w_{\tau^{i+1}(\beta)}. \tag{13.189.6}$$

Now, we claim that

$$\tau^i(\alpha) < \tau^i(\beta) \qquad \text{for every } i \in \{0, 1, \ldots, j\}. \tag{13.189.7}$$

[*Proof of (13.189.7):* We shall prove (13.189.7) by induction on i:

Induction base: We have $\alpha < \beta$. In view of $\underbrace{\tau^0}_{=\mathrm{id}}(\alpha) = \alpha$ and $\underbrace{\tau^0}_{=\mathrm{id}}(\beta) = \beta$, this rewrites as $\tau^0(\alpha) < \tau^0(\beta)$. In other words, (13.189.7) holds for $i = 0$. This completes the induction base.

Induction step: Let $s \in \{0, 1, \ldots, j-1\}$. Assume that (13.189.7) holds for $i = s$. We must prove that (13.189.7) holds for $i = s + 1$.

We have assumed that (13.189.7) holds for $i = s$. In other words, we have $\tau^s(\alpha) < \tau^s(\beta)$. Moreover, (13.189.6) (applied to $i = s$) yields $w_{\tau^{s+1}(\alpha)} = w_{\tau^{s+1}(\beta)}$. In view of $\tau^{s+1}(\alpha) = \tau(\tau^s(\alpha))$ and $\tau^{s+1}(\beta) = \tau(\tau^s(\beta))$, this rewrites as $w_{\tau(\tau^s(\alpha))} = w_{\tau(\tau^s(\beta))}$. Hence, Lemma 6.6.6(e) (applied to $\tau^s(\alpha)$ and $\tau^s(\beta)$ instead of α and β) yields $\tau(\tau^s(\alpha)) < \tau(\tau^s(\beta))$ (since $\tau^s(\alpha) < \tau^s(\beta)$). In view of $\tau^{s+1}(\alpha) = \tau(\tau^s(\alpha))$ and $\tau^{s+1}(\beta) = \tau(\tau^s(\beta))$, this rewrites as $\tau^{s+1}(\alpha) < \tau^{s+1}(\beta)$. In other words, (13.189.7) holds for $i = s + 1$. This completes the induction step. Thus, (13.189.7) is proved by induction.]

Now, $j \in \{0, 1, \ldots, j\}$. Hence, (13.189.7) (applied to $i = j$) yields $\tau^j(\alpha) < \tau^j(\beta)$. Thus, Lemma 6.6.6(c) (applied to $\tau^j(\alpha)$ and $\tau^j(\beta)$ instead of α and β) yields $w_{\tau(\tau^j(\alpha))} \leq w_{\tau(\tau^j(\beta))}$. In view of $\tau\left(\tau^j(\alpha)\right) = \underbrace{\left(\tau \circ \tau^j\right)}_{=\tau^{j+1}}(\alpha) = \tau^{j+1}(\alpha)$ and $\tau\left(\tau^j(\beta)\right) = \underbrace{\left(\tau \circ \tau^j\right)}_{=\tau^{j+1}}(\beta) = \tau^{j+1}(\beta)$, this rewrites as $w_{\tau^{j+1}(\alpha)} \leq w_{\tau^{j+1}(\beta)}$. Thus, Lemma 6.6.6(h) is proved. □

Proof of Proposition 6.6.7. We have $w = (w_1, w_2, \ldots, w_n)$ (since $w \in \mathfrak{A}^n$).

Let k be the size of z. Thus, $k = |z|$.

(a) Let $h \in z$. We must prove that $[w_{\tau,h}] = \{w_{\tau,i} \mid i \in z\}$.

For each $u \in \mathbb{N}$, we set $h_u := \tau^u(h)$. Lemma 13.189.1(i) (applied to $p = h$ and $p_u = h_u$) yields $\operatorname{ord}_\tau(h) = k$.

Proposition 6.6.5(a) yields that the word $w_{\tau,h}$ is nonempty and has length $\operatorname{ord}_\tau(h)$. In other words, the word $w_{\tau,h}$ is nonempty and has length k (since $\operatorname{ord}_\tau(h) = k$). Thus, $[w_{\tau,h}]$ is the k-necklace that contains $w_{\tau,h}$ (by the definition of $[w_{\tau,h}]$). In other words, $[w_{\tau,h}]$ is the orbit of the C-action on $\mathfrak{A}^k$ that contains $w_{\tau,h}$ (since a k-necklace is the same thing as an orbit of the C-action on $\mathfrak{A}^k$). In other words, $[w_{\tau,h}] = C \cdot w_{\tau,h}$.

But $C = \left\{c^i \mid i \in \mathbb{Z}\right\}$ (since c is a generator of the cyclic group C). Hence,

$$[w_{\tau,h}] = \underbrace{C}_{=\{c^i \mid i \in \mathbb{Z}\}} \cdot w_{\tau,h} = \left\{c^i \mid i \in \mathbb{Z}\right\} \cdot w_{\tau,h}$$

$$= \left\{c^i \cdot w_{\tau,h} \mid i \in \mathbb{Z}\right\}. \tag{13.189.8}$$

On the other hand, z is the cycle of τ that contains h (since z is a cycle of τ, and since $h \in z$). Hence, $z = \left\{\tau^i(h) \mid i \in \mathbb{Z}\right\}$, and thus

$$\{w_{\tau,u} \mid u \in z\} = \left\{w_{\tau,u} \mid u \in \left\{\tau^i(h) \mid i \in \mathbb{Z}\right\}\right\}$$
$$= \left\{ \underbrace{w_{\tau,\tau^i(h)}}_{\substack{=c^i \cdot w_{\tau,h} \\ \text{(by Proposition 6.6.5(e))}}} \mid i \in \mathbb{Z} \right\} = \left\{c^i \cdot w_{\tau,h} \mid i \in \mathbb{Z}\right\}.$$

Comparing this with (13.189.8), we obtain $[w_{\tau,h}] = \{w_{\tau,u} \mid u \in z\} = \{w_{\tau,i} \mid i \in z\}$ (here, we have renamed the index u as i). This proves Proposition 6.6.7(a).

(b) Let α and β be two distinct elements of z. We must prove that $w_{\tau,\alpha} \neq w_{\tau,\beta}$.

We have $\alpha \neq \beta$ (since α and β are distinct). Thus, either $\alpha < \beta$ or $\alpha > \beta$. We WLOG assume that $\alpha < \beta$ (since otherwise, it suffices to swap α with β).

We must prove that $w_{\tau,\alpha} \neq w_{\tau,\beta}$. Assume the contrary. Thus, $w_{\tau,\alpha} = w_{\tau,\beta}$. Lemma 6.6.6(g) thus shows that

$$\tau^i(\alpha) < \tau^i(\beta) \qquad \text{for each } i \in \mathbb{N}. \tag{13.189.9}$$

On the other hand, we have

$$z = \left\{\tau^0(\gamma), \tau^1(\gamma), \ldots, \tau^{k-1}(\gamma)\right\} \qquad \text{for each } \gamma \in z. \tag{13.189.10}$$

[*Proof of (13.189.10):* Let $\gamma \in z$. For each $u \in \mathbb{N}$, we set $\gamma_u := \tau^u(\gamma)$. Then, Lemma 13.189.1(c) (applied to $p = \gamma$ and $p_u = \gamma_u$) yields

$$z = \{\gamma_0, \gamma_1, \ldots, \gamma_{k-1}\} = \left\{\tau^0(\gamma), \tau^1(\gamma), \ldots, \tau^{k-1}(\gamma)\right\}$$

(since each $u \in \{0, 1, \ldots, k-1\}$ satisfies $\gamma_u = \tau^u(\gamma)$ (by the definition of γ_u)). This proves (13.189.10).]

Applying (13.189.10) to $\gamma = \beta$, we obtain $z = \left\{\tau^0(\beta), \tau^1(\beta), \ldots, \tau^{k-1}(\beta)\right\}$.

Now, let m be the smallest element of z. Then, $m \in z = \left\{\tau^0(\beta), \tau^1(\beta), \ldots, \tau^{k-1}(\beta)\right\}$. Hence, there exists some $i \in \{0, 1, \ldots, k-1\}$ such that $m = \tau^i(\beta)$. Consider this i.

But (13.189.10) (applied to $\gamma = \alpha$) yields $z = \left\{\tau^0(\alpha), \tau^1(\alpha), \ldots, \tau^{k-1}(\alpha)\right\}$. From $i \in \{0, 1, \ldots, k-1\}$, we obtain $\tau^i(\alpha) \in \left\{\tau^0(\alpha), \tau^1(\alpha), \ldots, \tau^{k-1}(\alpha)\right\}$. In other words, $\tau^i(\alpha) \in z$ (since $z = \left\{\tau^0(\alpha), \tau^1(\alpha), \ldots, \tau^{k-1}(\alpha)\right\}$).

But m is the **smallest** element of z. Thus, $p \geq m$ for each $p \in z$. We can apply this to $p = \tau^i(\alpha)$ (since $\tau^i(\alpha) \in z$), and thus obtain $\tau^i(\alpha) \geq m = \tau^i(\beta)$. But (13.189.9) yields $\tau^i(\alpha) < \tau^i(\beta)$. These two inequalities clearly contradict one another.

This contradiction shows that our assumption was false. Hence, $w_{\tau,\alpha} \neq w_{\tau,\beta}$ is proved. Thus, Proposition 6.6.7(b) follows.

(c) There are k many elements $i \in z$ (since k is the size of z). The words $w_{\tau,i}$ corresponding to these k many elements i are all distinct (by Proposition 6.6.7(b)). Thus, the set $\{w_{\tau,i} \mid i \in z\}$ (consisting of these words) has size k. In other words, $|\{w_{\tau,i} \mid i \in z\}| = k$. But recall that $k = |z|$. Hence, $|\{w_{\tau,i} \mid i \in z\}| = k = |z|$. This proves Proposition 6.6.7(c).

(d) The set z is a cycle of τ, and thus is nonempty (since any cycle of a permutation is nonempty). Hence, there exists some $h \in z$. Consider this h. Proposition 6.6.7(a) yields $[w_{\tau,h}] = \{w_{\tau,i} \mid i \in z\}$.

But $[w_{\tau,h}]$ is the k-necklace that contains $w_{\tau,h}$ (as we have showed in the proof of Proposition 6.6.7(a) above). Hence, $[w_{\tau,h}]$ is a k-necklace.

Proposition 6.6.7(c) yields $|\{w_{\tau,i} \mid i \in z\}| = |z| = k$ (since $k = |z|$). In view of $[w_{\tau,h}] = \{w_{\tau,i} \mid i \in z\}$, this rewrites as $|[w_{\tau,h}]| = k$.

The period of the necklace $[w_{\tau,h}]$ is $|[w_{\tau,h}]|$ (by the definition of the period of a necklace). In other words, the period of the necklace $[w_{\tau,h}]$ is k (since $|[w_{\tau,h}]| = k$).

But $[w_{\tau,h}]$ is a k-necklace. Hence, this k-necklace $[w_{\tau,h}]$ is aperiodic if and only if its period is k (by the definition of "aperiodic"). Thus, this k-necklace $[w_{\tau,h}]$ is aperiodic (since its period is k). In view of $[w_{\tau,h}] = \{w_{\tau,i} \mid i \in z\}$, this rewrites as follows: The k-necklace $\{w_{\tau,i} \mid i \in z\}$ is aperiodic. Hence, the set $\{w_{\tau,i} \mid i \in z\}$ is an aperiodic necklace. This proves Proposition 6.6.7(d). □

We have now proved Proposition 6.6.5, Lemma 6.6.6 and Proposition 6.6.7. This solves Exercise 6.6.8.

13.190. **Solution to Exercise 6.6.23.** *Solution to Exercise 6.6.23.* In preparation for solving Exercise 6.6.23, we first show a few simple lemmas about necklaces:

Lemma 13.190.1. *Let $u \in \mathfrak{A}^*$ be a nonempty word. Then, $u \in [u]$.*

Proof of Lemma 13.190.1. The word u is nonempty. In other words, $u \in \mathfrak{A}^n$ for some positive integer n. Consider this n. Thus, $[u]$ is the n-necklace containing u (by the definition of $[u]$). Hence, $[u]$ contains u. In other words, $u \in [u]$. This proves Lemma 13.190.1. □

Lemma 13.190.2. *Let N be a necklace. Let $u \in N$. Then, $N = [u]$.*

Proof of Lemma 13.190.2. We know that N is a necklace. In other words, N is an n-necklace for some positive integer n (by the definition of a necklace). Consider this n. We now know that N is an n-necklace; in other words, N is an orbit of the C-action on $\mathfrak{A}^n$ (by the definition of an n-necklace). Hence, N is a set of words of length n. Thus, u is a word of length n (since $u \in N$).

Recall that N is an orbit of the C-action on $\mathfrak{A}^n$. Since N contains u (because $u \in N$), we thus conclude that N is the orbit of the C-action on $\mathfrak{A}^n$ that contains u. In other words, N is the n-necklace that contains u (because an n-necklace is the same thing as an orbit of the C-action on $\mathfrak{A}^n$). In other words, $N = [u]$ (since the n-necklace that contains u has been denoted by $[u]$). This proves Lemma 13.190.2. □

Lemma 13.190.3. *Let $u \in \mathfrak{A}^*$ be a Lyndon word. Then, the necklace $[u]$ is aperiodic.*

Proof of Lemma 13.190.3. The word u is Lyndon, and thus is nonempty (since any Lyndon word is nonempty by definition). Hence, Lemma 13.190.1 yields $u \in [u]$. In other words, the necklace $[u]$ contains u.

Thus, the necklace $[u]$ contains the word u, which is a Lyndon word. Hence, the necklace $[u]$ contains a Lyndon word (namely, u). If the necklace $[u]$ was not aperiodic, then Exercise 6.1.34(d) (applied to $N = [u]$) would show that $[u]$ contains no Lyndon word; but this would contradict the previous sentence. Hence, the necklace $[u]$ must be aperiodic. This proves Lemma 13.190.3. □

Lemma 13.190.4. *Let $u \in \mathfrak{A}^*$ be a nonempty word such that the necklace $[u]$ is aperiodic. Then, $|[u]| = \ell(u)$.*

Proof of Lemma 13.190.4. The word $u \in \mathfrak{A}^*$ is nonempty. Thus, there exists a positive integer p such that $u \in \mathfrak{A}^p$. Consider this p. Hence, $[u]$ is a p-necklace. Also, from $u \in \mathfrak{A}^p$, we obtain $\ell(u) = p$.

The p-necklace $[u]$ is aperiodic. In other words, its period is p (by the definition of an aperiodic p-necklace). Thus,

$$p = (\text{period of the } p\text{-necklace } [u]) = |[u]|$$

(by the definition of the period of a p-necklace). Therefore, $|[u]| = p = \ell(u)$ (since $\ell(u) = p$). This proves Lemma 13.190.4. □

Lemma 13.190.5. *Let k be a positive integer. Let $v \in \mathfrak{A}^k$ be such that the necklace $[v]$ is not aperiodic. Then, there exists some $h \in \{1, 2, \ldots, k-1\}$ such that $c^h \cdot v = v$.*

Proof of Lemma 13.190.5. We have $v \in \mathfrak{A}^k$. Thus, $[v]$ is a k-necklace. Hence, Exercise 6.1.34(a) (applied to k and $[v]$ instead of n and N) yields that $[v]$ is a finite nonempty set and satisfies $|[v]| \mid k$. From $|[v]| \mid k$, we obtain $|[v]| \leq k$ (since k is a positive integer).

Next, we claim the following:

Claim 1: The k elements $c^0 \cdot v, c^1 \cdot v, \ldots, c^{k-1} \cdot v$ are not distinct.

[*Proof of Claim 1:* Assume the contrary. Thus, the k elements $c^0 \cdot v, c^1 \cdot v, \ldots, c^{k-1} \cdot v$ are distinct.

The necklace $[v]$ is defined to be the k-necklace containing v (since $v \in \mathfrak{A}^k$). In other words, the necklace $[v]$ is the C-orbit on $\mathfrak{A}^k$ containing v (since a k-necklace is the same thing as a C-orbit on $\mathfrak{A}^k$). In other words, the necklace $[v]$ is the C-orbit of $v \in \mathfrak{A}^k$. In other words, the necklace $[v]$ is the set $C \cdot v$ (since the C-orbit of $v \in \mathfrak{A}^k$ is the set $C \cdot v$). In other words, $[v] = C \cdot v$. Hence, $|[v]| = |C \cdot v|$.

The k elements $c^0 \cdot v, c^1 \cdot v, \ldots, c^{k-1} \cdot v$ all belong to $C \cdot v$ (since $c^0, c^1, \ldots, c^{k-1}$ all belong to C) and are distinct (by our assumption). Hence, the set $C \cdot v$ contains (at least) k distinct elements (namely, $c^0 \cdot v, c^1 \cdot v, \ldots, c^{k-1} \cdot v$). Thus, $|C \cdot v| \geq k$. Thus, $|[v]| = |C \cdot v| \geq k$. Combining this with $|[v]| \leq k$, we obtain $|[v]| = k$. In other words, the period of $[v]$ is k (since the period of $[v]$ is defined as the integer $|[v]|$). But $[v]$ is a k-necklace. Hence, the necklace $[v]$ is aperiodic if and only if its period is k (by the definition of an aperiodic necklace). Thus, the necklace $[v]$ is aperiodic (since its period is k). This contradicts the fact that the necklace $[v]$ is not aperiodic. This contradiction shows that our assumption was wrong. This completes our proof of Claim 1.]

Claim 1 shows that the k elements $c^0 \cdot v, c^1 \cdot v, \ldots, c^{k-1} \cdot v$ are not distinct. In other words, there exist two elements x and y of $\{0, 1, \ldots, k-1\}$ such that $x < y$ and $c^x \cdot v = c^y \cdot v$. Consider these x and y. From $x \in \{0, 1, \ldots, k-1\}$ and $y \in \{0, 1, \ldots, k-1\}$ and $x < y$, we obtain $0 \leq x < y \leq k-1$, so that $y - x \in \{1, 2, \ldots, k-1\}$. Furthermore,

$$\underbrace{c^{y-x}}_{=c^{-x}c^y} \cdot v = \left(c^{-x}c^y\right) \cdot v = c^{-x} \cdot \underbrace{\left(c^y \cdot v\right)}_{=c^x \cdot v} = c^{-x} \cdot \left(c^x \cdot v\right) = \underbrace{\left(c^{-x}c^x\right)}_{=\mathrm{id}} \cdot v = v.$$

Hence, there exists some $h \in \{1, 2, \ldots, k-1\}$ such that $c^h \cdot v = v$ (namely, $h = y - x$). This proves Lemma 13.190.5. □

Lemma 13.190.6. *Let n be a positive integer. Let d be a positive divisor of n. Thus, n/d is a positive integer.*

Let $q \in \mathfrak{A}^{n/d}$. Then, the n/d-necklace $[q]$ and the n-necklace $\left[q^d\right]$ satisfy $\left|\left[q^d\right]\right| = |[q]| \leq n/d$.

Proof of Lemma 13.190.6. Exercise 6.1.34(a) (applied to n/d instead of n) shows that every n/d-necklace N is a finite nonempty set and satisfies $|N| \mid n/d$. Applying this to $N = [q]$, we conclude that $[q]$ is a finite nonempty set and satisfies $|[q]| \mid n/d$. From $|[q]| \mid n/d$, we obtain $|[q]| \leq n/d$ (since $|[q]| \in \mathbb{N}$ and since n/d is a positive integer).

Now, consider the map $\Delta : \mathfrak{A}^{n/d} \to \mathfrak{A}^n$ defined in Lemma 13.147.5. The definition of Δ yields $\Delta(q) = q^d$. Hence, $q^d = \Delta(q) \in \mathfrak{A}^n$.

The necklace $[q]$ is defined to be the n/d-necklace containing q (since $q \in \mathfrak{A}^{n/d}$). In other words, the necklace $[q]$ is the C-orbit on $\mathfrak{A}^{n/d}$ containing q (since an n/d-necklace is the same thing as a C-orbit on $\mathfrak{A}^{n/d}$). In other words, the necklace $[q]$ is the C-orbit of $q \in \mathfrak{A}^{n/d}$. In other words, the necklace $[q]$ is the set $C \cdot q$ (since the C-orbit of $q \in \mathfrak{A}^{n/d}$ is the set $C \cdot q$). In other words, $[q] = C \cdot q$. The same argument (applied to n and q^d instead of n/d and q) yields $\left[q^d\right] = C \cdot q^d$.

But Lemma 13.147.5(c) shows that the map Δ is injective. Hence, $|\Delta(S)| = |S|$ for any finite subset S of $\mathfrak{A}^{n/d}$. Applying this to $S = [q]$, we obtain $|\Delta([q])| = |[q]|$ (since $[q]$ is a finite subset of $\mathfrak{A}^{n/d}$).

On the other hand, Lemma 13.147.5(b) shows that the map Δ is C-equivariant. Therefore, $\Delta(C \cdot q) = C \cdot \underbrace{\Delta(q)}_{=q^d} = C \cdot q^d$. This rewrites as $\Delta([q]) = [q^d]$ (since $[q] = C \cdot q$ and $[q^d] = C \cdot q^d$). Hence, $|\Delta([q])| = |[q^d]|$. Therefore, $|[q^d]| = |\Delta([q])| = |[q]| \leq n/d$. This proves Lemma 13.190.6. □

We can now start proving the results required in Exercise 6.6.23:

Proof of Proposition 6.6.15. Define $n \in \mathbb{N}$ by $n = \ell(w)$. Then, $n = \ell(w) > 0$ (since w is nonempty), so that n is a positive integer. Also, $w \in \mathfrak{A}^n$ (since $\ell(w) = n$).

We must show that the word w is aperiodic if and only if the necklace $[w]$ is aperiodic. In other words, we must prove the following two claims:

Claim 1: If the word w is aperiodic, then the necklace $[w]$ is aperiodic.

Claim 2: If the necklace $[w]$ is aperiodic, then the word w is aperiodic.

[*Proof of Claim 1:* Assume that the word w is aperiodic. In other words, there exist no $m \geq 2$ and $u \in \mathfrak{A}^*$ satisfying $w = u^m$ (by the definition of aperiodic words).

We must prove that the necklace $[w]$ is aperiodic. Indeed, assume the contrary. Then, the necklace $[w]$ is not aperiodic. Hence, Lemma 13.190.5 (applied to $v = w$ and $k = n$) yields that there exists some $h \in \{1, 2, \ldots, n-1\}$ such that $c^h \cdot w = w$. Consider this h.

From $h \in \{1, 2, \ldots, n-1\}$, we obtain $1 \leq h \leq n-1$, so that $h \geq 1 > 0$ and $n - \underbrace{h}_{\leq n-1} \geq n - (n-1) = 1 > 0$.

We have $w = (w_1, w_2, \ldots, w_n)$ (since $w \in \mathfrak{A}^n$). Define two words p and q by $p = (w_1, w_2, \ldots, w_h)$ and $q = (w_{h+1}, w_{h+2}, \ldots, w_n)$. Then, these words p and q have lengths $\ell(p) = h$ and $\ell(q) = n - h$, and thus are nonempty (since $\ell(p) = h > 0$ and $\ell(q) = n - h > 0$). They further satisfy

$$\begin{aligned} w &= (w_1, w_2, \ldots, w_n) = (w_1, w_2, \ldots, w_h, w_{h+1}, w_{h+2}, \ldots, w_n) \\ &= \underbrace{(w_1, w_2, \ldots, w_h)}_{=p} \underbrace{(w_{h+1}, w_{h+2}, \ldots, w_n)}_{=q} = pq. \end{aligned}$$

Recall that c acts on $\mathfrak{A}^n$ by cyclically rotating n-tuples one step to the left (that is, we have $c \cdot (a_1, a_2, \ldots, a_n) = (a_2, a_3, \ldots, a_n, a_1)$ for each $(a_1, a_2, \ldots, a_n) \in \mathfrak{A}^n$). Thus, c^h acts on $\mathfrak{A}^n$ by cyclically rotating n-tuples h steps to the left. In other words,

$$c^h \cdot (a_1, a_2, \ldots, a_n) = (a_{h+1}, a_{h+2}, \ldots, a_n, a_1, a_2, \ldots, a_h)$$

for each $(a_1, a_2, \ldots, a_n) \in \mathfrak{A}^n$ (since $h \in \{1, 2, \ldots, n-1\} \subset \{0, 1, \ldots, n\}$). Applying this to $(a_1, a_2, \ldots, a_n) = (w_1, w_2, \ldots, w_n)$, we obtain

$$\begin{aligned} c^h \cdot (w_1, w_2, \ldots, w_n) &= (w_{h+1}, w_{h+2}, \ldots, w_n, w_1, w_2, \ldots, w_h) \\ &= \underbrace{(w_{h+1}, w_{h+2}, \ldots, w_n)}_{=q} \underbrace{(w_1, w_2, \ldots, w_h)}_{=p} = qp. \end{aligned}$$

Comparing this with

$$c^h \cdot \underbrace{(w_1, w_2, \ldots, w_n)}_{=w} = c^h \cdot w = w = pq,$$

we obtain $pq = qp$. Hence, Proposition 6.1.4 (applied to $u = p$ and $v = q$) yields that there exist a $t \in \mathfrak{A}^*$ and two nonnegative integers N and M [1192] such that $p = t^N$ and $q = t^M$. Consider these t, N and M. From $p = t^N$, we obtain $\ell(p) = \ell(t^N) = N\ell(t)$, so that $N\ell(t) = \ell(p) = h \neq 0$ (since $h > 0$). Hence, $N \neq 0$, so that $N \geq 1$ (because N is a nonnegative integer). Also, from $q = t^M$, we obtain $\ell(q) = \ell(t^M) = M\ell(t)$, so that $M\ell(t) = \ell(q) = n - h \neq 0$ (since $n - h > 0$). Hence, $M \neq 0$, so that $M \geq 1$ (since M is a nonnegative integer). Now, $\underbrace{N}_{\geq 1} + \underbrace{M}_{\geq 1} \geq 1 + 1 = 2$ and $w = \underbrace{p}_{=t^N} \underbrace{q}_{=t^M} = t^N t^M = t^{N+M}$. Hence, there exist $m \geq 2$ and $u \in \mathfrak{A}^*$ satisfying $w = u^m$ (namely, $m = N + M$ and $u = t$). This contradicts the fact that there exist no $m \geq 2$ and $u \in \mathfrak{A}^*$ satisfying $w = u^m$. This contradiction shows that our assumption was false. Hence, the necklace $[w]$ is aperiodic. This proves Claim 1.]

[1192] Here, we are using the letters N and M for what was called n and m in Proposition 6.1.4 (since the letter n already has a different meaning in our current situation).

[*Proof of Claim 2:* Assume that the necklace $[w]$ is aperiodic. We must prove that the word w is aperiodic. Let $m \geq 2$ and $u \in \mathfrak{A}^*$ be such that $w = u^m$. We shall derive a contradiction.

From $w = u^m$, we obtain $\ell(w) = \ell(u^m) = m\ell(u)$, so that $m\ell(u) = \ell(w) = n$. Hence, $\ell(u) = n/m$ (since $m \geq 2 > 0$). Thus, $u \in \mathfrak{A}^{n/m}$. Also, m is a divisor of n (since $m\ell(u) = n$ with $\ell(u) \in \mathbb{Z}$), and is positive (since $m \geq 2 > 0$).

Hence, Lemma 13.190.6 (applied to $d = m$ and $q = u$) yields that the n/m-necklace $[u]$ and the n-necklace $[u^m]$ satisfy $|[u^m]| = |[u]| \leq n/m$. From $w = u^m$, we obtain $|[w]| = |[u^m]| \leq n/\underbrace{m}_{\geq 2} \leq n/2 < n$ (since $n > 0$).

But Lemma 13.190.4 (applied to w instead of u) yields that $|[w]| = \ell(w)$ (since the necklace $[w]$ is aperiodic). Hence, $|[w]| = \ell(w) = n$. But this contradicts $|[w]| < n$.

Forget that we fixed m and u. We thus have obtained a contradiction for every $m \geq 2$ and $u \in \mathfrak{A}^*$ satisfying $w = u^m$. Hence, there exist no $m \geq 2$ and $u \in \mathfrak{A}^*$ satisfying $w = u^m$. In other words, the word w is aperiodic (by the definition of an aperiodic word). This proves Claim 2.]

We have now proved both Claim 1 and Claim 2. Combining these two claims, we conclude that the word w is aperiodic if and only if the necklace $[w]$ is aperiodic. This proves Proposition 6.6.15. □

Proof of Corollary 6.6.16. The word w is aperiodic and thus nonempty (since any aperiodic word is nonempty). In other words, $\ell(w) > 0$. Let $n = \ell(w)$. Hence, $w \in \mathfrak{A}^n$. Thus, $c \cdot w \in \mathfrak{A}^n$, so that $\ell(c \cdot w) = n > 0$. Hence, the word $c \cdot w$ is nonempty.

Now, $[w]$ is the n-necklace containing w. In other words, $[w]$ is the C-orbit containing w (since the n-necklaces are exactly the C-orbits on $\mathfrak{A}^n$). In other words, $[w] = C \cdot w$ (where we are using the notation $C \cdot w$ for the C-orbit $\{d \cdot w \mid d \in C\}$). The same argument (applied to $c \cdot w$ instead of w) yields that $[c \cdot w] = C \cdot (c \cdot w)$. Hence,

$$[c \cdot w] = C \cdot (c \cdot w) = \underbrace{(Cc)}_{\substack{=C \\ \text{(since } C \text{ is a group)}}} \cdot w = C \cdot w = [w].$$

But Proposition 6.6.15 shows that the word w is aperiodic if and only if the necklace $[w]$ is aperiodic. Hence, the necklace $[w]$ is aperiodic (since the word w is aperiodic). In other words, the necklace $[c \cdot w]$ is aperiodic (since $[c \cdot w] = [w]$).

However, Proposition 6.6.15 (applied to $c \cdot w$ instead of w) shows that the word $c \cdot w$ is aperiodic if and only if the necklace $[c \cdot w]$ is aperiodic. Hence, the word $c \cdot w$ is aperiodic (since the necklace $[c \cdot w]$ is aperiodic). This proves Corollary 6.6.16. □

Proof of Corollary 6.6.17. Let N be an aperiodic necklace. We must prove that N is a set of aperiodic words. In other words, we must prove that each $w \in N$ is an aperiodic word.

So let $w \in N$. We shall prove that w is an aperiodic word.

Indeed, N is a necklace, i.e., an n-necklace for some positive integer n. Consider this n. Lemma 13.190.2 (applied to $u = w$) shows that $N = [w]$. Hence, $[w]$ is an aperiodic n-necklace (since N is an aperiodic n-necklace).

The word w has length n (since $[w]$ is an n-necklace) and thus is nonempty (since n is positive). Moreover, the necklace $[w]$ is aperiodic. Hence, Proposition 6.6.15 shows that the word w is aperiodic.

Forget that we fixed w. Thus, we have proved that each $w \in N$ is an aperiodic word. This completes the proof of Corollary 6.6.17. □

Proof of Proposition 6.6.19. We shall prove the following four claims:

Claim 1: The relation $\leq_\omega$ is reflexive.

Claim 2: The relation $\leq_\omega$ is transitive.

Claim 3: The relation $\leq_\omega$ is antisymmetric.

Claim 4: The relation $\leq_\omega$ is total.

[*Proof of Claim 1:* Each $a \in \mathfrak{A}^{\mathfrak{a}}$ satisfies $aa \leq aa$. In other words, each $a \in \mathfrak{A}^{\mathfrak{a}}$ satisfies $a \leq_\omega a$ (by the definition of the relation $\leq_\omega$). In other words, the relation $\leq_\omega$ is reflexive. This proves Claim 1.]

[*Proof of Claim 2:* Let $x, y, z \in \mathfrak{A}^{\mathfrak{a}}$ satisfy $x \leq_\omega y$ and $y \leq_\omega z$. We shall show that $x \leq_\omega z$.

We have assumed that $x \leq_\omega y$. In other words, $xy \leq yx$ (by the definition of the relation $\leq_\omega$). In other words, $yx \geq xy$.

We have assumed that $y \leq_\omega z$. In other words, $yz \leq zy$ (by the definition of the relation $\leq_\omega$). In other words, $zy \geq yz$.

The word y is aperiodic (since $y \in \mathfrak{A}^{\mathfrak{a}}$) and thus nonempty (since any aperiodic word is nonempty). Hence, Corollary 6.1.6 (applied to $u = z$, $v = y$ and $w = x$) yields $zx \geq xz$. In other words, $xz \leq zx$. In other words, $x \leq_\omega z$ (by the definition of the relation $\leq_\omega$).

Now, forget that we fixed x, y, z. We thus have proved that if $x, y, z \in \mathfrak{A}^{\mathfrak{a}}$ satisfy $x \leq_\omega y$ and $y \leq_\omega z$, then $x \leq_\omega z$. In other words, the relation $\leq_\omega$ is transitive. This proves Claim 2.]

[*Proof of Claim 3:* Let $a, b \in \mathfrak{A}^{\mathfrak{a}}$ satisfy $a \leq_\omega b$ and $b \leq_\omega a$. We shall show that $a = b$.

The word b is aperiodic (since $b \in \mathfrak{A}^{\mathfrak{a}}$). In other words, b is a nonempty word with the property that

$$\text{there exist no } m \geq 2 \text{ and } u \in \mathfrak{A}^* \text{ satisfying } b = u^m \tag{13.190.1}$$

(by the definition of an aperiodic word).

We have assumed that $a \leq_\omega b$. In other words, $ab \leq ba$ (by the definition of the relation $\leq_\omega$). Likewise, $ba \leq ab$ (since $b \leq_\omega a$). Combining these two inequalities, we obtain $ab = ba$. Hence, Proposition 6.1.4 (applied to $u = a$ and $v = b$) yields that there exist a $t \in \mathfrak{A}^*$ and two nonnegative integers n and m such that $a = t^n$ and $b = t^m$. Consider these t, n and m.

From $b = t^m$, we obtain $\ell(b) = \ell(t^m) = m\ell(t)$, thus $m\ell(t) = \ell(b) \neq 0$ (since b is nonempty). Hence, $m \neq 0$, so that $m \geq 1$ (since m is a nonnegative integer). Note that the word t satisfies $b = t^m$; thus, there exists a word $u \in \mathfrak{A}^*$ satisfying $b = u^m$ (namely, $u = t$). If we had $m \geq 2$, then this would violate (13.190.1). Hence, we cannot have $m \geq 2$. Thus, $m < 2$. Combining this with $m \geq 1$, we obtain $m = 1$. Now, $b = t^m = t^1$ (since $m = 1$). Likewise, $a = t^1$. Hence, $a = t^1 = b$.

Now, forget that we fixed a, b. We thus have proved that if $a, b \in \mathfrak{A}^{\mathfrak{a}}$ satisfy $a \leq_\omega b$ and $b \leq_\omega a$, then $a = b$. In other words, the relation $\leq_\omega$ is antisymmetric. This proves Claim 3.]

[*Proof of Claim 4:* Let $a, b \in \mathfrak{A}^{\mathfrak{a}}$. We shall prove that we have $a \leq_\omega b$ or $b \leq_\omega a$.

Indeed, the relation $\leq$ on $\mathfrak{A}^*$ is total (since the lexicographic order on words is a total order). Thus, we have $ab \leq ba$ or $ba \leq ab$. In other words, we have $a \leq_\omega b$ or $b \leq_\omega a$ (because the relation $a \leq_\omega b$ is equivalent to $ab \leq ba$ [1193], whereas the relation $b \leq_\omega a$ is equivalent to $ba \leq ab$ [1194]).

Now, forget that we fixed a, b. We thus have proved that any $a, b \in \mathfrak{A}^{\mathfrak{a}}$ satisfy $a \leq_\omega b$ or $b \leq_\omega a$. In other words, the relation $\leq_\omega$ is total. This proves Claim 4.]

Now, the relation $\leq_\omega$ is reflexive (by Claim 1), transitive (by Claim 2) and antisymmetric (by Claim 3). Hence, it is the smaller-or-equal relation of a partial order. This partial order must furthermore be a total order, since the relation $\leq_\omega$ is total (by Claim 4). Thus, the relation $\leq_\omega$ is the smaller-or-equal relation of a total order. This proves Proposition 6.6.19. □

Before we prove Proposition 6.6.20, let us show some more basic lemmas:

Lemma 13.190.7. *Let $u \in \mathfrak{A}^*$ be a nonempty word. Let m be a positive integer. Then, $\left(c^i \cdot u\right)^m = c^i \cdot (u^m)$ for each $i \in \mathbb{N}$.*

Proof of Lemma 13.190.7. Let $i \in \mathbb{N}$.

Let $n = \ell(u) \cdot m$ and $d = m$. We know that m is a positive integer; in other words, d is a positive integer (since $d = m$). Also, $n = \ell(u) \cdot \underbrace{m}_{=d} = \ell(u) \cdot d$; hence, d is a divisor of n. Finally, $\ell(u)$ is a positive integer (since u is nonempty). Now, both $\ell(u)$ and m are positive integers. Hence, the product $\ell(u) \cdot m$ is a positive integer as well. In other words, n is a positive integer (since $n = \ell(u) \cdot m$).

Consider the map $\Delta : \mathfrak{A}^{n/d} \to \mathfrak{A}^n$ defined in Lemma 13.147.5.

We have $n = \ell(u) \cdot d$, so that $\ell(u) = n/d$ (since $d \neq 0$ (because d is positive)). Hence, $u \in \mathfrak{A}^{n/d}$. Thus, the definition of Δ yields $\Delta(u) = u^d = u^m$ (since $d = m$). Also, the definition of Δ yields $\Delta\left(c^i u\right) = \left(\underbrace{c^i u}_{=c^i \cdot u}\right)^d = \left(c^i \cdot u\right)^d = \left(c^i \cdot u\right)^m$ (since $d = m$). But Lemma 13.147.5(b) shows that the map Δ is C-equivariant; in other words, we have $\Delta(gw) = g \cdot \Delta(w)$ for every $g \in C$ and $w \in \mathfrak{A}^{n/d}$. Applying this to $g = c^i$ and

[1193] by the definition of the relation $\leq_\omega$

[1194] by the definition of the relation $\leq_\omega$

$w = u$, we obtain $\Delta\left(c^i u\right) = c^i \cdot \underbrace{\Delta(u)}_{=u^m} = c^i \cdot (u^m)$. Comparing this with $\Delta\left(c^i u\right) = \left(c^i \cdot u\right)^m$, we obtain $\left(c^i \cdot u\right)^m = c^i \cdot (u^m)$. This proves Lemma 13.190.7. □

Lemma 13.190.8. *Let n be a positive integer. Let $w \in \mathfrak{A}^n$ be a word. Then, $\left(c^i \cdot w\right)_j = w_{i+j}$ for each $i \in \mathbb{N}$ and each $j \in \{1, 2, \ldots, n-i\}$.*

Proof of Lemma 13.190.8. Let $i \in \mathbb{N}$, and let $j \in \{1, 2, \ldots, n-i\}$. Then, $1 \leq j \leq n-i$, so that $n-i \geq 1 \geq 0$ and thus $n \geq i$. Hence, $i \leq n$, so that $i \in \{0, 1, \ldots, n\}$ (since $i \in \mathbb{N}$).

Recall that c acts on $\mathfrak{A}^n$ by cyclically rotating n-tuples one step to the left (that is, we have $c \cdot (a_1, a_2, \ldots, a_n) = (a_2, a_3, \ldots, a_n, a_1)$ for each $(a_1, a_2, \ldots, a_n) \in \mathfrak{A}^n$). Thus, c^i acts on $\mathfrak{A}^n$ by cyclically rotating n-tuples i steps to the left. In other words, for each $(a_1, a_2, \ldots, a_n) \in \mathfrak{A}^n$, we have

$$c^i \cdot (a_1, a_2, \ldots, a_n) = (a_{i+1}, a_{i+2}, \ldots, a_n, a_1, a_2, \ldots, a_i)$$

(since $i \in \{0, 1, \ldots, n\}$). Applying this to $a_k = w_k$, we obtain

$$c^i \cdot (w_1, w_2, \ldots, w_n) = (w_{i+1}, w_{i+2}, \ldots, w_n, w_1, w_2, \ldots, w_i).$$

But $w = (w_1, w_2, \ldots, w_n)$ (since $w \in \mathfrak{A}^n$), and thus

$$c^i \cdot w = c^i \cdot (w_1, w_2, \ldots, w_n) = (w_{i+1}, w_{i+2}, \ldots, w_n, w_1, w_2, \ldots, w_i).$$

Hence, the first $n-i$ letters of the word $c^i \cdot w$ are $w_{i+1}, w_{i+2}, \ldots, w_n$ (in this order). In other words, for each $k \in \{1, 2, \ldots, n-i\}$, we have

$$\left(\text{the } k\text{-th letter of } c^i \cdot w\right) = w_{i+k}.$$

Applying this to $k = j$, we obtain

$$\left(\text{the } j\text{-th letter of } c^i \cdot w\right) = w_{i+j}$$

(since $j \in \{1, 2, \ldots, n-i\}$). Hence,

$$\left(c^i \cdot w\right)_j = \left(\text{the } j\text{-th letter of } c^i \cdot w\right) = w_{i+j}.$$

This proves Lemma 13.190.8. □

Lemma 13.190.9. *Let $w \in \mathfrak{A}^*$ be a nonempty word. Let $m \in \mathbb{N}$. Then, $\left(c^i \cdot w\right)_1 = (w^m)_{i+1}$ for each $i \in \{0, 1, \ldots, m\ell(w) - 1\}$.*

Proof of Lemma 13.190.9. Let $n = \ell(w)$; thus, $w \in \mathfrak{A}^n$. The word w is nonempty. Hence, $\ell(w)$ is a positive integer. In other words, n is a positive integer (since $n = \ell(w)$).

Let $v = c^i \cdot w$. Then, $v \in \mathfrak{A}^n$ (since $w \in \mathfrak{A}^n$). Thus, $\ell(v) = n$.

Let $i \in \{0, 1, \ldots, m\ell(w) - 1\}$. Then, $0 \leq i < m\ell(w)$, so that $m\ell(w) > 0$ and thus $m\ell(w) \neq 0$. This entails $m \neq 0$, so that $m > 0$ (since $m \in \mathbb{N}$). Hence, the word v is a prefix of v^m, and thus has the same first letter as v^m (since the word v is nonempty[1195], and thus has a first letter). In other words, $v_1 = (v^m)_1$.

Let $N = m\ell(w)$. Thus, $i \in \{0, 1, \ldots, m\ell(w) - 1\} = \{0, 1, \ldots, N-1\}$ (since $m\ell(w) = N$), so that $i + 1 \in \{1, 2, \ldots, N\}$. Hence, $1 \leq i + 1 \leq N$, so that $N \geq 1$. This shows that N is a positive integer. Also, $i + 1 \leq N$, thus $N \geq i + 1$, hence $N - i \geq 1$ and thus $1 \leq N - i$. Hence, $1 \in \{1, 2, \ldots, N - i\}$.

We have $\ell(w^m) = m\ell(w) = N$ (since $N = m\ell(w)$), so that $w^m \in \mathfrak{A}^N$. Lemma 13.190.8 (applied to N, w^m and 1 instead of n, w and j) thus yields $\left(c^i \cdot (w^m)\right)_1 = (w^m)_{i+1}$. But Lemma 13.190.7 (applied to $u = w$) yields $\left(c^i \cdot w\right)^m = c^i \cdot (w^m)$. In view of $c^i \cdot w = v$, this rewrites as $v^m = c^i \cdot (w^m)$. Hence, $(v^m)_1 = \left(c^i \cdot (w^m)\right)_1 = (w^m)_{i+1}$. Hence, $v_1 = (v^m)_1 = (w^m)_{i+1}$. In view of $v = c^i \cdot w$, this rewrites as $\left(c^i \cdot w\right)_1 = (w^m)_{i+1}$. This proves Lemma 13.190.9. □

Lemma 13.190.10. *Let $u, v \in \mathfrak{A}^*$ be two aperiodic words such that $u \neq v$ and $u \leq_\omega v$. Let g, n and m be three positive integers such that $g = n\ell(v) = m\ell(u)$. Then, there exists a $k \in \{1, 2, \ldots, g\}$ such that $(u^m)_k < (v^n)_k$, and such that every $j \in \{1, 2, \ldots, k-1\}$ satisfies $(u^m)_j = (v^n)_j$.*

[1195] because $\ell(v) = n > 0$

Proof of Lemma 13.190.10. We have $\ell\left(u^m\right) = m\ell\left(u\right) = g$ (since $g = m\ell\left(u\right)$) and $\ell\left(v^n\right) = n\ell\left(v\right) = g$ (since $g = n\ell\left(v\right)$). Hence, $\ell\left(u^m\right) = g = \ell\left(v^n\right)$; in other words, the words u^m and v^n have the same length.

Proposition 6.6.19 shows that the relation $\leq_\omega$ on the set $\mathfrak{A}^{\mathfrak{a}}$ is the smaller-or-equal relation of a total order. Hence, in particular, this relation $\leq_\omega$ is antisymmetric. Thus, we don't have $v \leq_\omega u$ (because otherwise, we could combine $u \leq_\omega v$ with $v \leq_\omega u$ to obtain $u = v$, which would contradict $u \neq v$). In other words, we don't have $vu \leq uv$ (since the relation $v \leq_\omega u$ means the same as $vu \leq uv$). In other words, we don't have $uv \geq vu$.

From $n\ell\left(v\right) = m\ell\left(u\right)$, we obtain $m\ell\left(u\right) = n\ell\left(v\right)$. Hence, Exercise 6.1.11 (applied to n and m instead of m and n) shows that $uv \geq vu$ holds if and only if $u^m \geq v^n$ holds. Therefore, we don't have $u^m \geq v^n$ (since we don't have $uv \geq vu$). Thus, we have $u^m < v^n$ (since the lexicographic order on $\mathfrak{A}^*$ is a total order). Hence, the word u^m is not a prefix of v^n [1196].

But we have $u^m \leq v^n$ (since $u^m < v^n$). In other words,

> **either** there exists an $i \in \left\{1, 2, \ldots, \min\left\{\ell\left(u^m\right), \ell\left(v^n\right)\right\}\right\}$
>
> such that $\left(\left(u^m\right)_i < \left(v^n\right)_i, \text{ and every } j \in \left\{1, 2, \ldots, i-1\right\} \text{ satisfies } \left(u^m\right)_j = \left(v^n\right)_j\right)$,
>
> **or** the word u^m is a prefix of v^n

(by the definition of the lexicographic order). Thus, there exists an $i \in \left\{1, 2, \ldots, \min\left\{\ell\left(u^m\right), \ell\left(v^n\right)\right\}\right\}$ such that

$$\left(\left(u^m\right)_i < \left(v^n\right)_i, \text{ and every } j \in \left\{1, 2, \ldots, i-1\right\} \text{ satisfies } \left(u^m\right)_j = \left(v^n\right)_j\right)$$

(since the word u^m is not a prefix of v^n). Consider this i.

We have $i \in \left\{1, 2, \ldots, \min\left\{\ell\left(u^m\right), \ell\left(v^n\right)\right\}\right\} = \left\{1, 2, \ldots, g\right\}$ (since $\min\left\{\underbrace{\ell\left(u^m\right)}_{=g}, \underbrace{\ell\left(v^n\right)}_{=g}\right\} = \min\left\{g, g\right\} = g$) and $\left(u^m\right)_i < \left(v^n\right)_i$. Moreover, every $j \in \left\{1, 2, \ldots, i-1\right\}$ satisfies $\left(u^m\right)_j = \left(v^n\right)_j$. Hence, there exists a $k \in \left\{1, 2, \ldots, g\right\}$ such that $\left(u^m\right)_k < \left(v^n\right)_k$, and such that every $j \in \left\{1, 2, \ldots, k-1\right\}$ satisfies $\left(u^m\right)_j = \left(v^n\right)_j$ (namely, $k = i$). This proves Lemma 13.190.10. □

Lemma 13.190.11. *Let $u, v \in \mathfrak{A}^*$ be two aperiodic words such that $u \neq v$. Let $g = \ell\left(u\right) \cdot \ell\left(v\right)$. Then:*

(a) *There exists some $i \in \left\{0, 1, \ldots, g-1\right\}$ satisfying $\left(c^i \cdot u\right)_1 \neq \left(c^i \cdot v\right)_1$.*

(b) *If $u \leq_\omega v$, then the smallest $i \in \left\{0, 1, \ldots, g-1\right\}$ satisfying $\left(c^i \cdot u\right)_1 \neq \left(c^i \cdot v\right)_1$ satisfies $\left(c^i \cdot u\right)_1 < \left(c^i \cdot v\right)_1$.*

(c) *If $v \leq_\omega u$, then the smallest $i \in \left\{0, 1, \ldots, g-1\right\}$ satisfying $\left(c^i \cdot u\right)_1 \neq \left(c^i \cdot v\right)_1$ satisfies $\left(c^i \cdot u\right)_1 > \left(c^i \cdot v\right)_1$.*

Proof of Lemma 13.190.11. The words u and v are aperiodic and thus nonempty (since any aperiodic word is nonempty). Let $n = \ell\left(u\right)$ and $m = \ell\left(v\right)$. Then, $n = \ell\left(u\right) > 0$ (since u is nonempty), so that n is a positive integer. Likewise, m is a positive integer. Moreover, combining the equalities $\underbrace{n}_{=\ell(u)} \ell\left(v\right) = \ell\left(u\right) \cdot \ell\left(v\right) = g$ and $\underbrace{m}_{=\ell(v)} \ell\left(u\right) = \ell\left(v\right) \cdot \ell\left(u\right) = \ell\left(u\right) \cdot \ell\left(v\right) = g$, we obtain $g = n\ell\left(v\right) = m\ell\left(u\right)$. Also, $g = \underbrace{\ell\left(u\right)}_{=n} \cdot \underbrace{\ell\left(v\right)}_{=m} = nm$; this shows that g is a positive integer (since n and m are positive integers). Note also that $v \neq u$ (since $u \neq v$) and $g = \ell\left(u\right) \cdot \ell\left(v\right) = \ell\left(v\right) \cdot \ell\left(u\right)$. Thus, u and v play symmetric roles in our situation.

The words u and v are aperiodic; i.e., we have $u, v \in \mathfrak{A}^{\mathfrak{a}}$ (since $\mathfrak{A}^{\mathfrak{a}}$ is the set of all aperiodic words).

We shall now prove the following fact:

> *Claim 1:* Assume that $u \leq_\omega v$. Then, there exists some $i \in \left\{0, 1, \ldots, g-1\right\}$ satisfying $\left(c^i \cdot u\right)_1 \neq \left(c^i \cdot v\right)_1$.

[*Proof of Claim 1:* Lemma 13.190.10 shows that there exists a $k \in \left\{1, 2, \ldots, g\right\}$ such that $\left(u^m\right)_k < \left(v^n\right)_k$, and such that every $j \in \left\{1, 2, \ldots, k-1\right\}$ satisfies $\left(u^m\right)_j = \left(v^n\right)_j$. Consider this k. But $k \in \left\{1, 2, \ldots, g\right\}$, thus

[1196] *Proof.* Assume the contrary. Thus, the word u^m is a prefix of v^n. Hence, $u^m = v^n$ (since the words u^m and v^n have the same length). This contradicts $u^m < v^n$. This contradiction shows that our assumption was false, qed.

$k-1 \in \{0,1,\ldots,g-1\} = \{0,1,\ldots,m\ell(u)-1\}$ (since $g = m\ell(u)$). Hence, Lemma 13.190.9 (applied to u and $k-1$ instead of w and i) yields $\left(c^{k-1}\cdot u\right)_1 = (u^m)_{(k-1)+1} = (u^m)_k$ (since $(k-1)+1 = k$).

Also, $k-1 \in \{0,1,\ldots,g-1\} = \{0,1,\ldots,n\ell(v)-1\}$ (since $g = n\ell(v)$). Hence, Lemma 13.190.9 (applied to v, n and $k-1$ instead of w, m and i) yields $\left(c^{k-1}\cdot v\right)_1 = (v^n)_{(k-1)+1} = (v^n)_k$ (since $(k-1)+1 = k$).

Now, recall that $(u^m)_k < (v^n)_k$. In view of $\left(c^{k-1}\cdot u\right)_1 = (u^m)_k$ and $\left(c^{k-1}\cdot v\right)_1 = (v^n)_k$, we can rewrite this as $\left(c^{k-1}\cdot u\right)_1 < \left(c^{k-1}\cdot v\right)_1$. Hence, $\left(c^{k-1}\cdot u\right)_1 \neq \left(c^{k-1}\cdot v\right)_1$. Since $k-1 \in \{0,1,\ldots,g-1\}$, this shows that there exists some $i \in \{0,1,\ldots,g-1\}$ satisfying $\left(c^{i}\cdot u\right)_1 \neq \left(c^{i}\cdot v\right)_1$ (namely, $i = k-1$). This proves Claim 1.]

We can now easily obtain the following counterpart of Claim 1 for the case when $v \leq_\omega u$:

> *Claim 2:* Assume that $v \leq_\omega u$. Then, there exists some $i \in \{0,1,\ldots,g-1\}$ satisfying $\left(c^{i}\cdot u\right)_1 \neq \left(c^{i}\cdot v\right)_1$.

[*Proof of Claim 2:* We have $v \neq u$ and $g = \ell(v)\cdot\ell(u)$. Hence, Claim 1 (applied to v and u instead of u and v) yields that there exists some $i \in \{0,1,\ldots,g-1\}$ satisfying $\left(c^{i}\cdot v\right)_1 \neq \left(c^{i}\cdot u\right)_1$. In other words, there exists some $i \in \{0,1,\ldots,g-1\}$ satisfying $\left(c^{i}\cdot u\right)_1 \neq \left(c^{i}\cdot v\right)_1$. This proves Claim 2.]

(a) Proposition 6.6.19 shows that the relation $\leq_\omega$ on the set $\mathfrak{A}^{\mathfrak{a}}$ is the smaller-or-equal relation of a total order. Thus, we have $u \leq_\omega v$ or $v \leq_\omega u$ (since $u, v \in \mathfrak{A}^{\mathfrak{a}}$). In each of these two cases, we can conclude that there exists some $i \in \{0,1,\ldots,g-1\}$ satisfying $\left(c^{i}\cdot u\right)_1 \neq \left(c^{i}\cdot v\right)_1$ (indeed, in the case $u \leq_\omega v$, this follows from Claim 1, while in the case $v \leq_\omega u$, this follows from Claim 2). This proves Lemma 13.190.11(a).

(b) Assume that $u \leq_\omega v$. Thus, Lemma 13.190.10 shows that there exists a $k \in \{1,2,\ldots,g\}$ such that $(u^m)_k < (v^n)_k$, and such that every $j \in \{1,2,\ldots,k-1\}$ satisfies $(u^m)_j = (v^n)_j$. Consider this k. We have $(u^m)_k < (v^n)_k$. This rewrites as $(u^m)_{(k-1)+1} < (v^n)_{(k-1)+1}$ (since $(k-1)+1 = k$).

It is easy to see the following:

> *Claim 3:* The number $k-1$ is the smallest $i \in \{0,1,\ldots,g-1\}$ satisfying $(u^m)_{i+1} \neq (v^n)_{i+1}$.

[*Proof of Claim 3:* The number $k-1$ is an element of $\{0,1,\ldots,g-1\}$ (since $k \in \{1,2,\ldots,g\}$) and satisfies $(u^m)_{(k-1)+1} \neq (v^n)_{(k-1)+1}$ (since $(u^m)_{(k-1)+1} < (v^n)_{(k-1)+1}$). In other words, the number $k-1$ is an $i \in \{0,1,\ldots,g-1\}$ satisfying $(u^m)_{i+1} \neq (v^n)_{i+1}$. It remains to show that $k-1$ is the **smallest** such i.

Indeed, let $i \in \{0,1,\ldots,g-1\}$ be such that $(u^m)_{i+1} \neq (v^n)_{i+1}$. We shall show that $i \geq k-1$.

Indeed, assume the contrary. Hence, $i < k-1$. Thus, $i+1 < k$, so that $i+1 \leq k-1$ (since $i+1$ and k are integers). Combining this with $i+1 > i \geq 0$ (since $i \in \{0,1,\ldots,g-1\}$), we obtain $i+1 \in \{1,2,\ldots,k-1\}$. But recall that every $j \in \{1,2,\ldots,k-1\}$ satisfies $(u^m)_j = (v^n)_j$. Applying this to $j = i+1$, we obtain $(u^m)_{i+1} = (v^n)_{i+1}$ (since $i+1 \in \{1,2,\ldots,k-1\}$). This contradicts $(u^m)_{i+1} \neq (v^n)_{i+1}$. This contradiction shows that our assumption was false. Hence, $i \geq k-1$ is proved.

Now, forget that we fixed i. We thus have shown that $i \geq k-1$ for each $i \in \{0,1,\ldots,g-1\}$ such that $(u^m)_{i+1} \neq (v^n)_{i+1}$. Thus, $k-1$ is the **smallest** $i \in \{0,1,\ldots,g-1\}$ satisfying $(u^m)_{i+1} \neq (v^n)_{i+1}$ (since we already know that $k-1$ is such an i). This proves Claim 3.]

From Claim 3, we quickly obtain the following:

> *Claim 4:* The smallest $i \in \{0,1,\ldots,g-1\}$ satisfying $(u^m)_{i+1} \neq (v^n)_{i+1}$ satisfies $(u^m)_{i+1} < (v^n)_{i+1}$.

[*Proof of Claim 4:* We must show that the smallest $i \in \{0,1,\ldots,g-1\}$ satisfying $(u^m)_{i+1} \neq (v^n)_{i+1}$ satisfies $(u^m)_{i+1} < (v^n)_{i+1}$. But we know (from Claim 3) that this smallest i is $k-1$. Hence, we need to prove that $i = k-1$ satisfies $(u^m)_{i+1} < (v^n)_{i+1}$. In other words, we need to prove that $(u^m)_{(k-1)+1} < (v^n)_{(k-1)+1}$. But we have proven this already. Thus, Claim 4 is proved.]

On the other hand, each $i \in \{0,1,\ldots,g-1\}$ satisfies

$$(u^m)_{i+1} = \left(c^{i}\cdot u\right)_1 . \tag{13.190.2}$$

[*Proof of (13.190.2):* Let $i \in \{0,1,\ldots,g-1\}$. Then, $i \in \{0,1,\ldots,g-1\} = \{0,1,\ldots,m\ell(u)-1\}$ (since $g = m\ell(u)$). Hence, Lemma 13.190.9 (applied to $w = u$) yields $\left(c^{i}\cdot u\right)_1 = (u^m)_{i+1}$. This proves (13.190.2).]

Furthermore, each $i \in \{0,1,\ldots,g-1\}$ satisfies

$$(v^n)_{i+1} = \left(c^{i}\cdot v\right)_1 . \tag{13.190.3}$$

[*Proof of (13.190.3):* The proof of (13.190.3) proceeds in the same way as the proof of (13.190.2), but using v and n instead of u and m.]

Using the equalities (13.190.2) and (13.190.3), we can rewrite Claim 4 as follows:

Claim 5: The smallest $i \in \{0, 1, \ldots, g-1\}$ satisfying $\left(c^i \cdot u\right)_1 \neq \left(c^i \cdot v\right)_1$ satisfies $\left(c^i \cdot u\right)_1 < \left(c^i \cdot v\right)_1$.

But Claim 5 is precisely the statement of Lemma 13.190.11(b). Thus, Lemma 13.190.11(b) is proven.

(c) Assume that $v \leq_\omega u$. Thus, Lemma 13.190.11(b) (applied to v and u instead of u and v) yields the following:

Claim 6: The smallest $i \in \{0, 1, \ldots, g-1\}$ satisfying $\left(c^i \cdot v\right)_1 \neq \left(c^i \cdot u\right)_1$ satisfies $\left(c^i \cdot v\right)_1 < \left(c^i \cdot u\right)_1$.

In other words, we have the following:

Claim 7: The smallest $i \in \{0, 1, \ldots, g-1\}$ satisfying $\left(c^i \cdot u\right)_1 \neq \left(c^i \cdot v\right)_1$ satisfies $\left(c^i \cdot u\right)_1 > \left(c^i \cdot v\right)_1$.

(Indeed, Claim 7 is a restatement of Claim 6, since the statement "$\left(c^i \cdot u\right)_1 \neq \left(c^i \cdot v\right)_1$" is equivalent to "$\left(c^i \cdot v\right)_1 \neq \left(c^i \cdot u\right)_1$", whereas the statement "$\left(c^i \cdot u\right)_1 > \left(c^i \cdot v\right)_1$" is equivalent to "$\left(c^i \cdot v\right)_1 < \left(c^i \cdot u\right)_1$".)

But Claim 7 is precisely the statement of Lemma 13.190.11(c). Thus, Lemma 13.190.11(c) is proven. □

Lemma 13.190.12. *Let $a, b \in \mathfrak{A}$ be two letters. Let $p, q \in \mathfrak{A}^*$ be two words such that $\ell(p) = \ell(q)$. Then, we have $ap \leq bq$ if and only if either $a < b$ or ($a = b$ and $pa \leq qb$).*

Proof of Lemma 13.190.12. Recall that $a \in \mathfrak{A}$; hence, a is identified with the one-letter word (a). Thus, $\ell(a) = 1$. Similarly, $\ell(b) = 1$. Hence, $\ell(a) = 1 = \ell(b)$, so that $\ell(a) \leq \ell(b)$.

We shall next prove the following:

Claim 1: If $ap \leq bq$, then we have either $a < b$ or ($a = b$ and $pa \leq qb$).

[*Proof of Claim 1:* Assume that $ap \leq bq$. We must prove that we have either $a < b$ or ($a = b$ and $pa \leq qb$). If $a < b$, then this is obviously true. Thus, for the rest of this proof of Claim 1, we can WLOG assume that we don't have $a < b$. Assume this.

We don't have $a < b$. Hence, we have $a \geq b$ (since the order on $\mathfrak{A}$ is a total order).

But Proposition 6.1.2(f) (applied to p, b and q instead of b, c and d) yields $a \leq b$ (since $ap \leq bq$ and $\ell(a) \leq \ell(b)$). Combining this with $a \geq b$, we obtain $a = b$. Thus, $ap = bp$, so that $bp = ap \leq bq$. Hence, Proposition 6.1.2(c) (applied to b, p and q instead of a, c and d) yields $p \leq q$. Therefore, Proposition 6.1.2(j) (applied to p, q and b instead of a, b and c) yields $pb \leq qb$ (since $\ell(p) \geq \ell(q)$ (because $\ell(p) = \ell(q)$)). But $a = b$, so that $pa = pb \leq qb$. Hence, we have shown that $a = b$ and $pa \leq qb$. Thus, we have either $a < b$ or ($a = b$ and $pa \leq qb$). This proves Claim 1.]

Claim 2: If $a < b$, then $ap \leq bq$.

[*Proof of Claim 2:* Assume that $a < b$. If a was a prefix of b, then we would have $a = b$ (since a and b are one-letter words), which would contradict $a < b$. Thus, a is not a prefix of b.

But $a \leq b$ (since $a < b$). Hence, Proposition 6.1.2(d) (applied to p, b and q instead of b, c and d) yields that either we have $ap \leq bq$ or the word a is a prefix of b. Since the word a is not a prefix of b, we thus conclude that $ap \leq bq$. This proves Claim 2.]

Claim 3: If $a = b$ and $pa \leq qb$, then $ap \leq bq$.

[*Proof of Claim 3:* Assume that $a = b$ and $pa \leq qb$. We must prove that $ap \leq bq$.

We have $\ell(p) \leq \ell(q)$ (since $\ell(p) = \ell(q)$) and $pa \leq qb$. Thus, Proposition 6.1.2(f) (applied to p, a, q and b instead of a, b, c and d) yields $p \leq q$. Hence, Proposition 6.1.2(b) (applied to p and q instead of c and d) yields $ap \leq aq = bq$ (since $a = b$). This proves Claim 3.]

Claim 4: If we have either $a < b$ or ($a = b$ and $pa \leq qb$), then we have $ap \leq bq$.

[*Proof of Claim 4:* This follows by combining Claim 2 and Claim 3.]

Combining Claim 1 with Claim 4, we conclude that we have $ap \leq bq$ if and only if either $a < b$ or ($a = b$ and $pa \leq qb$). (Indeed, the "only if" part of this statement follows from Claim 1, whereas the "if" part follows from Claim 4.) This proves Lemma 13.190.12. □

Proof of Proposition 6.6.20. (b) Assume that $u \neq v$. Let $g = \ell(u) \cdot \ell(v)$. Then, Lemma 13.190.11(a) yields that there exists some $i \in \{0, 1, \ldots, g-1\}$ satisfying $\left(c^i \cdot u\right)_1 \neq \left(c^i \cdot v\right)_1$. This i must satisfy $i \in$

$\{0,1,\ldots,g-1\} \subset \mathbb{N}$ and $\left(c^i \cdot u\right)_1 \neq \left(c^i \cdot v\right)_1$. Hence, there exists some $i \in \mathbb{N}$ satisfying $\left(c^i \cdot u\right)_1 \neq \left(c^i \cdot v\right)_1$. This proves Proposition 6.6.20(b).

(d) From $n\ell(u) = m\ell(v)$, we obtain $m\ell(v) = n\ell(u)$. Hence, Exercise 6.1.11 (applied to n, m, u and v instead of m, n, v and u) shows that $vu \geq uv$ holds if and only if $v^n \geq u^m$ holds. In other words, we have the logical equivalence $(vu \geq uv) \iff (v^m \geq u^n)$.

Now, we have the following chain of logical equivalences:

$$(u \leq_\omega v) \iff (uv \leq vu) \qquad \text{(by the definition of the relation } \leq_\omega\text{)}$$
$$\iff (vu \geq uv) \iff (v^m \geq u^n) \iff (u^n \leq v^m).$$

In other words, we have $u \leq_\omega v$ if and only if $u^n \leq v^m$. This proves Proposition 6.6.20(d).

(c) Our proof relies on the following five claims:

> *Claim 1:* If $u \leq_\omega v$, then the smallest $i \in \mathbb{N}$ satisfying $\left(c^i \cdot u\right)_1 \neq \left(c^i \cdot v\right)_1$ **either** does not exist **or** satisfies $\left(c^i \cdot u\right)_1 < \left(c^i \cdot v\right)_1$.

[*Proof of Claim 1:* Assume that $u \leq_\omega v$. We must prove that the smallest $i \in \mathbb{N}$ satisfying $\left(c^i \cdot u\right)_1 \neq \left(c^i \cdot v\right)_1$ **either** does not exist **or** satisfies $\left(c^i \cdot u\right)_1 < \left(c^i \cdot v\right)_1$.

If $u = v$, then this is obvious[1197]. Hence, for the rest of this proof of Claim 1, we WLOG assume that $u \neq v$. Let $g = \ell(u) \cdot \ell(v)$.

Lemma 13.190.11(a) yields that there exists some $i \in \{0,1,\ldots,g-1\}$ satisfying $\left(c^i \cdot u\right)_1 \neq \left(c^i \cdot v\right)_1$. Let j denote the **smallest** such i. Thus, $j \in \{0,1,\ldots,g-1\}$ and $\left(c^j \cdot u\right)_1 \neq \left(c^j \cdot v\right)_1$. From $j \in \{0,1,\ldots,g-1\}$, we obtain $j \leq g-1$; hence, $g-1 \geq j$.

Furthermore, Lemma 13.190.11(b) yields that the smallest $i \in \{0,1,\ldots,g-1\}$ satisfying $\left(c^i \cdot u\right)_1 \neq \left(c^i \cdot v\right)_1$ satisfies $\left(c^i \cdot u\right)_1 < \left(c^i \cdot v\right)_1$. Since we have denoted this smallest i by j, we can restate this as follows: The number $i = j$ satisfies $\left(c^i \cdot u\right)_1 < \left(c^i \cdot v\right)_1$. In other words, we have $\left(c^j \cdot u\right)_1 < \left(c^j \cdot v\right)_1$.

We have defined j to be the **smallest** $i \in \{0,1,\ldots,g-1\}$ satisfying $\left(c^i \cdot u\right)_1 \neq \left(c^i \cdot v\right)_1$. Thus, any $i \in \{0,1,\ldots,g-1\}$ satisfying $\left(c^i \cdot u\right)_1 \neq \left(c^i \cdot v\right)_1$ is greater or equal to j. In other words, for any $i \in \{0,1,\ldots,g-1\}$ satisfying $\left(c^i \cdot u\right)_1 \neq \left(c^i \cdot v\right)_1$, we have

$$i \geq j. \tag{13.190.4}$$

Now, recall that $j \in \{0,1,\ldots,g-1\} \subset \mathbb{N}$ and $\left(c^j \cdot u\right)_1 \neq \left(c^j \cdot v\right)_1$. In other words, j is an $i \in \mathbb{N}$ satisfying $\left(c^i \cdot u\right)_1 \neq \left(c^i \cdot v\right)_1$. Moreover, j is the **smallest** such i [1198].

So we know that j is the smallest $i \in \mathbb{N}$ satisfying $\left(c^i \cdot u\right)_1 \neq \left(c^i \cdot v\right)_1$. Thus, in particular, this smallest i exists. Moreover, this smallest i satisfies $\left(c^i \cdot u\right)_1 < \left(c^i \cdot v\right)_1$ (because this smallest i is j, but we know that $\left(c^j \cdot u\right)_1 < \left(c^j \cdot v\right)_1$). Thus, we have shown that the smallest $i \in \mathbb{N}$ satisfying $\left(c^i \cdot u\right)_1 \neq \left(c^i \cdot v\right)_1$ satisfies $\left(c^i \cdot u\right)_1 < \left(c^i \cdot v\right)_1$. This proves Claim 1.]

> *Claim 2:* If $v \leq_\omega u$, then the smallest $i \in \mathbb{N}$ satisfying $\left(c^i \cdot v\right)_1 \neq \left(c^i \cdot u\right)_1$ **either** does not exist **or** satisfies $\left(c^i \cdot v\right)_1 < \left(c^i \cdot u\right)_1$.

[*Proof of Claim 2:* Claim 2 is just Claim 1 with the roles of u and v swapped.]

> *Claim 3:* If $v \leq_\omega u$, then the smallest $i \in \mathbb{N}$ satisfying $\left(c^i \cdot u\right)_1 \neq \left(c^i \cdot v\right)_1$ **either** does not exist **or** satisfies $\left(c^i \cdot u\right)_1 > \left(c^i \cdot v\right)_1$.

[1197] Indeed, in this case, the smallest $i \in \mathbb{N}$ satisfying $\left(c^i \cdot u\right)_1 \neq \left(c^i \cdot v\right)_1$ does not exist, because $\left(c^i \cdot \underbrace{u}_{=v}\right)_1 = \left(c^i \cdot v\right)_1$ for each $i \in \mathbb{N}$.

[1198] *Proof.* We shall show that if $i \in \mathbb{N}$ satisfies $\left(c^i \cdot u\right)_1 \neq \left(c^i \cdot v\right)_1$, then $i \geq j$.

Indeed, let $i \in \mathbb{N}$ satisfy $\left(c^i \cdot u\right)_1 \neq \left(c^i \cdot v\right)_1$. We must prove that $i \geq j$. If $i \geq g-1$, then this is obvious (because in this case, we have $i \geq g-1 \geq j$). Hence, for the rest of this proof, we WLOG assume that $i < g-1$. Hence, $i \leq g-1$ and therefore $i \in \{0,1,\ldots,g-1\}$. Hence, (13.190.4) yields $i \geq j$.

Now, forget that we fixed i. We thus have shown that if $i \in \mathbb{N}$ satisfies $\left(c^i \cdot u\right)_1 \neq \left(c^i \cdot v\right)_1$, then $i \geq j$. In other words, any $i \in \mathbb{N}$ satisfying $\left(c^i \cdot u\right)_1 \neq \left(c^i \cdot v\right)_1$ must satisfy $i \geq j$. Therefore, j is the **smallest** $i \in \mathbb{N}$ satisfying $\left(c^i \cdot u\right)_1 \neq \left(c^i \cdot v\right)_1$ (because we already know that j is such an i).

[*Proof of Claim 3:* Claim 3 is just a restatement of Claim 2, since the statement "$\left(c^i \cdot u\right)_1 \neq \left(c^i \cdot v\right)_1$" is equivalent to "$\left(c^i \cdot u\right)_1 \neq \left(c^i \cdot v\right)_1$" and since the statement "$\left(c^i \cdot u\right)_1 > \left(c^i \cdot v\right)_1$" is equivalent to "$\left(c^i \cdot v\right)_1 < \left(c^i \cdot u\right)_1$".]

> *Claim 4:* If we don't have $u \leq_\omega v$, then the smallest $i \in \mathbb{N}$ satisfying $\left(c^i \cdot u\right)_1 \neq \left(c^i \cdot v\right)_1$ exists but does **not** satisfy $\left(c^i \cdot u\right)_1 < \left(c^i \cdot v\right)_1$.

[*Proof of Claim 4:* Assume that we don't have $u \leq_\omega v$. We must prove that the smallest $i \in \mathbb{N}$ satisfying $\left(c^i \cdot u\right)_1 \neq \left(c^i \cdot v\right)_1$ exists but does **not** satisfy $\left(c^i \cdot u\right)_1 < \left(c^i \cdot v\right)_1$.

Proposition 6.6.19 shows that the relation $\leq_\omega$ on the set $\mathfrak{A}^{\mathfrak{a}}$ is the smaller-or-equal relation of a total order. Hence, if we had $u = v$, then we would have $u \leq_\omega v$, which would contradict our assumption that we don't have $u \leq_\omega v$. Hence, we must have $u \neq v$.

Thus, Proposition 6.6.20(b) shows that there exists some $i \in \mathbb{N}$ satisfying $\left(c^i \cdot u\right)_1 \neq \left(c^i \cdot v\right)_1$. Hence, the **smallest** such i exists.

Recall that the relation $\leq_\omega$ on the set $\mathfrak{A}^{\mathfrak{a}}$ is the smaller-or-equal relation of a total order. Hence, we must have $v \leq_\omega u$ (since we don't have $u \leq_\omega v$). Thus, Claim 3 yields that the smallest $i \in \mathbb{N}$ satisfying $\left(c^i \cdot u\right)_1 \neq \left(c^i \cdot v\right)_1$ **either** does not exist **or** satisfies $\left(c^i \cdot u\right)_1 > \left(c^i \cdot v\right)_1$. Since we know that this smallest i exists (indeed, we have proved this in the previous paragraph), we thus conclude that the smallest $i \in \mathbb{N}$ satisfying $\left(c^i \cdot u\right)_1 \neq \left(c^i \cdot v\right)_1$ satisfies $\left(c^i \cdot u\right)_1 > \left(c^i \cdot v\right)_1$. Hence, this smallest i does **not** satisfy $\left(c^i \cdot u\right)_1 < \left(c^i \cdot v\right)_1$ (because $\left(c^i \cdot u\right)_1 < \left(c^i \cdot v\right)_1$ would contradict $\left(c^i \cdot u\right)_1 > \left(c^i \cdot v\right)_1$).

Altogether, we thus have shown that the smallest $i \in \mathbb{N}$ satisfying $\left(c^i \cdot u\right)_1 \neq \left(c^i \cdot v\right)_1$ exists but does **not** satisfy $\left(c^i \cdot u\right)_1 < \left(c^i \cdot v\right)_1$. This proves Claim 4.]

> *Claim 5:* If the smallest $i \in \mathbb{N}$ satisfying $\left(c^i \cdot u\right)_1 \neq \left(c^i \cdot v\right)_1$ **either** does not exist **or** satisfies $\left(c^i \cdot u\right)_1 < \left(c^i \cdot v\right)_1$, then we have $u \leq_\omega v$.

[*Proof of Claim 5:* Claim 5 is just the contrapositive of Claim 4, and thus holds (since Claim 4 holds).]

Now, Proposition 6.6.20(c) follows by combining Claim 1 with Claim 5. (Indeed, Claim 1 is the "only if" part of Proposition 6.6.20(c), whereas Claim 5 is the "if" part.)

(a) The words u and v are aperiodic and thus nonempty (since any aperiodic word is nonempty).

Let $n = \ell(v)$ and $m = \ell(u)$. Then, $n = \ell(v) > 0$ (since v is nonempty), so that n is a positive integer. Likewise, m is a positive integer. Hence, $m > 0$.

Let $g = nm$. Then, g is a positive integer (since n and m are positive integers). Hence, $g \geq 1$, so that $1 \in \{1, 2, \ldots, g\}$.

Moreover, combining the equalities $n \underbrace{\ell(u)}_{=m} = nm = g$ and $m \underbrace{\ell(v)}_{=n} = mn = nm = g$, we obtain $g = n\ell(u) = m\ell(v)$. Hence, Proposition 6.6.20(d) shows that we have $u \leq_\omega v$ if and only if $u^n \leq v^m$. In other words, we have the logical equivalence

$$(u \leq_\omega v) \iff (u^n \leq v^m). \tag{13.190.5}$$

The word u is aperiodic. Thus, Corollary 6.6.16 (applied to $w = u$) yields that the word $c \cdot u$ is aperiodic. Likewise, the word $c \cdot v$ is aperiodic. Moreover, $u \in \mathfrak{A}^m$ (since $\ell(u) = m$), so that $c \cdot u \in \mathfrak{A}^m$ and thus $\ell(c \cdot u) = m$. Likewise, $\ell(c \cdot v) = n$ (since $\ell(v) = n$). Furthermore, comparing $n \underbrace{\ell(c \cdot u)}_{=m} = nm$ with $m \underbrace{\ell(c \cdot v)}_{=n} = mn = nm$, we obtain $n\ell(c \cdot u) = m\ell(c \cdot v)$. Hence, Proposition 6.6.20(d) (applied to $c \cdot u$ and $c \cdot v$ instead of u and v) shows that we have $c \cdot u \leq_\omega c \cdot v$ if and only if $(c \cdot u)^n \leq (c \cdot v)^m$. In other words, we have the logical equivalence

$$(c \cdot u \leq_\omega c \cdot v) \iff ((c \cdot u)^n \leq (c \cdot v)^m). \tag{13.190.6}$$

Lemma 13.190.7 (applied to v and 1 instead of u and i) yields $\left(c^1 \cdot v\right)^m = c^1 \cdot (v^m)$. In view of $c^1 = c$, this rewrites as $(c \cdot v)^m = c \cdot (v^m)$. The same argument (with v and m replaced by u and n) yields $(c \cdot u)^n = c \cdot (u^n)$.

The word u^n has length $\ell(u^n) = n\ell(u) = g$, and thus can be written as $u^n = (u^n)_1 (u^n)_2 \cdots (u^n)_g$. Define a word $p \in \mathfrak{A}^*$ by $p = (u^n)_2 (u^n)_3 \cdots (u^n)_g$. Since $g \geq 1$, we thus have $p \in \mathfrak{A}^{g-1}$ (since $(u^n)_2, (u^n)_3, \ldots, (u^n)_g$ are single letters), so that $\ell(p) = g - 1$. Also, define a letter $a \in \mathfrak{A}$ by $a = (u^n)_1$. (This is well-defined, since

$1 \in \{1, 2, \ldots, g\}$.) Thus,

$$\underbrace{a}_{=(u^n)_1} \underbrace{p}_{=(u^n)_2 (u^n)_3 \cdots (u^n)_g} = (u^n)_1 \left((u^n)_2 (u^n)_3 \cdots (u^n)_g \right) = (u^n)_1 (u^n)_2 \cdots (u^n)_g$$

$$= u^n. \tag{13.190.7}$$

Lemma 13.147.4 (applied to g, a and p instead of m, u and v) yields $c(ap) = pa$. In view of $ap = u^n$, this rewrites as $c(u^n) = pa$. Hence,

$$(c \cdot u)^n = c \cdot (u^n) = c(u^n) = pa. \tag{13.190.8}$$

The word v^m has length $\ell(v^m) = m\ell(v) = g$, and thus can be written as $v^m = (v^m)_1 (v^m)_2 \cdots (v^m)_g$. Define a word $q \in \mathfrak{A}^*$ by $q = (v^m)_2 (v^m)_3 \cdots (v^m)_g$. Since $g \geq 1$, we thus have $q \in \mathfrak{A}^{g-1}$ (since $(v^m)_2, (v^m)_3, \ldots, (v^m)_g$ are single letters), so that $\ell(q) = g - 1$. Also, define a letter $b \in \mathfrak{A}$ by $b = (v^m)_1$. (This is well-defined, since $1 \in \{1, 2, \ldots, g\}$.) Thus,

$$\underbrace{b}_{=(v^m)_1} \underbrace{q}_{=(v^m)_2 (v^m)_3 \cdots (v^m)_g} = (v^m)_1 \left((v^m)_2 (v^m)_3 \cdots (v^m)_g \right) = (v^m)_1 (v^m)_2 \cdots (v^m)_g$$

$$= v^m. \tag{13.190.9}$$

Lemma 13.147.4 (applied to g, b and q instead of m, u and v) yields $c(bq) = qb$. In view of $bq = v^m$, this rewrites as $c(v^m) = qb$. Hence,

$$(c \cdot v)^m = c \cdot (v^m) = c(v^m) = qb. \tag{13.190.10}$$

Comparing $\ell(p) = g - 1$ with $\ell(q) = g - 1$, we obtain $\ell(p) = \ell(q)$. Thus, Lemma 13.190.12 yields that we have $ap \leq bq$ if and only if either $a < b$ or ($a = b$ and $pa \leq qb$). In other words, we have the following equivalence:

$$(ap \leq bq) \iff (\text{either } a < b \text{ or } (a = b \text{ and } pa \leq qb)). \tag{13.190.11}$$

The word u^n starts with the word u (since n is a positive integer), and thus has the same first letter as the word u (since the word u is nonempty). In other words, $(u^n)_1 = u_1$. Hence, $a = (u^n)_1 = u_1$. The same argument (applied to v, m and b instead of u, n and a) yields $b = v_1$.

Now, we have the following chain of logical equivalences:

$$\begin{aligned}
(u \leq_\omega v) &\iff \left(\underbrace{u^n}_{\substack{=ap \\ (\text{by } (13.190.7))}} \leq \underbrace{v^m}_{\substack{=bq \\ (\text{by } (13.190.9))}} \right) \qquad (\text{by } (13.190.5)) \\
&\iff (ap \leq bq) \\
&\iff \left(\text{either } \underbrace{a}_{=u_1} < \underbrace{b}_{=v_1} \text{ or } \left(\underbrace{a}_{=u_1} = \underbrace{b}_{=v_1} \text{ and } \underbrace{pa}_{\substack{=(c \cdot u)^n \\ (\text{by } (13.190.8))}} \leq \underbrace{qb}_{\substack{=(c \cdot v)^m \\ (\text{by } (13.190.10))}} \right) \right) \\
&\qquad (\text{by } (13.190.11)) \\
&\iff \left(\text{either } u_1 < v_1 \text{ or } \left(u_1 = v_1 \text{ and } \underbrace{(c \cdot u)^n \leq (c \cdot v)^m}_{\substack{\iff (c \cdot u \leq_\omega c \cdot v) \\ (\text{by } (13.190.6))}} \right) \right) \\
&\iff (\text{either } u_1 < v_1 \text{ or } (u_1 = v_1 \text{ and } c \cdot u \leq_\omega c \cdot v)).
\end{aligned}$$

In other words, we have $u \leq_\omega v$ if and only if either $u_1 < v_1$ or ($u_1 = v_1$ and $c \cdot u \leq_\omega c \cdot v$). This proves Proposition 6.6.20(a). □

Finally, let us prove Proposition 6.6.22:

Proof of Proposition 6.6.22. We have $w = (w_1, w_2, \ldots, w_n)$ (since $w \in \mathfrak{A}^n$).

(a) Let $h \in \{1, 2, \ldots, n\}$. Let z be the cycle of τ that contains h. Thus, $h \in z$. Hence, Proposition 6.6.7(a) yields that $[w_{\tau,h}] = \{w_{\tau,i} \mid i \in z\}$. Moreover, Proposition 6.6.7(d) shows that the set $\{w_{\tau,i} \mid i \in z\}$ is an aperiodic necklace. In other words, the set $[w_{\tau,h}]$ is an aperiodic necklace (since $[w_{\tau,h}] = \{w_{\tau,i} \mid i \in z\}$).

But Proposition 6.6.5(a) shows that the word $w_{\tau,h}$ is nonempty and has length $\operatorname{ord}_\tau(h)$. Hence, Proposition 6.6.15 (applied to $w_{\tau,h}$ instead of w) shows that the word $w_{\tau,h}$ is aperiodic if and only if the necklace $[w_{\tau,h}]$ is aperiodic. Hence, the word $w_{\tau,h}$ is aperiodic (since the necklace $[w_{\tau,h}]$ is aperiodic).

Forget that we fixed h. We thus have proved that the word $w_{\tau,h}$ is aperiodic for each $h \in \{1, 2, \ldots, n\}$. In other words, the words $w_{\tau,1}, w_{\tau,2}, \ldots, w_{\tau,n}$ are aperiodic. This proves Proposition 6.6.22(a).

(b) Proposition 6.6.22(a) shows that the words $w_{\tau,1}, w_{\tau,2}, \ldots, w_{\tau,n}$ are aperiodic. Thus, these words belong to $\mathfrak{A}^{\mathfrak{a}}$ (since $\mathfrak{A}^{\mathfrak{a}}$ is the set of all aperiodic words). Hence, the chain of inequalities $w_{\tau,1} \leq_\omega w_{\tau,2} \leq_\omega \cdots \leq_\omega w_{\tau,n}$ makes sense. It remains to prove that it holds.

Let α and β be two elements of $\{1, 2, \ldots, n\}$ such that $\alpha < \beta$. We shall show that $w_{\tau,\alpha} \leq_\omega w_{\tau,\beta}$.

Indeed, assume the contrary. Thus, it is **not** true that $w_{\tau,\alpha} \leq_\omega w_{\tau,\beta}$.

Let us notice that each $i \in \mathbb{N}$ satisfies

$$\left(c^i \cdot w_{\tau,\alpha}\right)_1 = w_{\tau^{i+1}(\alpha)} \qquad (13.190.12)$$

and

$$\left(c^i \cdot w_{\tau,\beta}\right)_1 = w_{\tau^{i+1}(\beta)} \qquad (13.190.13)$$

[*Proof of (13.190.12) and (13.190.13):* Let $i \in \mathbb{N}$. Then, Proposition 6.6.5(b) (applied to $h = \tau^i(\alpha)$) yields that the first letter of the word $w_{\tau,\tau^i(\alpha)}$ is $w_{\tau(\tau^i(\alpha))}$.

But Proposition 6.6.5(e) (applied to $h = \alpha$) yields $w_{\tau,\tau^i(\alpha)} = c^i \cdot w_{\tau,\alpha}$. Hence, $\left(w_{\tau,\tau^i(\alpha)}\right)_1 = \left(c^i \cdot w_{\tau,\alpha}\right)_1$. Thus,

$$\begin{aligned}\left(c^i \cdot w_{\tau,\alpha}\right)_1 &= \left(w_{\tau,\tau^i(\alpha)}\right)_1 = \left(\text{the first letter of the word } w_{\tau,\tau^i(\alpha)}\right) = w_{\tau(\tau^i(\alpha))} \\ &\qquad \left(\text{since the first letter of the word } w_{\tau,\tau^i(\alpha)} \text{ is } w_{\tau(\tau^i(\alpha))}\right) \\ &= w_{\tau^{i+1}(\alpha)} \qquad \left(\text{since } \tau\left(\tau^i(\alpha)\right) = \tau^{i+1}(\alpha)\right).\end{aligned}$$

This proves (13.190.12). The same argument (but applied to β instead of α) proves (13.190.13).]

Proposition 6.6.22(a) shows that the words $w_{\tau,1}, w_{\tau,2}, \ldots, w_{\tau,n}$ are aperiodic. Hence, in particular, the words $w_{\tau,\alpha}$ and $w_{\tau,\beta}$ are aperiodic. Thus, Proposition 6.6.20(c) (applied to $u = w_{\tau,\alpha}$ and $v = w_{\tau,\beta}$) yields that we have $w_{\tau,\alpha} \leq_\omega w_{\tau,\beta}$ if and only if the smallest $i \in \mathbb{N}$ satisfying $\left(c^i \cdot w_{\tau,\alpha}\right)_1 \neq \left(c^i \cdot w_{\tau,\beta}\right)_1$ **either** does not exist **or** satisfies $\left(c^i \cdot w_{\tau,\alpha}\right)_1 < \left(c^i \cdot w_{\tau,\beta}\right)_1$. Thus, it is **not** true that the smallest $i \in \mathbb{N}$ satisfying $\left(c^i \cdot w_{\tau,\alpha}\right)_1 \neq \left(c^i \cdot w_{\tau,\beta}\right)_1$ **either** does not exist **or** satisfies $\left(c^i \cdot w_{\tau,\alpha}\right)_1 < \left(c^i \cdot w_{\tau,\beta}\right)_1$ (because it is **not** true that $w_{\tau,\alpha} \leq_\omega w_{\tau,\beta}$). In other words, the smallest $i \in \mathbb{N}$ satisfying $\left(c^i \cdot w_{\tau,\alpha}\right)_1 \neq \left(c^i \cdot w_{\tau,\beta}\right)_1$ exists but does **not** satisfy $\left(c^i \cdot w_{\tau,\alpha}\right)_1 < \left(c^i \cdot w_{\tau,\beta}\right)_1$. In view of (13.190.12) and (13.190.13), we can rewrite this fact as follows: The smallest $i \in \mathbb{N}$ satisfying $w_{\tau^{i+1}(\alpha)} \neq w_{\tau^{i+1}(\beta)}$ exists but does **not** satisfy $w_{\tau^{i+1}(\alpha)} < w_{\tau^{i+1}(\beta)}$.

Consider this i, and denote it by j. Thus, $i = j$ does **not** satisfy $w_{\tau^{i+1}(\alpha)} < w_{\tau^{i+1}(\beta)}$. In other words, we do **not** have $w_{\tau^{j+1}(\alpha)} < w_{\tau^{j+1}(\beta)}$. In other words, we have $w_{\tau^{j+1}(\alpha)} \geq w_{\tau^{j+1}(\beta)}$.

We have defined j to be the smallest $i \in \mathbb{N}$ satisfying $w_{\tau^{i+1}(\alpha)} \neq w_{\tau^{i+1}(\beta)}$. Hence, j is an $i \in \mathbb{N}$ satisfying $w_{\tau^{i+1}(\alpha)} \neq w_{\tau^{i+1}(\beta)}$. In other words, j is an element of $\mathbb{N}$ and satisfies $w_{\tau^{j+1}(\alpha)} \neq w_{\tau^{j+1}(\beta)}$. Combining $w_{\tau^{j+1}(\alpha)} \geq w_{\tau^{j+1}(\beta)}$ with $w_{\tau^{j+1}(\alpha)} \neq w_{\tau^{j+1}(\beta)}$, we obtain

$$w_{\tau^{j+1}(\alpha)} > w_{\tau^{j+1}(\beta)}. \qquad (13.190.14)$$

On the other hand, j is the **smallest** $i \in \mathbb{N}$ satisfying $w_{\tau^{i+1}(\alpha)} \neq w_{\tau^{i+1}(\beta)}$. Thus, every such $i \in \mathbb{N}$ is $\geq j$. In other words,

$$\text{if } i \in \mathbb{N} \text{ satisfies } w_{\tau^{i+1}(\alpha)} \neq w_{\tau^{i+1}(\beta)}, \text{ then } i \geq j. \qquad (13.190.15)$$

Hence,

$$\text{every } i \in \{0, 1, \ldots, j-1\} \text{ satisfies } w_{\tau^{i+1}(\alpha)} = w_{\tau^{i+1}(\beta)}. \qquad (13.190.16)$$

[*Proof of (13.190.16):* Let $i \in \{0, 1, \ldots, j-1\}$. Thus, $i \geq 0$ and $i \leq j - 1 < j$. If we had $w_{\tau^{i+1}(\alpha)} \neq w_{\tau^{i+1}(\beta)}$, then we would have $i \geq j$ (by (13.190.15)), which would contradict $i < j$. Thus, we cannot have $w_{\tau^{i+1}(\alpha)} \neq w_{\tau^{i+1}(\beta)}$. Hence, we must have $w_{\tau^{i+1}(\alpha)} = w_{\tau^{i+1}(\beta)}$. This proves (13.190.16).]

Thus, Lemma 6.6.6(h) yields $w_{\tau^{j+1}(\alpha)} \leq w_{\tau^{j+1}(\beta)}$. This contradicts (13.190.14). This contradiction shows that our assumption was wrong. Hence, $w_{\tau,\alpha} \leq_\omega w_{\tau,\beta}$ is proven.

Now, forget that we fixed α and β. We thus have shown that if α and β are two elements of $\{1, 2, \ldots, n\}$ such that $\alpha < \beta$, then $w_{\tau,\alpha} \leq_\omega w_{\tau,\beta}$. In other words, we have $w_{\tau,1} \leq_\omega w_{\tau,2} \leq_\omega \cdots \leq_\omega w_{\tau,n}$. This proves Proposition 6.6.22(b). □

We have now proven Proposition 6.6.15, Corollary 6.6.16, Corollary 6.6.17, Proposition 6.6.19, Proposition 6.6.20 and Proposition 6.6.22. Thus, Exercise 6.6.23 is solved.

13.191. **Solution to Exercise 6.6.30.** *Solution to Exercise 6.6.30.* Let us first show a simple lemma about multisets of necklaces:

Lemma 13.191.1. *Let M be a finite multiset of necklaces. Let $M' = \biguplus_{N \in M} N$. (We are here using the fact that each necklace $N \in M$ is a set, thus a multiset.)*

Let $u \in \mathfrak{A}^*$ be a nonempty word. Then,

$$(\textit{multiplicity of } u \textit{ in } M') = (\textit{multiplicity of } [u] \textit{ in } M).$$

Proof of Lemma 13.191.1. We note the following simple fact:

> *Claim 1:* Let N be a necklace. Then, the statement $u \in N$ is equivalent to the statement $N = [u]$.

[*Proof of Claim 1:* The implication $(u \in N) \Longrightarrow (N = [u])$ follows from Lemma 13.190.2. The converse implication $(N = [u]) \Longrightarrow (u \in N)$ follows from Lemma 13.190.1. Combining these two implications, we obtain the equivalence $(u \in N) \Longleftrightarrow (N = [u])$. Thus, Claim 1 is proven.]

Let us write the finite multiset M in the form $M = \{N_1, N_2, \ldots, N_k\}_{\text{multiset}}$, where $N_1, N_2, \ldots, N_k$ are necklaces. Thus,

$$\begin{aligned} &(\text{multiplicity of } [u] \text{ in } M) \\ &= (\text{the number of all } i \in \{1, 2, \ldots, k\} \text{ such that } N_i = [u]) \end{aligned} \tag{13.191.1}$$

and

$$\biguplus_{N \in M} N = N_1 \uplus N_2 \uplus \cdots \uplus N_k$$

(where we are using the notation $M_1 \uplus M_2 \uplus \cdots \uplus M_k$ for a multiset union $\biguplus_{s \in \{1,2,\ldots,k\}} M_s$). Hence,

$$M' = \biguplus_{N \in M} N = N_1 \uplus N_2 \uplus \cdots \uplus N_k,$$

so that

$$\begin{aligned} &(\text{multiplicity of } u \text{ in } M') \\ &= (\text{multiplicity of } u \text{ in } N_1 \uplus N_2 \uplus \cdots \uplus N_k) \\ &= \sum_{i=1}^{k} \underbrace{(\text{multiplicity of } u \text{ in } N_i)}_{\substack{= \begin{cases} 1, & \text{if } u \in N_i; \\ 0, & \text{if } u \notin N_i \end{cases} \\ \text{(since } N_i \text{ is a set (because any necklace is a set))}}} = \sum_{i=1}^{k} \begin{cases} 1, & \text{if } u \in N_i; \\ 0, & \text{if } u \notin N_i \end{cases} \\ &= (\text{the number of all } i \in \{1, 2, \ldots, k\} \text{ such that } u \in N_i) \\ &= (\text{the number of all } i \in \{1, 2, \ldots, k\} \text{ such that } N_i = [u]) \\ &\qquad \left(\begin{array}{c} \text{because for any } i \in \{1, 2, \ldots, k\}, \text{ the statement } u \in N_i \text{ is equivalent} \\ \text{to the statement } N_i = [u] \text{ (by Claim 1, applied to } N = N_i) \end{array}\right) \\ &= (\text{multiplicity of } [u] \text{ in } M) \qquad (\text{by } (13.191.1)). \end{aligned}$$

This proves Lemma 13.191.1. □

We can now prove one "half" of Theorem 6.6.29:

Lemma 13.191.2. *We have* $\operatorname{RG} \circ \operatorname{GR} = \operatorname{id}$.

Proof of Lemma 13.191.2. Let $w \in \mathfrak{A}^*$. Let $n = \ell(w)$; thus, $w \in \mathfrak{A}^n$ and $w = (w_1, w_2, \ldots, w_n)$. Let τ be the permutation $(\operatorname{std} w)^{-1} \in \mathfrak{S}_n$. Then, the definition of the map GR shows that

$$\operatorname{GR} w = \left\{[w]_z \mid z \text{ is a cycle of } \tau\right\}_{\text{multiset}}.$$

Let us denote this multiset $\operatorname{GR} w$ by M. Thus,

$$M = \operatorname{GR} w = \left\{[w]_z \mid z \text{ is a cycle of } \tau\right\}_{\text{multiset}}.$$

Also, $M = \operatorname{GR} w \in \mathfrak{M}\mathfrak{N}^{\mathfrak{a}}$ (as we have seen in Definition 6.6.12); in other words, M is a finite multiset of aperiodic necklaces (since $\mathfrak{M}\mathfrak{N}^{\mathfrak{a}}$ is the set of all finite multisets of aperiodic necklaces).

Next, we shall show the following:

Claim 1: Let u be a nonempty word. Let z be a cycle of τ. Then,

$$(\text{the number of all } i \in z \text{ such that } u = w_{\tau,i}) = \delta_{[u],[w]_z}.$$

[*Proof of Claim 1:* The definition of the aperiodic necklace $[w]_z$ yields $[w]_z = \{w_{\tau,i} \mid i \in z\}$. We are in one of the following two cases:

Case 1: We have $[u] = [w]_z$.

Case 2: We have $[u] \neq [w]_z$.

Let us first consider Case 1. In this case, we have $[u] = [w]_z$. Hence, $\delta_{[u],[w]_z} = 1$.

Lemma 13.190.1 yields $u \in [u] = [w]_z = \{w_{\tau,i} \mid i \in z\}$. Proposition 6.6.7(b) shows that any two distinct elements α and β of z satisfy $w_{\tau,\alpha} \neq w_{\tau,\beta}$. In other words, the words $w_{\tau,i}$ for all $i \in z$ are distinct. Hence, there exists **at most one** $i \in z$ such that $u = w_{\tau,i}$. But we also know that there exists **at least one** such i (because $u \in \{w_{\tau,i} \mid i \in z\}$).

Combining the results of the previous two sentences, we conclude that there exists **exactly one** $i \in z$ such that $u = w_{\tau,i}$. In other words,

$$(\text{the number of all } i \in z \text{ such that } u = w_{\tau,i}) = 1.$$

Comparing this with $\delta_{[u],[w]_z} = 1$, we obtain

$$(\text{the number of all } i \in z \text{ such that } u = w_{\tau,i}) = \delta_{[u],[w]_z}.$$

Thus, Claim 1 is proved in Case 1.

Let us now consider Case 2. In this case, we have $[u] \neq [w]_z$. Hence, $\delta_{[u],[w]_z} = 0$.

Also, there exists no $i \in z$ such that $u = w_{\tau,i}$ [1199]. In other words,

$$(\text{the number of all } i \in z \text{ such that } u = w_{\tau,i}) = 0.$$

Comparing this with $\delta_{[u],[w]_z} = 0$, we obtain

$$(\text{the number of all } i \in z \text{ such that } u = w_{\tau,i}) = \delta_{[u],[w]_z}.$$

Thus, Claim 1 is proved in Case 2.

We have now proved Claim 1 in both Cases 1 and 2. Hence, Claim 1 always holds.]

Let $M' = \biguplus_{N \in M} N$. (We are here using the fact that each necklace $N \in M$ is a set, thus a multiset.) Then, M' is a multiset of aperiodic words[1200].

Next, we claim that

$$M' = \left\{w_{\tau,1}, w_{\tau,2}, \ldots, w_{\tau,n}\right\}_{\text{multiset}}. \tag{13.191.2}$$

[*Proof of (13.191.2):* Proposition 6.6.22(a) shows that the words $w_{\tau,1}, w_{\tau,2}, \ldots, w_{\tau,n}$ are aperiodic. Hence, $\{w_{\tau,1}, w_{\tau,2}, \ldots, w_{\tau,n}\}_{\text{multiset}}$ is a multiset of aperiodic words. Thus, both M' and $\{w_{\tau,1}, w_{\tau,2}, \ldots, w_{\tau,n}\}_{\text{multiset}}$ are multisets of aperiodic words.

[1199] *Proof.* Assume the contrary. Thus, there exists some $i \in z$ such that $u = w_{\tau,i}$. In other words, $u \in \{w_{\tau,i} \mid i \in z\}$. This rewrites as $u \in [w]_z$ (since $[w]_z = \{w_{\tau,i} \mid i \in z\}$). Since $[w]_z$ is a necklace, we can thus apply Lemma 13.190.2 to $N = [w]_z$. As a result, we obtain $[w]_z = [u]$. In other words, $[u] = [w]_z$. But this contradicts $[u] \neq [w]_z$. This contradiction shows that our assumption was false, qed.

[1200] Indeed, this was shown in Definition 6.6.26.

Now, let u be an aperiodic word. Hence, u is a nonempty word (since any aperiodic word is nonempty). Lemma 13.191.1 thus yields that

$$\begin{aligned}
(\text{multiplicity of } u \text{ in } M') &= \left(\text{multiplicity of } [u] \text{ in } \underbrace{M}_{=\left\{[w]_z \mid z \text{ is a cycle of } \tau\right\}_{\text{multiset}}}\right) \\
&= \left(\text{multiplicity of } [u] \text{ in } \left\{[w]_z \mid z \text{ is a cycle of } \tau\right\}_{\text{multiset}}\right) \\
&= \left(\text{the number of all cycles } z \text{ of } \tau \text{ such that } [u] = [w]_z\right)
\end{aligned}$$

(by the definition of the multiset $\left\{[w]_z \mid z \text{ is a cycle of } \tau\right\}_{\text{multiset}}$). Comparing this with

$$\begin{aligned}
\sum_{z \text{ is a cycle of } \tau} \delta_{[u],[w]_z} &= \sum_{\substack{z \text{ is a cycle of } \tau;\\ [u]=[w]_z}} \underbrace{\delta_{[u],[w]_z}}_{\substack{=1\\ (\text{since } [u]=[w]_z)}} + \sum_{\substack{z \text{ is a cycle of } \tau;\\ [u]\neq[w]_z}} \underbrace{\delta_{[u],[w]_z}}_{\substack{=0\\ (\text{since } [u]\neq[w]_z)}} \\
&\qquad \left(\begin{array}{c}\text{since each cycle } z \text{ of } \tau \text{ satisfies}\\ \text{either } [u]=[w]_z \text{ or } [u]\neq[w]_z \text{ (but not both)}\end{array}\right) \\
&= \sum_{\substack{z \text{ is a cycle of } \tau;\\ [u]=[w]_z}} 1 + \underbrace{\sum_{\substack{z \text{ is a cycle of } \tau;\\ [u]\neq[w]_z}} 0}_{=0} = \sum_{\substack{z \text{ is a cycle of } \tau;\\ [u]=[w]_z}} 1 \\
&= \left(\text{the number of all cycles } z \text{ of } \tau \text{ such that } [u] = [w]_z\right) \cdot 1 \\
&= \left(\text{the number of all cycles } z \text{ of } \tau \text{ such that } [u] = [w]_z\right),
\end{aligned}$$

we obtain

$$(\text{multiplicity of } u \text{ in } M') = \sum_{z \text{ is a cycle of } \tau} \delta_{[u],[w]_z}. \tag{13.191.3}$$

On the other hand, the definition of the multiset $\left\{w_{\tau,1}, w_{\tau,2}, \ldots, w_{\tau,n}\right\}_{\text{multiset}}$ yields

$$\begin{aligned}
&\left(\text{multiplicity of } u \text{ in } \left\{w_{\tau,1}, w_{\tau,2}, \ldots, w_{\tau,n}\right\}_{\text{multiset}}\right) \\
&= \left(\text{the number of all } i \in \{1,2,\ldots,n\} \text{ such that } u = w_{\tau,i}\right) \\
&= \sum_{z \text{ is a cycle of } \tau} \left(\text{the number of all } i \in z \text{ such that } u = w_{\tau,i}\right)
\end{aligned}$$

(since the cycles of τ are subsets of $\{1,2,\ldots,n\}$, and since each $i \in \{1,2,\ldots,n\}$ belongs to exactly one of these cycles). Hence,

$$\begin{aligned}
&\left(\text{multiplicity of } u \text{ in } \left\{w_{\tau,1}, w_{\tau,2}, \ldots, w_{\tau,n}\right\}_{\text{multiset}}\right) \\
&= \sum_{z \text{ is a cycle of } \tau} \underbrace{\left(\text{the number of all } i \in z \text{ such that } u = w_{\tau,i}\right)}_{\substack{=\delta_{[u],[w]_z}\\ (\text{by Claim 1})}} \\
&= \sum_{z \text{ is a cycle of } \tau} \delta_{[u],[w]_z}.
\end{aligned}$$

Comparing this with (13.191.3), we obtain

$$(\text{multiplicity of } u \text{ in } M') = \left(\text{multiplicity of } u \text{ in } \left\{w_{\tau,1}, w_{\tau,2}, \ldots, w_{\tau,n}\right\}_{\text{multiset}}\right).$$

Now, forget that we fixed u. We thus have proved that

$$(\text{multiplicity of } u \text{ in } M') = \left(\text{multiplicity of } u \text{ in } \left\{w_{\tau,1}, w_{\tau,2}, \ldots, w_{\tau,n}\right\}_{\text{multiset}}\right)$$

for each aperiodic word u. In other words, each aperiodic word u appears in the multisets M' and $\left\{w_{\tau,1}, w_{\tau,2}, \ldots, w_{\tau,n}\right\}_{\text{multiset}}$ with equal multiplicity. Since both M' and $\left\{w_{\tau,1}, w_{\tau,2}, \ldots, w_{\tau,n}\right\}_{\text{multiset}}$ are multisets of aperiodic words, we thus conclude that the multisets M' and $\left\{w_{\tau,1}, w_{\tau,2}, \ldots, w_{\tau,n}\right\}_{\text{multiset}}$ contain the same elements with equal multiplicities. In other words, these multisets M' and $\left\{w_{\tau,1}, w_{\tau,2}, \ldots, w_{\tau,n}\right\}_{\text{multiset}}$ are equal. In other words, $M' = \left\{w_{\tau,1}, w_{\tau,2}, \ldots, w_{\tau,n}\right\}_{\text{multiset}}$. This proves (13.191.2).]

Proposition 6.6.22(b) yields $w_{\tau,1} \leq_\omega w_{\tau,2} \leq_\omega \cdots \leq_\omega w_{\tau,n}$. From $M' = \{w_{\tau,1}, w_{\tau,2}, \ldots, w_{\tau,n}\}_{\text{multiset}}$ and $w_{\tau,1} \leq_\omega w_{\tau,2} \leq_\omega \cdots \leq_\omega w_{\tau,n}$, we conclude that $(w_{\tau,1}, w_{\tau,2}, \ldots, w_{\tau,n})$ is the $\leq_\omega$-increasing list of M'. Thus, the $\leq_\omega$-increasing list of M' has n elements.

Recall that $M' = \biguplus_{N \in M} N$. Hence, Definition 6.6.26 tells us that $\operatorname{RG}(M)$ can be constructed as follows: Let $(m_1, m_2, \ldots, m_n)$ be the $\leq_\omega$-increasing list of M'. (This is well-defined, since the $\leq_\omega$-increasing list of M' has n elements.) For each $i \in \{1, 2, \ldots, n\}$, let ℓ_i be the last letter of the nonempty word m_i. Then, the definition of RG yields $\operatorname{RG}(M) = (\ell_1, \ell_2, \ldots, \ell_n)$.

We shall now use this to show that $\operatorname{RG}(M) = w$.

Indeed, we first observe that $(m_1, m_2, \ldots, m_n) = (w_{\tau,1}, w_{\tau,2}, \ldots, w_{\tau,n})$, since both lists $(m_1, m_2, \ldots, m_n)$ and $(w_{\tau,1}, w_{\tau,2}, \ldots, w_{\tau,n})$ are the $\leq_\omega$-increasing list of M'.

Now, for each $i \in \{1, 2, \ldots, n\}$, we have

$$\begin{aligned}
\ell_i &= \left(\text{the last letter of the nonempty word } \underbrace{m_i}_{\substack{=w_{\tau,i} \\ (\text{since } (m_1,m_2,\ldots,m_n)=(w_{\tau,1},w_{\tau,2},\ldots,w_{\tau,n}))}}\right) \\
&\qquad (\text{since } \ell_i \text{ was defined to be the last letter of the nonempty word } m_i) \\
&= (\text{the last letter of the nonempty word } w_{\tau,i}) \\
&= w_i \qquad \left(\begin{array}{c}\text{since the last letter of the word } w_{\tau,i} \text{ is } w_i \\ (\text{by Proposition 6.6.5(c), applied to } h = i)\end{array}\right).
\end{aligned}$$

In other words, $(\ell_1, \ell_2, \ldots, \ell_n) = (w_1, w_2, \ldots, w_n)$. Comparing this with $w = (w_1, w_2, \ldots, w_n)$, we obtain $(\ell_1, \ell_2, \ldots, \ell_n) = w$. Now, $\operatorname{RG}(M) = (\ell_1, \ell_2, \ldots, \ell_n) = w$.

But $\operatorname{GR} w = M$ (since M was defined as $\operatorname{GR} w$). Now,

$$(\operatorname{RG} \circ \operatorname{GR})(w) = \operatorname{RG}\left(\underbrace{\operatorname{GR} w}_{=M}\right) = \operatorname{RG}(M) = w = \operatorname{id}(w).$$

Forget that we fixed w. We thus have shown that $(\operatorname{RG} \circ \operatorname{GR})(w) = \operatorname{id}(w)$ for each $w \in \mathfrak{A}^*$. In other words, we have $\operatorname{RG} \circ \operatorname{GR} = \operatorname{id}$. This proves Lemma 13.191.2. □

Now, we shall approach the proof of $\operatorname{GR} \circ \operatorname{RG} = \operatorname{id}$. As was explained in the Hint, we can now try to obtain $\operatorname{RG} \circ \operatorname{GR} = \operatorname{id}$ by finding a bijection $\mathfrak{A}^* \to \mathfrak{MN}^{\mathfrak{a}}$, or, more precisely, a bijection between certain finite subsets of $\mathfrak{A}^*$ and $\mathfrak{MN}^{\mathfrak{a}}$. In order to define these subsets, we first introduce a notation:

Definition 13.191.3. Let M be a finite multiset of finite sets (for example, of necklaces). Then, $\operatorname{sum} M$ shall denote the sum of the sizes of all sets $N \in M$ (counted with multiplicities). Formally speaking, this can be defined as follows: Set

$$\operatorname{sum} M = \sum_{N \in \operatorname{Supp} M} (\text{multiplicity of } N \text{ in } M) \cdot |N|. \tag{13.191.4}$$

Equivalently, if $M = \{N_1, N_2, \ldots, N_k\}_{\text{multiset}}$, then

$$\operatorname{sum} M = |N_1| + |N_2| + \cdots + |N_k|. \tag{13.191.5}$$

Example 13.191.4. Applying (13.191.5), we obtain

$$\operatorname{sum}(\{\{1,4\}, \{2,3\}, \{1,2,5\}, \{2,3\}\}_{\text{multiset}}) = |\{1,4\}| + |\{2,3\}| + |\{1,2,5\}| + |\{2,3\}| = 2 + 2 + 3 + 2 = 9.$$

Definition 13.191.5. Let $\mathfrak{B}$ be a subset of $\mathfrak{A}$.

(a) For any $n \in \mathbb{N}$, we have $\mathfrak{B}^n \subset \mathfrak{B}^* \subset \mathfrak{A}^*$ (since $\mathfrak{B} \subset \mathfrak{A}$). Thus, elements of $\mathfrak{B}^n$ are words in $\mathfrak{A}^*$.

(b) A $\mathfrak{B}$-*necklace* will mean a necklace that is a subset of $\mathfrak{B}^*$ (that is, a necklace consisting of words in $\mathfrak{B}^*$).

(c) For any $n \in \mathbb{N}$, we let $\mathfrak{MN}^{\mathfrak{a}}_{\mathfrak{B},n}$ denote the set of all finite multisets M of aperiodic $\mathfrak{B}$-necklaces satisfying $\operatorname{sum} M = n$.

Example 13.191.6. For this example, let $\mathfrak{A} = \{1, 2, \ldots, 9\}$ and $\mathfrak{B} = \{2, 4, 6, 8\}$. Then, the necklace $[286] = \{286, 862, 628\}$ is an aperiodic $\mathfrak{B}$-necklace, but the necklace $[25] = \{25, 52\}$ is not a $\mathfrak{B}$-necklace (since $25 \notin \mathfrak{B}^*$). The multiset $M = \{[286], [24], [24]\}_{\text{multiset}}$ is a finite multiset of aperiodic $\mathfrak{B}$-necklaces and satisfies $\operatorname{sum} M = |[286]| + |[24]| + |[24]| = 3 + 2 + 2 = 7$; thus, it belongs to $\mathfrak{MN}^{\mathfrak{a}}_{\mathfrak{B},7}$.

The sets we introduced in Definition 13.191.5 interact nicely with the Gessel-Reutenauer bijection: For any subset $\mathfrak{B}$ of $\mathfrak{A}$ and any $n \in \mathbb{N}$, the map $\operatorname{GR} : \mathfrak{A}^* \to \mathfrak{MN}^{\mathfrak{a}}$ restricts to a map from $\mathfrak{B}^n$ to $\mathfrak{MN}^{\mathfrak{a}}_{\mathfrak{B},n}$. More precisely:

Lemma 13.191.7. *Let $\mathfrak{B}$ be a subset of $\mathfrak{A}$. Let $n \in \mathbb{N}$. Then,* $\operatorname{GR}(\mathfrak{B}^n) \subset \mathfrak{MN}^{\mathfrak{a}}_{\mathfrak{B},n}$.

Proof of Lemma 13.191.7. Let $w \in \mathfrak{B}^n$. We shall show that $\operatorname{GR} w \in \mathfrak{MN}^{\mathfrak{a}}_{\mathfrak{B},n}$.

Indeed, $w \in \mathfrak{B}^n \subset \mathfrak{A}^n$ (since $\mathfrak{B} \subset \mathfrak{A}$). Hence, $\ell(w) = n$. Let τ be the permutation $(\operatorname{std} w)^{-1} \in \mathfrak{S}_n$. The definition of GR yields

$$\operatorname{GR} w = \{[w]_z \mid z \text{ is a cycle of } \tau\}_{\text{multiset}}. \tag{13.191.6}$$

Now, let z be a cycle of τ. Then, the necklace $[w]_z$ is a subset of $\mathfrak{B}^*$ [1201]. Hence, the necklace $[w]_z$ is a $\mathfrak{B}$-necklace, and thus an aperiodic $\mathfrak{B}$-necklace (since we have seen in the definition of $[w]_z$ that $[w]_z$ is aperiodic).

Forget that we fixed z. We thus have proved that $[w]_z$ is an aperiodic $\mathfrak{B}$-necklace whenever z is a cycle of τ. Hence, the multiset $\{[w]_z \mid z \text{ is a cycle of } \tau\}_{\text{multiset}}$ is a multiset of aperiodic $\mathfrak{B}$-necklaces. In view of (13.191.6), this rewrites as follows: The multiset $\operatorname{GR} w$ is a multiset of aperiodic $\mathfrak{B}$-necklaces. Moreover, this multiset $\operatorname{GR} w$ is clearly finite.

Let us now show that $\operatorname{sum}(\operatorname{GR} w) = n$. Indeed, the cycles of τ form a set partition of the set $\{1, 2, \ldots, n\}$ (that is, they are disjoint nonempty sets, and their union is $\{1, 2, \ldots, n\}$). Hence, the sum of their sizes is $|\{1, 2, \ldots, n\}|$. In other words, $\sum_{z \text{ is a cycle of } \tau} |z| = |\{1, 2, \ldots, n\}|$.

But let us recall that $\operatorname{sum}(\operatorname{GR} w)$ was defined to be the sum of the sizes of all sets $N \in \operatorname{GR} w$ (counted with multiplicities). Since $\operatorname{GR} w = \{[w]_z \mid z \text{ is a cycle of } \tau\}_{\text{multiset}}$, these sets $N \in \operatorname{GR} w$ are precisely the necklaces $[w]_z$ where z ranges over all cycles of τ; hence, the previous sentence says that $\operatorname{sum}(\operatorname{GR} w)$ is the sum of the sizes of the latter necklaces. In other words,

$$\operatorname{sum}(\operatorname{GR} w) = \sum_{z \text{ is a cycle of } \tau} \left| \underbrace{[w]_z}_{\substack{=\{w_{\tau,i} \mid i \in z\} \\ \text{(by the definition of } [w]_z)}} \right| = \sum_{z \text{ is a cycle of } \tau} \underbrace{|\{w_{\tau,i} \mid i \in z\}|}_{\substack{=|z| \\ \text{(by Proposition 6.6.7(c))}}}$$
$$= \sum_{z \text{ is a cycle of } \tau} |z| = |\{1, 2, \ldots, n\}| = n.$$

Altogether, we now know that the multiset $\operatorname{GR} w$ is a finite multiset of aperiodic $\mathfrak{B}$-necklaces and satisfies $\operatorname{sum}(\operatorname{GR} w) = n$. In other words, $\operatorname{GR} w$ is a finite multiset M of aperiodic $\mathfrak{B}$-necklaces satisfying $\operatorname{sum} M = n$. In other words, $\operatorname{GR} w \in \mathfrak{MN}^{\mathfrak{a}}_{\mathfrak{B},n}$ (since $\mathfrak{MN}^{\mathfrak{a}}_{\mathfrak{B},n}$ was defined as the set of all finite multisets M of aperiodic $\mathfrak{B}$-necklaces satisfying $\operatorname{sum} M = n$).

Forget that we fixed w. We thus have shown that $\operatorname{GR} w \in \mathfrak{MN}^{\mathfrak{a}}_{\mathfrak{B},n}$ for each $w \in \mathfrak{B}^n$. In other words, $\operatorname{GR}(\mathfrak{B}^n) \subset \mathfrak{MN}^{\mathfrak{a}}_{\mathfrak{B},n}$. This proves Lemma 13.191.7. □

The following fact is almost trivial:

Lemma 13.191.8. *Let $\mathfrak{B}$ be a subset of $\mathfrak{A}$. Let $u \in \mathfrak{B}^*$ be a nonempty word. Then, $[u]$ is a $\mathfrak{B}$-necklace.*

Proof of Lemma 13.191.8. The word $u \in \mathfrak{B}^*$ is nonempty. Thus, there exists a positive integer p such that $u \in \mathfrak{B}^p$. Consider this p. Hence, $u \in \mathfrak{B}^p \subset \mathfrak{A}^p$ (since $\mathfrak{B} \subset \mathfrak{A}$). Therefore, $[u]$ is a well-defined p-necklace (since p is a positive integer), hence a necklace.

[1201] This follows in a straightforward way from the definition of $[w]_z$. (In more detail: The definition of $[w]_z$ yields $[w]_z = \{w_{\tau,i} \mid i \in z\}$. But all letters of the word w belong to $\mathfrak{B}$ (since $w \in \mathfrak{B}^n$). Hence, for each $i \in z$, the word $w_{\tau,i}$ consists entirely of letters in $\mathfrak{B}$ (since the definition of $w_{\tau,i}$ shows that $w_{\tau,i}$ consists of some letters of w, but all letters of w belong to $\mathfrak{B}$), and thus belongs to $\mathfrak{B}^*$. In other words, $\{w_{\tau,i} \mid i \in z\}$ is a subset of $\mathfrak{B}^*$. In other words, $[w]_z$ is a subset of $\mathfrak{B}^*$ (since $[w]_z = \{w_{\tau,i} \mid i \in z\}$). Qed.)

Lemma 13.190.1 yields that $u \in [u]$. In other words, $[u]$ contains u.

The cyclic group C acts on both $\mathfrak{B}^p$ and $\mathfrak{A}^p$ in the same way; thus, the C-set $\mathfrak{B}^p$ is a subset of the C-set $\mathfrak{A}^p$. Hence, the orbit $C \cdot u$ of u is the same no matter whether we are considering u as an element of $\mathfrak{B}^p$ or as an element of the larger C-set $\mathfrak{A}^p$. But considering u as an element of $\mathfrak{B}^p$, we clearly see that this orbit $C \cdot u$ is an orbit of the C-set $\mathfrak{B}^p$; thus, $C \cdot u \subset \mathfrak{B}^p$.

The set $[u]$ is a p-necklace. In other words, $[u]$ is an orbit of the C-action on $\mathfrak{A}^p$ (since this is what "p-necklace" means). Since $[u]$ contains u, we thus conclude that $[u]$ is the orbit of the C-action on $\mathfrak{A}^p$ that contains u. In other words, $[u] = C \cdot u$. Hence, $[u] = C \cdot u \subset \mathfrak{B}^p \subset \mathfrak{B}^*$. In other words, $[u]$ is a subset of $\mathfrak{B}^*$.

Thus, $[u]$ is a necklace that is a subset of $\mathfrak{B}^*$ (since $[u]$ is a necklace). In other words, $[u]$ is a $\mathfrak{B}$-necklace (since a $\mathfrak{B}$-necklace is defined to mean a necklace that is a subset of $\mathfrak{B}^*$). This proves Lemma 13.191.8. □

We shall now use the CFL factorization (introduced in Definition 6.1.25) to establish a bijection between $\mathfrak{B}^n$ and $\mathfrak{M}\mathfrak{N}^{\mathfrak{a}}_{\mathfrak{B},n}$. We recall that every word has a unique CFL factorization (as we have seen in Theorem 6.1.27); thus, we can speak of "the" CFL factorization of a word.

Proposition 13.191.9. *Let $\mathfrak{B}$ be a subset of $\mathfrak{A}$. Let $n \in \mathbb{N}$. Let $w \in \mathfrak{B}^n$. Let $(a_1, a_2, \ldots, a_k)$ be the CFL factorization of the word w. (This factorization exists and is unique, as we proved in Theorem 6.1.27.) Then, $\{[a_1], [a_2], \ldots, [a_k]\}_{\text{multiset}}$ is a well-defined element of $\mathfrak{M}\mathfrak{N}^{\mathfrak{a}}_{\mathfrak{B},n}$.*

Proof of Proposition 13.191.9. We have $w \in \mathfrak{B}^n \subset \mathfrak{B}^*$. Hence, each letter of w belongs to $\mathfrak{B}$. Moreover, from $w \in \mathfrak{B}^n$, we obtain $\ell(w) = n$.

Recall that $(a_1, a_2, \ldots, a_k)$ is the CFL factorization of the word w. According to Definition 6.1.25, this means that $(a_1, a_2, \ldots, a_k)$ is a tuple of Lyndon words satisfying $w = a_1 a_2 \cdots a_k$ and $a_1 \geq a_2 \geq \cdots \geq a_k$. Thus, in particular, $(a_1, a_2, \ldots, a_k)$ is a tuple of Lyndon words. In other words, $a_1, a_2, \ldots, a_k$ are Lyndon words.

Now, we claim the following:

Claim 1: Let $i \in \{1, 2, \ldots, k\}$. Then:
(a) The set $[a_i]$ is a well-defined aperiodic $\mathfrak{B}$-necklace.
(b) We have $|[a_i]| = \ell(a_i)$.

[*Proof of Claim 1:* The word a_i is a Lyndon word (since $a_1, a_2, \ldots, a_k$ are Lyndon words), and thus nonempty (since any Lyndon word is nonempty by definition). Thus, the necklace $[a_i]$ is well-defined.

From $w = a_1 a_2 \cdots a_k$, we conclude that each letter of a_i is a letter of w. Thus, each letter of a_i belongs to $\mathfrak{B}$ (since each letter of w belongs to $\mathfrak{B}$). Hence, $a_i \in \mathfrak{B}^*$. Thus, Lemma 13.191.8 (applied to $u = a_i$) shows that $[a_i]$ is a $\mathfrak{B}$-necklace.

Lemma 13.190.3 (applied to $u = a_i$) shows that the necklace $[a_i]$ is aperiodic (since a_i is a Lyndon word). Thus, $[a_i]$ is an aperiodic $\mathfrak{B}$-necklace (since we already know that $[a_i]$ is a $\mathfrak{B}$-necklace). This proves Claim 1 **(a)**.

(b) The word a_i is nonempty and the necklace $[a_i]$ is aperiodic (as we know); thus, Lemma 13.190.4 (applied to $u = a_i$) yields $|[a_i]| = \ell(a_i)$. This proves Claim 1 **(b)**.]

Now, for each $i \in \{1, 2, \ldots, k\}$, the set $[a_i]$ is a well-defined aperiodic $\mathfrak{B}$-necklace (by Claim 1 **(a)**). In other words, $[a_1], [a_2], \ldots, [a_k]$ are well-defined aperiodic $\mathfrak{B}$-necklaces. Thus, $\{[a_1], [a_2], \ldots, [a_k]\}_{\text{multiset}}$ is a finite multiset of aperiodic $\mathfrak{B}$-necklaces. Moreover, (13.191.5) (applied to $M = \{[a_1], [a_2], \ldots, [a_k]\}_{\text{multiset}}$ and $N_i = [a_i]$) yields

$$\operatorname{sum}(\{[a_1], [a_2], \ldots, [a_k]\}_{\text{multiset}}) = |[a_1]| + |[a_2]| + \cdots + |[a_k]| = \sum_{i=1}^{k} \underbrace{|[a_i]|}_{\substack{=\ell(a_i) \\ \text{(by Claim 1 (b))}}} = \sum_{i=1}^{k} \ell(a_i).$$

Comparing this with

$$\ell\left(\underbrace{w}_{=a_1 a_2 \cdots a_k}\right) = \ell(a_1 a_2 \cdots a_k) = \ell(a_1) + \ell(a_2) + \cdots + \ell(a_k) = \sum_{i=1}^{k} \ell(a_i),$$

we obtain $\operatorname{sum}(\{[a_1], [a_2], \ldots, [a_k]\}_{\text{multiset}}) = \ell(w) = n$. Hence, $\{[a_1], [a_2], \ldots, [a_k]\}_{\text{multiset}}$ is a finite multiset M of aperiodic $\mathfrak{B}$-necklaces satisfying $\operatorname{sum} M = n$ (since we already know that $\{[a_1], [a_2], \ldots, [a_k]\}_{\text{multiset}}$

is a finite multiset of aperiodic $\mathfrak{B}$-necklaces). In other words, $\{[a_1], [a_2], \ldots, [a_k]\}_{\text{multiset}} \in \mathfrak{MN}^{\mathfrak{a}}_{\mathfrak{B},n}$ (since $\mathfrak{MN}^{\mathfrak{a}}_{\mathfrak{B},n}$ was defined as the set of all such multisets M).

Thus, we have showed that $\{[a_1], [a_2], \ldots, [a_k]\}_{\text{multiset}}$ is a well-defined element of $\mathfrak{MN}^{\mathfrak{a}}_{\mathfrak{B},n}$. This proves Proposition 13.191.9. □

Definition 13.191.10. Let $\mathfrak{B}$ be a subset of $\mathfrak{A}$. Let $n \in \mathbb{N}$. Then, we define a map $\operatorname{CFL} : \mathfrak{B}^n \to \mathfrak{MN}^{\mathfrak{a}}_{\mathfrak{B},n}$ as follows: For each $w \in \mathfrak{B}^n$, we set $\operatorname{CFL} w = \{[a_1], [a_2], \ldots, [a_k]\}_{\text{multiset}}$, where $(a_1, a_2, \ldots, a_k)$ is the CFL factorization of the word w. (This is well-defined, since Proposition 13.191.9 shows that $\{[a_1], [a_2], \ldots, [a_k]\}_{\text{multiset}}$ is a well-defined element of $\mathfrak{MN}^{\mathfrak{a}}_{\mathfrak{B},n}$.)

Example 13.191.11. If $\mathfrak{B} = \{2 < 4 < 6 < 8\}$ and $n = 9$ and $w = 262642424$, then the map CFL sends the word w to $\operatorname{CFL} w = \{[26264], [24], [24]\}_{\text{multiset}}$, since the CFL factorization of w is $(26264, 24, 24)$.

Recall that $\mathfrak{L}$ denotes the set of all Lyndon words, whereas $\mathfrak{N}^{\mathfrak{a}}$ denotes the set of all aperiodic necklaces.

Definition 13.191.12. We define a map $\operatorname{lynd} : \mathfrak{N}^{\mathfrak{a}} \to \mathfrak{L}$ as follows:

Let $N \in \mathfrak{N}^{\mathfrak{a}}$. Thus, N is an aperiodic necklace. Hence, N is a necklace. In other words, N is an n-necklace for some positive integer n (by the definition of a "necklace"). Consider this n. Hence, N is an aperiodic n-necklace. Thus, Exercise 6.1.34(c) shows that N contains exactly one Lyndon word. We define $\operatorname{lynd} N$ to be this Lyndon word. (Thus, $\operatorname{lynd} N \in \mathfrak{L}$; hence, the map $\operatorname{lynd} : \mathfrak{N}^{\mathfrak{a}} \to \mathfrak{L}$ is well-defined.)

Example 13.191.13. If $\mathfrak{A} = \{1 < 2 < 3 < 4 < \cdots\}$ and $N = [4121312]$, then $\operatorname{lynd} N = 1213124$.

Lemma 13.191.14. *Let N be an aperiodic necklace. Then:*

(a) *We have $\ell(\operatorname{lynd} N) = |N|$.*
(b) *We have $[\operatorname{lynd} N] = N$.*
(c) *If u is a Lyndon word such that $u \in N$, then $u = \operatorname{lynd} N$.*
(d) *Let $\mathfrak{B}$ be a subset of $\mathfrak{A}$. Assume that N is a $\mathfrak{B}$-necklace. Then, $\operatorname{lynd} N \in \mathfrak{B}^*$.*

Proof of Lemma 13.191.14. The definition of $\operatorname{lynd} N$ yields that $\operatorname{lynd} N$ is the unique Lyndon word that N contains. In other words, $\operatorname{lynd} N$ is the unique Lyndon word w such that $w \in N$. Hence, $\operatorname{lynd} N$ is a Lyndon word and satisfies $\operatorname{lynd} N \in N$. Hence, Lemma 13.190.2 (applied to $u = \operatorname{lynd} N$) yields $N = [\operatorname{lynd} N]$. This proves Lemma 13.191.14(b).

(a) The word $\operatorname{lynd} N$ is a Lyndon word, and thus is nonempty (since any Lyndon word is nonempty). Moreover, the necklace N is aperiodic. Since $N = [\operatorname{lynd} N]$, this rewrites as follows: The necklace $[\operatorname{lynd} N]$ is aperiodic. Hence, Lemma 13.190.4 (applied to $u = \operatorname{lynd} N$) yields $|[\operatorname{lynd} N]| = \ell(\operatorname{lynd} N)$. Since $N = [\operatorname{lynd} N]$, we now have $|N| = |[\operatorname{lynd} N]| = \ell(\operatorname{lynd} N)$. This proves Lemma 13.191.14(a).

(c) Recall that $\operatorname{lynd} N$ is the unique Lyndon word w such that $w \in N$. Thus, $\operatorname{lynd} N$ is the only such Lyndon word w. In other words, if w is a Lyndon word such that $w \in N$, then $w = \operatorname{lynd} N$. Renaming w as u in this statement, we obtain precisely the claim of Lemma 13.191.14(c).

(d) We know that N is a $\mathfrak{B}$-necklace. In other words, N is a necklace that is a subset of $\mathfrak{B}^*$ (because this is what "$\mathfrak{B}$-necklace" means). Thus, N is a subset of $\mathfrak{B}^*$. In other words, $N \subset \mathfrak{B}^*$. Thus, $\operatorname{lynd} N \in N \subset \mathfrak{B}^*$. This proves Lemma 13.191.14(d). □

Lemma 13.191.15. *Let u be a Lyndon word. Then,* $\operatorname{lynd}[u] = u$.

Proof of Lemma 13.191.15. The necklace $[u]$ is aperiodic (by Lemma 13.190.3). Hence, $\operatorname{lynd}[u]$ is well-defined. Moreover, the word u is Lyndon and thus nonempty (since every Lyndon word is nonempty). Thus, Lemma 13.190.1 yields $u \in [u]$. Hence, Lemma 13.191.14(c) (applied to $N = [u]$) yields $u = \operatorname{lynd}[u]$. This proves Lemma 13.191.15. □

The next definition introduces a notation for a simple and very basic concept: the concept of applying a map $f : X \to Y$ to a finite multiset M of elements of X (by applying f to each element of M).

Definition 13.191.16. Let X and Y be two sets, and let $f : X \to Y$ be any map. Let M be a finite multiset of elements of X. Then, $f_* M$ shall denote the finite multiset of elements of Y obtained by applying f to each element of M. More formally, this multiset $f_* M$ can be defined by the requirement that for any object y, we have

$$(\text{multiplicity of } y \text{ in } f_* M) = \sum_{\substack{x \in X; \\ f(x) = y}} (\text{multiplicity of } x \text{ in } M).$$

More explicitly, f_*M can be described as follows: If $M = \{m_1, m_2, \ldots, m_k\}_{\text{multiset}}$, then

$$f_*M = \{f(m_1), f(m_2), \ldots, f(m_k)\}_{\text{multiset}}. \tag{13.191.7}$$

Example 13.191.17. Let $X = \mathbb{Z}$ and $Y = \mathbb{N}$, and let $f : X \to Y$ be the map that sends each $x \in \mathbb{Z}$ to $x^2 \in \mathbb{N}$. Then, (13.191.7) yields $f_*(\{-2, 2, 3, 3\}_{\text{multiset}}) = \{f(-2), f(2), f(3), f(3)\}_{\text{multiset}} = \{4, 4, 9, 9\}_{\text{multiset}}$.

Example 13.191.18. Let M be the finite multiset $\{[13], [13], [4252]\}_{\text{multiset}}$ of aperiodic necklaces. Then, (13.191.7) yields

$$\operatorname{lynd}_* M = \{\operatorname{lynd}[13], \operatorname{lynd}[13], \operatorname{lynd}[4252]\}_{\text{multiset}} = \{13, 13, 2425\}_{\text{multiset}}.$$

Lemma 13.191.19. *Let M be a finite multiset of aperiodic necklaces. Let $\widetilde{M}$ be the multiset $\operatorname{lynd}_* M$. Then:*

(a) *The multiset $\widetilde{M}$ is a finite multiset of Lyndon words.*

Now, let $b_1, b_2, \ldots, b_k$ be some Lyndon words such that $\widetilde{M} = \{b_1, b_2, \ldots, b_k\}_{\text{multiset}}$. Then:

(b) *We have $M = \{[b_1], [b_2], \ldots, [b_k]\}_{\text{multiset}}$.*

(c) *We have $\operatorname{sum} M = \ell(b_1) + \ell(b_2) + \cdots + \ell(b_k)$.*

Proof of Lemma 13.191.19. We know that M is a finite multiset of aperiodic necklaces, i.e., a finite multiset of elements of $\mathfrak{N}^{\mathfrak{a}}$. Hence, $\operatorname{lynd}_* M$ is well-defined and is a finite multiset of elements of $\mathfrak{L}$ (since lynd is a map from $\mathfrak{N}^{\mathfrak{a}}$ to $\mathfrak{L}$). In other words, $\operatorname{lynd}_* M$ is well-defined and is a finite multiset of elements of Lyndon words (since the elements of $\mathfrak{L}$ are the Lyndon words). In other words, $\widetilde{M}$ is well-defined and is a finite multiset of elements of Lyndon words (because $\widetilde{M} = \operatorname{lynd}_* M$). This proves Lemma 13.191.19(a).

Write the multiset M in the form $M = \{N_1, N_2, \ldots, N_i\}_{\text{multiset}}$ for some aperiodic necklaces $N_1, N_2, \ldots, N_i$. (This can be done, since M is a finite multiset of aperiodic necklaces.)

From $M = \{N_1, N_2, \ldots, N_i\}_{\text{multiset}}$, we obtain $\operatorname{lynd}_* M = \{\operatorname{lynd} N_1, \operatorname{lynd} N_2, \ldots, \operatorname{lynd} N_i\}_{\text{multiset}}$ (by (13.191.7), applied to $\mathfrak{N}^{\mathfrak{a}}$, $\mathfrak{L}$, lynd, i and N_j instead of X, Y, f, k and m_j). Hence,

$$\widetilde{M} = \operatorname{lynd}_* M = \{\operatorname{lynd} N_1, \operatorname{lynd} N_2, \ldots, \operatorname{lynd} N_i\}_{\text{multiset}}. \tag{13.191.8}$$

(b) For each $r \in \{1, 2, \ldots, i\}$, the set N_r is an aperiodic necklace (since $N_1, N_2, \ldots, N_i$ are aperiodic necklaces) and therefore satisfies $[\operatorname{lynd} N_r] = N_r$ (by Lemma 13.191.14(b), applied to $N = N_r$). In other words,

$$([\operatorname{lynd} N_1], [\operatorname{lynd} N_2], \ldots, [\operatorname{lynd} N_i]) = (N_1, N_2, \ldots, N_i). \tag{13.191.9}$$

But comparing (13.191.8) with $\widetilde{M} = \{b_1, b_2, \ldots, b_k\}_{\text{multiset}}$, we obtain

$$\{b_1, b_2, \ldots, b_k\}_{\text{multiset}} = \{\operatorname{lynd} N_1, \operatorname{lynd} N_2, \ldots, \operatorname{lynd} N_i\}_{\text{multiset}}.$$

Thus, the k-tuple $(b_1, b_2, \ldots, b_k)$ is a permutation of the i-tuple $(\operatorname{lynd} N_1, \operatorname{lynd} N_2, \ldots, \operatorname{lynd} N_i)$. Hence,

$$\{[b_1], [b_2], \ldots, [b_k]\}_{\text{multiset}} = \{[\operatorname{lynd} N_1], [\operatorname{lynd} N_2], \ldots, [\operatorname{lynd} N_i]\}_{\text{multiset}} = \{N_1, N_2, \ldots, N_i\}_{\text{multiset}}$$

(by (13.191.9)). Comparing this with $M = \{N_1, N_2, \ldots, N_i\}_{\text{multiset}}$, we obtain $M = \{[b_1], [b_2], \ldots, [b_k]\}_{\text{multiset}}$. This proves Lemma 13.191.19(b).

(c) For each $r \in \{1, 2, \ldots, i\}$, the set N_r is an aperiodic necklace (since $N_1, N_2, \ldots, N_i$ are aperiodic necklaces) and therefore satisfies

$$\ell(\operatorname{lynd} N_r) = |N_r| \tag{13.191.10}$$

(by Lemma 13.191.14(a), applied to $N = N_r$).

In our above proof of Lemma 13.191.19(b), we have showed that the k-tuple $(b_1, b_2, \ldots, b_k)$ is a permutation of the i-tuple $(\operatorname{lynd} N_1, \operatorname{lynd} N_2, \ldots, \operatorname{lynd} N_i)$. Hence,

$$\begin{aligned}\ell(b_1) + \ell(b_2) + \cdots + \ell(b_k) &= \ell(\operatorname{lynd} N_1) + \ell(\operatorname{lynd} N_2) + \cdots + \ell(\operatorname{lynd} N_i) \\ &= \sum_{r=1}^{i} \underbrace{\ell(\operatorname{lynd} N_r)}_{\substack{=|N_r| \\ \text{(by (13.191.10))}}} = \sum_{r=1}^{i} |N_r| = |N_1| + |N_2| + \cdots + |N_i|.\end{aligned}$$

But (13.191.5) (applied to i instead of k) yields

$$\operatorname{sum} M = |N_1| + |N_2| + \cdots + |N_i| \qquad (\text{since } M = \{N_1, N_2, \ldots, N_i\}_{\text{multiset}}) .$$

Comparing these two equalities, we find $\operatorname{sum} M = \ell(b_1) + \ell(b_2) + \cdots + \ell(b_k)$. This proves Lemma 13.191.19(c). □

Proposition 13.191.20. *Let $\mathfrak{B}$ be a subset of $\mathfrak{A}$. Let $n \in \mathbb{N}$. Let $M \in \mathfrak{M}\mathfrak{N}^{\mathfrak{a}}_{\mathfrak{B},n}$. Then, M is a finite multiset of aperiodic $\mathfrak{B}$-necklaces. Let $\widetilde{M}$ be the multiset $\operatorname{lynd}_* M$. This $\widetilde{M}$ is a finite multiset of Lyndon words (by Lemma 13.191.19(a)). Let $(b_1, b_2, \ldots, b_k)$ be the $\leq$-increasing list of $\widetilde{M}$. (Here we are using the lexicographic order $\leq$, not the relation $\leq_\omega$.)*

Then, $b_k b_{k-1} \cdots b_1 \in \mathfrak{B}^n$.

Proof of Proposition 13.191.20. We have $M \in \mathfrak{M}\mathfrak{N}^{\mathfrak{a}}_{\mathfrak{B},n}$. In other words, M is a finite multiset of aperiodic $\mathfrak{B}$-necklaces satisfying $\operatorname{sum} M = n$ (since $\mathfrak{M}\mathfrak{N}^{\mathfrak{a}}_{\mathfrak{B},n}$ was defined to be the set of all such multisets). Hence, in particular, M is a finite multiset of aperiodic necklaces (since each $\mathfrak{B}$-necklace is a necklace). Thus, $\operatorname{lynd}_* M$ is well-defined.

The list $(b_1, b_2, \ldots, b_k)$ is the $\leq$-increasing list of $\widetilde{M}$ (by its definition). Hence, $\widetilde{M} = \{b_1, b_2, \ldots, b_k\}_{\text{multiset}}$. Thus, $b_1, b_2, \ldots, b_k$ are elements of $\widetilde{M}$, and therefore are Lyndon words (since $\widetilde{M}$ is a multiset of Lyndon words). Thus, Lemma 13.191.19(b) yields $M = \{[b_1], [b_2], \ldots, [b_k]\}_{\text{multiset}}$.

Now, let $i \in \{1, 2, \ldots, k\}$. Then, $[b_i]$ is an element of M (since $M = \{[b_1], [b_2], \ldots, [b_k]\}_{\text{multiset}}$). Thus, $[b_i]$ is an aperiodic $\mathfrak{B}$-necklace (since M is a finite multiset of aperiodic $\mathfrak{B}$-necklaces), hence a $\mathfrak{B}$-necklace, and thus a subset of $\mathfrak{B}^*$ (by the definition of a $\mathfrak{B}$-necklace). In other words, $[b_i] \subset \mathfrak{B}^*$. But b_i is a Lyndon word (since $b_1, b_2, \ldots, b_k$ are Lyndon words), and thus a nonempty word (since any Lyndon word is nonempty). Hence, Lemma 13.190.1 (applied to $u = b_i$) yields $b_i \in [b_i] \subset \mathfrak{B}^*$.

Forget that we fixed i. We thus have shown that $b_i \in \mathfrak{B}^*$ for each $i \in \{1, 2, \ldots, k\}$. Hence, $b_k b_{k-1} \cdots b_1 \in \mathfrak{B}^*$.

But Lemma 13.191.19(c) yields $\operatorname{sum} M = \ell(b_1) + \ell(b_2) + \cdots + \ell(b_k)$. Comparing this with $\operatorname{sum} M = n$, we obtain

$$\ell(b_1) + \ell(b_2) + \cdots + \ell(b_k) = n.$$

Now, combining $b_k b_{k-1} \cdots b_1 \in \mathfrak{B}^*$ with

$$\ell(b_k b_{k-1} \cdots b_1) = \ell(b_k) + \ell(b_{k-1}) + \cdots + \ell(b_1) = \ell(b_1) + \ell(b_2) + \cdots + \ell(b_k) = n,$$

we obtain $b_k b_{k-1} \cdots b_1 \in \mathfrak{B}^n$. This proves Proposition 13.191.20. □

Definition 13.191.21. Let $\mathfrak{B}$ be a subset of $\mathfrak{A}$. Let $n \in \mathbb{N}$. Then, we define a map $\operatorname{LFC} : \mathfrak{M}\mathfrak{N}^{\mathfrak{a}}_{\mathfrak{B},n} \to \mathfrak{B}^n$ as follows: Let $M \in \mathfrak{M}\mathfrak{N}^{\mathfrak{a}}_{\mathfrak{B},n}$. Then, M is a finite multiset of aperiodic $\mathfrak{B}$-necklaces. Let $\widetilde{M}$ be the multiset $\operatorname{lynd}_* M$. This $\widetilde{M}$ is a finite multiset of Lyndon words (by Lemma 13.191.19(a)). Let $(b_1, b_2, \ldots, b_k)$ be the $\leq$-increasing list of $\widetilde{M}$. Then, we set $\operatorname{LFC}(M) = b_k b_{k-1} \cdots b_1$. (This is well-defined, since Proposition 13.191.20 shows that $b_k b_{k-1} \cdots b_1 \in \mathfrak{B}^n$.)

Example 13.191.22. If $\mathfrak{B} = \{2 < 4 < 6 < 8\}$ and $n = 9$ and $M = \{[26264], [24], [24]\}_{\text{multiset}}$, then the multiset $\widetilde{M}$ in Definition 13.191.21 is $\{26264, 24, 24\}_{\text{multiset}}$, and its $\leq$-increasing list is $(24, 24, 26264)$, so that $\operatorname{LFC}(M) = 262642424$.

Proposition 13.191.23. *Let $\mathfrak{B}$ be a subset of $\mathfrak{A}$. Let $n \in \mathbb{N}$. Then, the maps $\operatorname{CFL} : \mathfrak{B}^n \to \mathfrak{M}\mathfrak{N}^{\mathfrak{a}}_{\mathfrak{B},n}$ and $\operatorname{LFC} : \mathfrak{M}\mathfrak{N}^{\mathfrak{a}}_{\mathfrak{B},n} \to \mathfrak{B}^n$ (introduced in Definition 13.191.10 and in Definition 13.191.21) are mutually inverse.*

Proof of Proposition 13.191.23. We shall prove two claims:

Claim 1: We have $\operatorname{CFL} \circ \operatorname{LFC} = \operatorname{id}$.

[*Proof of Claim 1:* Let $M \in \mathfrak{M}\mathfrak{N}^{\mathfrak{a}}_{\mathfrak{B},n}$. Thus, M is a finite multiset of aperiodic $\mathfrak{B}$-necklaces satisfying $\operatorname{sum} M = n$ (since $\mathfrak{M}\mathfrak{N}^{\mathfrak{a}}_{\mathfrak{B},n}$ was defined to be the set of all such multisets). Hence, M is a finite multiset of elements of $\mathfrak{N}^{\mathfrak{a}}$. Thus, $\operatorname{lynd}_* M$ is well-defined.

Let $\widetilde{M}$ be the multiset $\operatorname{lynd}_* M$. This $\widetilde{M}$ is a finite multiset of Lyndon words (by Lemma 13.191.19(a)). Let $(b_1, b_2, \ldots, b_k)$ be the $\leq$-increasing list of $\widetilde{M}$. According to the definition of a $\leq$-increasing list, this means that $\widetilde{M} = \{b_1, b_2, \ldots, b_k\}_{\text{multiset}}$ and $b_1 \leq b_2 \leq \cdots \leq b_k$. Now, $b_1, b_2, \ldots, b_k$ are Lyndon words (since

$\{b_1, b_2, \ldots, b_k\}_{\text{multiset}} = \widetilde{M}$ is a multiset of Lyndon words). Hence, $(b_k, b_{k-1}, \ldots, b_1)$ is a tuple of Lyndon words.

Define $w \in \mathfrak{B}^*$ by $w = \operatorname{LFC}(M)$. Thus, $w = \operatorname{LFC}(M) = b_k b_{k-1} \cdots b_1$ (by the definition of the map LFC).

We shall now show that $\operatorname{CFL} w = M$.

Indeed, $b_1 \leq b_2 \leq \cdots \leq b_k$. In other words, $b_k \geq b_{k-1} \geq \cdots \geq b_1$. Combining this with the fact that $(b_k, b_{k-1}, \ldots, b_1)$ is a tuple of Lyndon words, we conclude that $(b_k, b_{k-1}, \ldots, b_1)$ is the CFL factorization of w (by the definition of the CFL factorization).

On the other hand, M is a finite multiset of aperiodic necklaces. Moreover, $b_1, b_2, \ldots, b_k$ are Lyndon words and satisfy $\widetilde{M} = \{b_1, b_2, \ldots, b_k\}_{\text{multiset}}$. Hence, Lemma 13.191.19(b) yields $M = \{[b_1], [b_2], \ldots, [b_k]\}_{\text{multiset}}$.

But the definition of the map CFL yields that if $(a_1, a_2, \ldots, a_k)$ is the CFL factorization of the word w, then

$$\operatorname{CFL} w = \{[a_1], [a_2], \ldots, [a_k]\}_{\text{multiset}}.$$

Applying this to $(a_1, a_2, \ldots, a_k) = (b_k, b_{k-1}, \ldots, b_1)$, we obtain

$$\begin{aligned} \operatorname{CFL} w &= \{[b_k], [b_{k-1}], \ldots, [b_1]\}_{\text{multiset}} && (\text{since } (b_k, b_{k-1}, \ldots, b_1) \text{ is the CFL factorization of } w) \\ &= \{[b_1], [b_2], \ldots, [b_k]\}_{\text{multiset}} = M && (\text{since } M = \{[b_1], [b_2], \ldots, [b_k]\}_{\text{multiset}}). \end{aligned}$$

In view of $w = \operatorname{LFC}(M)$, this rewrites as $\operatorname{CFL}(\operatorname{LFC}(M)) = M$. Thus, $(\operatorname{CFL} \circ \operatorname{LFC})(M) = \operatorname{CFL}(\operatorname{LFC}(M)) = M = \operatorname{id}(M)$.

Forget that we fixed M. We thus have showed that $(\operatorname{CFL} \circ \operatorname{LFC})(M) = \operatorname{id}(M)$ for each $M \in \mathfrak{MN}^{\mathfrak{a}}_{\mathfrak{B},n}$. In other words, $\operatorname{CFL} \circ \operatorname{LFC} = \operatorname{id}$. Thus, Claim 1 is proven.]

Claim 2: We have $\operatorname{LFC} \circ \operatorname{CFL} = \operatorname{id}$.

[*Proof of Claim 2:* Let $w \in \mathfrak{B}^n$. Let $(a_1, a_2, \ldots, a_k)$ be the CFL factorization of the word w. Then, the definition of CFL yields $\operatorname{CFL} w = \{[a_1], [a_2], \ldots, [a_k]\}_{\text{multiset}}$.

We know that $(a_1, a_2, \ldots, a_k)$ is a CFL factorization of w. In other words, $(a_1, a_2, \ldots, a_k)$ is a tuple of Lyndon words satisfying $w = a_1 a_2 \cdots a_k$ and $a_1 \geq a_2 \geq \cdots \geq a_k$ (because this is how a "CFL factorization of w" is defined). In particular, $a_1, a_2, \ldots, a_k$ are Lyndon words (since $(a_1, a_2, \ldots, a_k)$ is a tuple of Lyndon words).

Hence, for each $i \in \{1, 2, \ldots, k\}$, the word a_i is a Lyndon word, and therefore satisfies $\operatorname{lynd}[a_i] = a_i$ (by Lemma 13.191.15, applied to $u = a_i$). In other words,

$$(\operatorname{lynd}[a_1], \operatorname{lynd}[a_2], \ldots, \operatorname{lynd}[a_k]) = (a_1, a_2, \ldots, a_k). \tag{13.191.11}$$

Let M denote the multiset $\operatorname{CFL} w \in \mathfrak{MN}^{\mathfrak{a}}_{\mathfrak{B},n}$.

Thus, $M \in \mathfrak{MN}^{\mathfrak{a}}_{\mathfrak{B},n}$. In other words, M is a finite multiset of aperiodic $\mathfrak{B}$-necklaces satisfying $\operatorname{sum} M = n$ (since $\mathfrak{MN}^{\mathfrak{a}}_{\mathfrak{B},n}$ was defined to be the set of all such multisets). Hence, M is a finite multiset of elements of $\mathfrak{N}^{\mathfrak{a}}$. Thus, $\operatorname{lynd}_* M$ is well-defined.

The definition of M yields $M = \operatorname{CFL} w = \{[a_1], [a_2], \ldots, [a_k]\}_{\text{multiset}}$. Hence,

$$\operatorname{lynd}_* M = \{\operatorname{lynd}[a_1], \operatorname{lynd}[a_2], \ldots, \operatorname{lynd}[a_k]\}_{\text{multiset}}$$

(by (13.191.7), applied to $\mathfrak{N}^{\mathfrak{a}}$, $\mathfrak{L}$, lynd and $[a_j]$ instead of X, Y, f and m_j).

Let $\widetilde{M}$ be the multiset $\operatorname{lynd}_* M$. This $\widetilde{M}$ is a finite multiset of Lyndon words (by Lemma 13.191.19(a)), and satisfies

$$\begin{aligned} \widetilde{M} &= \operatorname{lynd}_* M = \{\operatorname{lynd}[a_1], \operatorname{lynd}[a_2], \ldots, \operatorname{lynd}[a_k]\}_{\text{multiset}} \\ &= \{a_1, a_2, \ldots, a_k\}_{\text{multiset}} && (\text{by } (13.191.11)) \\ &= \{a_k, a_{k-1}, \ldots, a_1\}_{\text{multiset}}. \end{aligned}$$

Also, recall that $a_1 \geq a_2 \geq \cdots \geq a_k$. In other words, $a_k \leq a_{k-1} \leq \cdots \leq a_1$.

Now, $(a_k, a_{k-1}, \ldots, a_1)$ is the $\leq$-increasing list of $\widetilde{M}$ (because $\widetilde{M} = \{a_k, a_{k-1}, \ldots, a_1\}_{\text{multiset}}$ and $a_k \leq a_{k-1} \leq \cdots \leq a_1$). But the definition of LFC shows that if $(b_1, b_2, \ldots, b_k)$ is the $\leq$-increasing list of $\widetilde{M}$, then

$$\operatorname{LFC}(M) = b_k b_{k-1} \cdots b_1.$$

We can apply this to $(b_1, b_2, \ldots, b_k) = (a_k, a_{k-1}, \ldots, a_1)$ (since $(a_k, a_{k-1}, \ldots, a_1)$ is the $\leq$-increasing list of $\widetilde{M}$); thus, we obtain $\operatorname{LFC}(M) = a_1 a_2 \cdots a_k$. Comparing this with $w = a_1 a_2 \cdots a_k$, we obtain $\operatorname{LFC}(M) = w$.

In view of $M = \operatorname{CFL} w$, this rewrites as $\operatorname{LFC}(\operatorname{CFL} w) = w$. Hence,

$$(\operatorname{LFC} \circ \operatorname{CFL})(w) = \operatorname{LFC}(\operatorname{CFL} w) = w = \operatorname{id}(w).$$

Forget that we fixed w. We thus have shown that $(\operatorname{LFC} \circ \operatorname{CFL})(w) = \operatorname{id}(w)$ for each $w \in \mathfrak{B}^n$. In other words, $\operatorname{LFC} \circ \operatorname{CFL} = \operatorname{id}$. This proves Claim 2.]

We have $\operatorname{CFL} \circ \operatorname{LFC} = \operatorname{id}$ (by Claim 1) and $\operatorname{LFC} \circ \operatorname{CFL} = \operatorname{id}$ (by Claim 2). Combining these two equalities, we conclude that the maps $\operatorname{CFL} : \mathfrak{B}^n \to \mathfrak{MN}^{\mathfrak{a}}_{\mathfrak{B},n}$ and $\operatorname{LFC} : \mathfrak{MN}^{\mathfrak{a}}_{\mathfrak{B},n} \to \mathfrak{B}^n$ are mutually inverse. This proves Proposition 13.191.23. □

Corollary 13.191.24. *Let $\mathfrak{B}$ be a finite subset of $\mathfrak{A}$. Let $n \in \mathbb{N}$. Then:*

(a) *The sets $\mathfrak{B}^n$ and $\mathfrak{MN}^{\mathfrak{a}}_{\mathfrak{B},n}$ are finite and satisfy $|\mathfrak{B}^n| = \left|\mathfrak{MN}^{\mathfrak{a}}_{\mathfrak{B},n}\right|$.*

(b) *Any injective map $f : \mathfrak{B}^n \to \mathfrak{MN}^{\mathfrak{a}}_{\mathfrak{B},n}$ is bijective.*

Proof of Corollary 13.191.24. The set $\mathfrak{B}$ is finite. Hence, the set $\mathfrak{B}^n$ is finite as well.

Proposition 13.191.23 shows that the maps $\operatorname{CFL} : \mathfrak{B}^n \to \mathfrak{MN}^{\mathfrak{a}}_{\mathfrak{B},n}$ and $\operatorname{LFC} : \mathfrak{MN}^{\mathfrak{a}}_{\mathfrak{B},n} \to \mathfrak{B}^n$ (defined as in Proposition 13.191.23) are mutually inverse. Thus, these two maps are bijections. Hence, there exists a bijection from $\mathfrak{B}^n$ to $\mathfrak{MN}^{\mathfrak{a}}_{\mathfrak{B},n}$. This entails $|\mathfrak{B}^n| = \left|\mathfrak{MN}^{\mathfrak{a}}_{\mathfrak{B},n}\right|$. Since $\mathfrak{B}^n$ is finite, this shows that $\mathfrak{MN}^{\mathfrak{a}}_{\mathfrak{B},n}$ is finite as well. Thus, Corollary 13.191.24(a) is proven.

(b) Let $f : \mathfrak{B}^n \to \mathfrak{MN}^{\mathfrak{a}}_{\mathfrak{B},n}$ be any injective map. We must prove that f is bijective.

Corollary 13.191.24(a) shows that $\mathfrak{B}^n$ and $\mathfrak{MN}^{\mathfrak{a}}_{\mathfrak{B},n}$ are two finite sets of equal sizes. Hence, f is a map between two finite sets of equal sizes. But it is well-known that any injective map between two finite sets of equal sizes is bijective[1202]. Thus, f is bijective. This proves Corollary 13.191.24(b).

□

Now, at last we can prove the following:

Lemma 13.191.25. *We have* $\operatorname{GR} \circ \operatorname{RG} = \operatorname{id}$.

Proof of Lemma 13.191.25. Lemma 13.191.2 yields $\operatorname{RG} \circ \operatorname{GR} = \operatorname{id}$. Thus, the map GR has a left inverse (namely, RG). Therefore, the map GR is injective.

Let $M \in \mathfrak{MN}^{\mathfrak{a}}$. We shall show that $(\operatorname{GR} \circ \operatorname{RG})(M) = M$.

We have $M \in \mathfrak{MN}^{\mathfrak{a}}$. In other words, M is a finite multiset of aperiodic necklaces (by the definition of $\mathfrak{MN}^{\mathfrak{a}}$). Thus, each element of M is an aperiodic necklace.

Let $n = \operatorname{sum} M$; thus, $n \in \mathbb{N}$.

Define a subset $\mathfrak{B}$ of $\mathfrak{A}$ by

$$\mathfrak{B} = \bigcup_{N \in M} \bigcup_{w \in N} \{\text{all letters of } w\} \tag{13.191.12}$$

(where the "$\bigcup_{N \in M}$" symbol should be understood as "$\bigcup_{N \in \operatorname{Supp} M}$").

Thus, $\mathfrak{B}$ is a finite union of finite unions of finite sets[1203]. Thus, the set $\mathfrak{B}$ is finite. Furthermore, the definition of $\mathfrak{B}$ ensures that each $N \in M$ is an aperiodic $\mathfrak{B}$-necklace[1204]. Thus, M is a finite multiset of aperiodic $\mathfrak{B}$-necklaces. Hence, M is a finite multiset of aperiodic $\mathfrak{B}$-necklaces satisfying $\operatorname{sum} M = n$ (since $n = \operatorname{sum} M$). In other words, $M \in \mathfrak{MN}^{\mathfrak{a}}_{\mathfrak{B},n}$ (by the definition of $\mathfrak{MN}^{\mathfrak{a}}_{\mathfrak{B},n}$).

But Lemma 13.191.7 yields that $\operatorname{GR}(\mathfrak{B}^n) \subset \mathfrak{MN}^{\mathfrak{a}}_{\mathfrak{B},n}$. In other words, $\operatorname{GR}(w) \in \mathfrak{MN}^{\mathfrak{a}}_{\mathfrak{B},n}$ for each $w \in \mathfrak{B}^n$. Thus, we can define a map

$$\begin{aligned} \overline{\operatorname{GR}} : \mathfrak{B}^n &\to \mathfrak{MN}^{\mathfrak{a}}_{\mathfrak{B},n}, \\ w &\mapsto \operatorname{GR}(w). \end{aligned}$$

[1202] This is one of the basic facts known as the "pigeonhole principle".

[1203] To wit:

- There are only finitely many $N \in M$ (since M is finite).
- For each $N \in M$, there are only finitely many $w \in N$ (since N is a necklace, thus a finite set).
- For each $N \in M$ and each $w \in N$, the set $\{\text{all letters of } w\}$ is finite (since the word w has only finitely many letters).

[1204] *Proof.* Let $N \in M$. Then, N is an aperiodic necklace (since each element of M is an aperiodic necklace).

But (13.191.12) shows that $\bigcup_{w \in N} \{\text{all letters of } w\} \subset \mathfrak{B}$ (since $N \in M$). In other words, each $w \in N$ satisfies $\{\text{all letters of } w\} \subset \mathfrak{B}$. In other words, each $w \in N$ satisfies $w \in \mathfrak{B}^*$. In other words, $N \subset \mathfrak{B}^*$. Hence, N is a $\mathfrak{B}$-necklace (since N is a necklace), hence an aperiodic $\mathfrak{B}$-necklace (since N is aperiodic). Qed.

Consider this map $\overline{\mathrm{GR}}$. Clearly, this map $\overline{\mathrm{GR}}$ is a restriction of the map GR; hence, this map $\overline{\mathrm{GR}}$ is injective (since the map GR is injective). Hence, Corollary 13.191.24(b) (applied to $f = \overline{\mathrm{GR}}$) yields that the map $\overline{\mathrm{GR}}$ is bijective. Thus, $\mathfrak{MN}^{\mathfrak{a}}_{\mathfrak{B},n} = \overline{\mathrm{GR}}\left(\mathfrak{B}^n\right)$.

Now, $M \in \mathfrak{MN}^{\mathfrak{a}}_{\mathfrak{B},n} = \overline{\mathrm{GR}}\left(\mathfrak{B}^n\right)$. In other words, there exists some $w \in \mathfrak{B}^n$ satisfying $M = \overline{\mathrm{GR}}\ w$. Consider this w. We have $M = \overline{\mathrm{GR}}\ w = \mathrm{GR}\, w$ (by the definition of $\overline{\mathrm{GR}}$). Now,

$$(\mathrm{GR} \circ \mathrm{RG})\left(\underbrace{M}_{=\mathrm{GR}\, w}\right) = (\mathrm{GR} \circ \mathrm{RG})(\mathrm{GR}\, w) = \left(\mathrm{GR} \circ \underbrace{\mathrm{RG} \circ \mathrm{GR}}_{=\mathrm{id}}\right)(w) = \mathrm{GR}\, w = M = \mathrm{id}(M).$$

Forget that we fixed M. We thus have proved that $(\mathrm{GR} \circ \mathrm{RG})(M) = \mathrm{id}(M)$ for each $M \in \mathfrak{MN}^{\mathfrak{a}}$. In other words, $\mathrm{GR} \circ \mathrm{RG} = \mathrm{id}$. This proves Lemma 13.191.25. □

Proof of Theorem 6.6.29. We have $\mathrm{GR} \circ \mathrm{RG} = \mathrm{id}$ (by Lemma 13.191.25) and $\mathrm{RG} \circ \mathrm{GR} = \mathrm{id}$ (by Lemma 13.191.2). These two equalities combined yield the claim of Theorem 6.6.29. □

Hence, Exercise 6.6.30 is solved.

13.192. **Solution to Exercise 6.6.51.**

Solution to Exercise 6.6.51. In order to solve Exercise 6.6.51, we must prove the statements made in Subsection 6.6.2. We shall begin with Proposition 6.6.39. First, let us show a few simple lemmas about words $w \in \mathfrak{A}^*$ and the respective monomials $\mathbf{x}_w$:

Lemma 13.192.1. *Let u and v be two words in $\mathfrak{A}^*$. Then, $\mathbf{x}_{uv} = \mathbf{x}_u \mathbf{x}_v$.*

Proof of Lemma 13.192.1. Write the word u in the form $u = (u_1, u_2, \ldots, u_p)$. Then, the definition of $\mathbf{x}_u$ yields $\mathbf{x}_u = x_{u_1} x_{u_2} \cdots x_{u_p}$.

Write the word v in the form $v = (v_1, v_2, \ldots, v_q)$. Then, the definition of $\mathbf{x}_v$ yields $\mathbf{x}_v = x_{v_1} x_{v_2} \cdots x_{v_q}$.

From $u = (u_1, u_2, \ldots, u_p)$ and $v = (v_1, v_2, \ldots, v_q)$, we obtain

$$uv = (u_1, u_2, \ldots, u_p)(v_1, v_2, \ldots, v_q) = (u_1, u_2, \ldots, u_p, v_1, v_2, \ldots, v_q).$$

Hence, the definition of $\mathbf{x}_{uv}$ yields

$$\mathbf{x}_{uv} = x_{u_1} x_{u_2} \cdots x_{u_p} x_{v_1} x_{v_2} \cdots x_{v_q}.$$

Comparing this with

$$\underbrace{\mathbf{x}_u}_{=x_{u_1} x_{u_2} \cdots x_{u_p}} \underbrace{\mathbf{x}_v}_{=x_{v_1} x_{v_2} \cdots x_{v_q}} = x_{u_1} x_{u_2} \cdots x_{u_p} x_{v_1} x_{v_2} \cdots x_{v_q},$$

we obtain $\mathbf{x}_{uv} = \mathbf{x}_u \mathbf{x}_v$. This proves Lemma 13.192.1. □

Lemma 13.192.2. *Let $k \in \mathbb{N}$. Let $a_1, a_2, \ldots, a_k$ be words in $\mathfrak{A}^*$. Then, $\mathbf{x}_{a_1 a_2 \cdots a_k} = \mathbf{x}_{a_1} \mathbf{x}_{a_2} \cdots \mathbf{x}_{a_k}$.*

Proof of Lemma 13.192.2. Lemma 13.192.2 follows by induction on k, using Lemma 13.192.1 in the induction step. □

Lemma 13.192.3. *Let $n \in \mathbb{N}$. Then, $\sum_{w \in \mathfrak{A}^n} \mathbf{x}_w = p_1^n$.*

Proof of Lemma 13.192.3. The definition of p_1 yields

$$p_1 = x_1^1 + x_2^1 + x_3^1 + \cdots = \sum_{a \in \{1,2,3,\ldots\}} \underbrace{x_a^1}_{=x_a} = \sum_{a \in \{1,2,3,\ldots\}} x_a = \sum_{a \in \mathfrak{A}} x_a$$

(since $\{1, 2, 3, \ldots\} = \mathfrak{A}$). Taking both sides of this equality to the n-th power, we obtain

$$\begin{aligned} p_1^n &= \left(\sum_{a \in \mathfrak{A}} x_a\right)^n \\ &= \sum_{(a_1, a_2, \ldots, a_n) \in \mathfrak{A}^n} x_{a_1} x_{a_2} \cdots x_{a_n} \qquad \text{(by the product rule)} \\ &= \sum_{(w_1, w_2, \ldots, w_n) \in \mathfrak{A}^n} x_{w_1} x_{w_2} \cdots x_{w_n} \end{aligned} \tag{13.192.1}$$

(here, we have renamed the summation index $(a_1, a_2, \ldots, a_n)$ as $(w_1, w_2, \ldots, w_n)$).

Each $w \in \mathfrak{A}^n$ can be written uniquely in the form $w = (w_1, w_2, \ldots, w_n)$. Hence, we can substitute $(w_1, w_2, \ldots, w_n)$ for w in the sum $\sum_{w \in \mathfrak{A}^n} \mathbf{x}_w$. We thus obtain

$$\sum_{w \in \mathfrak{A}^n} \mathbf{x}_w = \sum_{(w_1, w_2, \ldots, w_n) \in \mathfrak{A}^n} \underbrace{\mathbf{x}_{(w_1, w_2, \ldots, w_n)}}_{\substack{= x_{w_1} x_{w_2} \cdots x_{w_n} \\ \text{(by the definition of } \mathbf{x}_{(w_1, w_2, \ldots, w_n)}\text{)}}} = \sum_{(w_1, w_2, \ldots, w_n) \in \mathfrak{A}^n} x_{w_1} x_{w_2} \cdots x_{w_n}.$$

Comparing this with (13.192.1), we obtain $\sum_{w \in \mathfrak{A}^n} \mathbf{x}_w = p_1^n$. This proves Lemma 13.192.3. □

Lemma 13.192.4. *We have*

$$\sum_{w \in \mathfrak{A}^*} \mathbf{x}_w t^{\ell(w)} = \frac{1}{1 - p_1 t} \qquad \text{in the ring } (\mathbf{k}[[\mathbf{x}]])[[t]].$$

Proof of Lemma 13.192.4. We have

$$\underbrace{\sum_{w \in \mathfrak{A}^*}}_{\substack{= \sum_{n \in \mathbb{N}} \sum_{w \in \mathfrak{A}^n} \\ \text{(since } \mathfrak{A}^* = \bigsqcup_{n \in \mathbb{N}} \mathfrak{A}^n\text{)}}} \mathbf{x}_w t^{\ell(w)} = \sum_{n \in \mathbb{N}} \sum_{w \in \mathfrak{A}^n} \mathbf{x}_w \underbrace{t^{\ell(w)}}_{\substack{= t^n \\ \text{(since } \ell(w) = n \\ \text{(since } w \in \mathfrak{A}^n\text{))}}} = \sum_{n \in \mathbb{N}} \underbrace{\sum_{w \in \mathfrak{A}^n} \mathbf{x}_w}_{\substack{= p_1^n \\ \text{(by Lemma 13.192.3)}}} t^n = \sum_{n \in \mathbb{N}} \underbrace{p_1^n t^n}_{= (p_1 t)^n}$$

$$= \sum_{n \in \mathbb{N}} (p_1 t)^n = \frac{1}{1 - p_1 t}.$$

This proves Lemma 13.192.4. □

Lemma 13.192.5. *We have*

$$\prod_{w \in \mathfrak{L}} \frac{1}{1 - \mathbf{x}_w t^{\ell(w)}} = \sum_{w \in \mathfrak{A}^*} \mathbf{x}_w t^{\ell(w)} \qquad \text{in the ring } (\mathbf{k}[[\mathbf{x}]])[[t]].$$

Proof of Lemma 13.192.5. Let $\mathfrak{M}$ denote the set of all finite multisets of Lyndon words. Define two maps $\mathbf{m} : \mathfrak{M} \to \mathfrak{A}^*$ and $\mathbf{n} : \mathfrak{A}^* \to \mathfrak{M}$ as in Proposition 13.143.1. Then, Proposition 13.143.1 shows that the maps $\mathbf{m}$ and $\mathbf{n}$ are mutually inverse bijections.

Recall that $\mathfrak{L}$ is the set of all Lyndon words. Thus, the Lyndon words are precisely the elements of $\mathfrak{L}$. But the definition of $\mathfrak{M}$ says that $\mathfrak{M}$ is the set of all finite multisets of Lyndon words. In other words, $\mathfrak{M}$ is the set of all finite multisets of elements of $\mathfrak{L}$ (since the Lyndon words are precisely the elements of $\mathfrak{L}$).

Let $\mathfrak{N}$ be the set of all families $(k_w)_{w \in \mathfrak{L}} \in \mathbb{N}^{\mathfrak{L}}$ of nonnegative integers (indexed by the Lyndon words) such that all but finitely many $w \in \mathfrak{L}$ satisfy $k_w = 0$. Thus,

$$\sum_{\substack{(k_w)_{w \in \mathfrak{L}} \in \mathbb{N}^{\mathfrak{L}}; \\ \text{all but finitely many } w \in \mathfrak{L} \\ \text{satisfy } k_w = 0}} = \sum_{(k_w)_{w \in \mathfrak{L}} \in \mathfrak{N}} \tag{13.192.2}$$

(an equality of summation signs). Proposition 13.143.2 (applied to $S = \mathfrak{L}$) shows that the map mult : $\mathfrak{M} \to \mathfrak{N}$ that sends each multiset $M \in \mathfrak{M}$ to the family

$$((\text{multiplicity of } w \text{ in the multiset } M))_{w \in \mathfrak{L}} \in \mathfrak{N}$$

is well-defined and is a bijection. Consider this map mult. Since mult : $\mathfrak{M} \to \mathfrak{N}$ is a bijection, its inverse $\text{mult}^{-1} : \mathfrak{N} \to \mathfrak{M}$ is well-defined and also a bijection.

For any $w \in \mathfrak{L}$, the monomial $t^{\ell(w)}$ has positive degree[1205]. Hence, for any $w \in \mathfrak{L}$, the formal power series $\frac{1}{1 - \mathbf{x}_w t^{\ell(w)}}$ is well-defined and satisfies the equality

$$\frac{1}{1 - \mathbf{x}_w t^{\ell(w)}} = \sum_{k \in \mathbb{N}} \left(\mathbf{x}_w t^{\ell(w)}\right)^k$$

[1205] *Proof.* Let $w \in \mathfrak{L}$. Thus, w is a Lyndon word (since $\mathfrak{L}$ is the set of all Lyndon words), therefore a nonempty word (since any Lyndon word is nonempty by definition). Hence, $\ell(w) > 0$. Therefore, the monomial $t^{\ell(w)}$ has positive degree. Qed.

(since $\dfrac{1}{1-q} = \sum_{k\in\mathbb{N}} q^k$ in the ring $\mathbf{k}[[q]]$ of formal power series).

Multiplying these equalities over all $w \in \mathfrak{L}$, we obtain

$$\begin{aligned}\prod_{w\in\mathfrak{L}} \frac{1}{1-\mathbf{x}_w t^{\ell(w)}} &= \prod_{w\in\mathfrak{L}} \sum_{k\in\mathbb{N}} \left(\mathbf{x}_w t^{\ell(w)}\right)^k \\ &= \sum_{\substack{(k_w)_{w\in\mathfrak{L}}\in\mathbb{N}^{\mathfrak{L}};\\ \text{all but finitely many } w\in\mathfrak{L}\\ \text{satisfy } k_w=0}} \prod_{w\in\mathfrak{L}} \left(\mathbf{x}_w t^{\ell(w)}\right)^{k_w} \qquad \text{(by the product rule)} \\ &= \sum_{(k_w)_{w\in\mathfrak{L}}\in\mathfrak{N}} \prod_{w\in\mathfrak{L}} \left(\mathbf{x}_w t^{\ell(w)}\right)^{k_w} \qquad \text{(by (13.192.2))}. \end{aligned} \tag{13.192.3}$$

The composition $\mathbf{m}\circ\operatorname{mult}^{-1} : \mathfrak{N} \to \mathfrak{A}^*$ of the bijections $\mathbf{m}$ and mult^{-1} is clearly a bijection. We now claim the following:

Claim 1: Let $(k_w)_{w\in\mathfrak{L}} \in \mathfrak{N}$. Then,

$$\mathbf{x}_{\left(\mathbf{m}\circ\operatorname{mult}^{-1}\right)\left((k_w)_{w\in\mathfrak{L}}\right)} t^{\ell\left(\left(\mathbf{m}\circ\operatorname{mult}^{-1}\right)\left((k_w)_{w\in\mathfrak{L}}\right)\right)} = \prod_{w\in\mathfrak{L}} \left(\mathbf{x}_w t^{\ell(w)}\right)^{k_w}.$$

[*Proof of Claim 1:* Let $M = \operatorname{mult}^{-1}\left((k_w)_{w\in\mathfrak{L}}\right)$. Then,

$$(k_w)_{w\in\mathfrak{L}} = \operatorname{mult} M = ((\text{multiplicity of } w \text{ in the multiset } M))_{w\in\mathfrak{L}}$$

(by the definition of the map mult). In other words, every $w \in \mathfrak{L}$ satisfies

$$k_w = (\text{multiplicity of } w \text{ in the multiset } M). \tag{13.192.4}$$

Moreover, $M = \operatorname{mult}^{-1}\left((k_w)_{w\in\mathfrak{L}}\right) \in \mathfrak{M}$. In other words, M is a finite multiset of elements of $\mathfrak{L}$ (since $\mathfrak{M}$ is the set of all finite multisets of elements of $\mathfrak{L}$). Let $a_1, a_2, \ldots, a_k$ denote the elements of this multiset M listed in decreasing order. Then, the definition of $\mathbf{m}$ yields $\mathbf{m}(M) = a_1 a_2 \cdots a_k$. Also, $a_1, a_2, \ldots, a_k$ are elements of M (by the definition of $a_1, a_2, \ldots, a_k$), and thus belong to $\mathfrak{L}$ (since M is a finite multiset of elements of $\mathfrak{L}$). In other words, every a_i belongs to $\mathfrak{L}$.

We have $M = \{a_1, a_2, \ldots, a_k\}_{\text{multiset}}$ (since $a_1, a_2, \ldots, a_k$ are the elements of the multiset M listed in decreasing order). Thus, each $w \in \mathfrak{L}$ satisfies

$$\begin{aligned}&\left(\text{multiplicity of } w \text{ in the multiset } \underbrace{M}_{=\{a_1,a_2,\ldots,a_k\}_{\text{multiset}}}\right) \\ &= \left(\text{multiplicity of } w \text{ in the multiset } \{a_1, a_2, \ldots, a_k\}_{\text{multiset}}\right) \\ &= (\text{the number of } i \in \{1, 2, \ldots, k\} \text{ satisfying } a_i = w).\end{aligned}$$

Hence, each $w \in \mathfrak{L}$ satisfies

$$\begin{aligned}&(\text{the number of } i \in \{1, 2, \ldots, k\} \text{ satisfying } a_i = w) \\ &= (\text{multiplicity of } w \text{ in the multiset } M) \\ &= k_w \qquad \text{(by (13.192.4))}.\end{aligned} \tag{13.192.5}$$

Now,

$$\begin{aligned}\ell\left(\underbrace{\mathbf{m}(M)}_{=a_1a_2\cdots a_k}\right) &= \ell(a_1a_2\cdots a_k) = \ell(a_1)+\ell(a_2)+\cdots+\ell(a_k)\\ &= \underbrace{\sum_{i\in\{1,2,\ldots,k\}}}_{\substack{=\sum_{w\in\mathfrak{L}}\sum_{i\in\{1,2,\ldots,k\};\ a_i=w}\\ \text{(since every } a_i \text{ belongs to } \mathfrak{L})}} \ell(a_i) = \sum_{w\in\mathfrak{L}}\sum_{\substack{i\in\{1,2,\ldots,k\};\\ a_i=w}} \ell\left(\underbrace{a_i}_{\substack{=w\\ (\text{since } a_i=w)}}\right)\\ &= \sum_{w\in\mathfrak{L}} \underbrace{\sum_{\substack{i\in\{1,2,\ldots,k\};\\ a_i=w}} \ell(w)}_{=(\text{the number of } i\in\{1,2,\ldots,k\} \text{ satisfying } a_i=w)\cdot\ell(w)}\\ &= \sum_{w\in\mathfrak{L}} \underbrace{(\text{the number of } i\in\{1,2,\ldots,k\} \text{ satisfying } a_i=w)}_{\substack{=k_w\\ (\text{by (13.192.5)})}}\cdot\ell(w)\\ &= \sum_{w\in\mathfrak{L}} k_w\cdot\ell(w).\end{aligned}$$

Hence,

$$t^{\ell(\mathbf{m}(M))} = t^{\sum_{w\in\mathfrak{L}} k_w\cdot\ell(w)} = \prod_{w\in\mathfrak{L}} \underbrace{t^{k_w\cdot\ell(w)}}_{=\left(t^{\ell(w)}\right)^{k_w}} = \prod_{w\in\mathfrak{L}} \left(t^{\ell(w)}\right)^{k_w}. \tag{13.192.6}$$

Moreover, from $\mathbf{m}(M) = a_1a_2\cdots a_k$, we obtain

$$\begin{aligned}\mathbf{x}_{\mathbf{m}(M)} &= \mathbf{x}_{a_1a_2\cdots a_k} = \mathbf{x}_{a_1}\mathbf{x}_{a_2}\cdots\mathbf{x}_{a_k} \qquad \text{(by Lemma 13.192.2)}\\ &= \underbrace{\prod_{i\in\{1,2,\ldots,k\}}}_{\substack{=\prod_{w\in\mathfrak{L}}\prod_{i\in\{1,2,\ldots,k\};\ a_i=w}\\ \text{(since every } a_i \text{ belongs to } \mathfrak{L})}} \mathbf{x}_{a_i} = \prod_{w\in\mathfrak{L}}\prod_{\substack{i\in\{1,2,\ldots,k\};\\ a_i=w}} \underbrace{\mathbf{x}_{a_i}}_{\substack{=\mathbf{x}_w\\ (\text{since } a_i=w)}}\\ &= \prod_{w\in\mathfrak{L}} \underbrace{\prod_{\substack{i\in\{1,2,\ldots,k\};\\ a_i=w}} \mathbf{x}_w}_{\substack{=\mathbf{x}_w^{(\text{the number of } i\in\{1,2,\ldots,k\} \text{ satisfying } a_i=w)}=\mathbf{x}_w^{k_w}\\ (\text{since (the number of } i\in\{1,2,\ldots,k\} \text{ satisfying } a_i=w)=k_w\\ (\text{by (13.192.5)})))}} = \prod_{w\in\mathfrak{L}} \mathbf{x}_w^{k_w}.\end{aligned}$$

Multiplying this equality by (13.192.6), we obtain

$$\mathbf{x}_{\mathbf{m}(M)} t^{\ell(\mathbf{m}(M))} = \left(\prod_{w\in\mathfrak{L}} \mathbf{x}_w^{k_w}\right)\prod_{w\in\mathfrak{L}}\left(t^{\ell(w)}\right)^{k_w} = \prod_{w\in\mathfrak{L}} \underbrace{\left(\mathbf{x}_w^{k_w}\left(t^{\ell(w)}\right)^{k_w}\right)}_{=\left(\mathbf{x}_w t^{\ell(w)}\right)^{k_w}} = \prod_{w\in\mathfrak{L}}\left(\mathbf{x}_w t^{\ell(w)}\right)^{k_w}.$$

In view of

$$\mathbf{m}\left(\underbrace{M}_{=\mathrm{mult}^{-1}\left((k_w)_{w\in\mathfrak{L}}\right)}\right) = \mathbf{m}\left(\mathrm{mult}^{-1}\left((k_w)_{w\in\mathfrak{L}}\right)\right) = \left(\mathbf{m}\circ\mathrm{mult}^{-1}\right)\left((k_w)_{w\in\mathfrak{L}}\right),$$

this rewrites as

$$\mathbf{x}_{\left(\mathbf{m}\circ\operatorname{mult}^{-1}\right)\left((k_w)_{w\in\mathfrak{L}}\right)} t^{\ell\left(\left(\mathbf{m}\circ\operatorname{mult}^{-1}\right)\left((k_w)_{w\in\mathfrak{L}}\right)\right)} = \prod_{w\in\mathfrak{L}} \left(\mathbf{x}_w t^{\ell(w)}\right)^{k_w}.$$

This proves Claim 1.]

Now, in the ring $(\mathbf{k}[[\mathbf{x}]])[[t]]$ of formal power series, we have

$$\begin{aligned}\sum_{w\in\mathfrak{A}^*} \mathbf{x}_w t^{\ell(w)} &= \sum_{k\in\mathfrak{N}} \mathbf{x}_{\left(\mathbf{m}\circ\operatorname{mult}^{-1}\right)(k)} t^{\ell\left(\left(\mathbf{m}\circ\operatorname{mult}^{-1}\right)(k)\right)} \\ &\qquad \left(\begin{array}{c}\text{here, we have substituted } \left(\mathbf{m}\circ\operatorname{mult}^{-1}\right)(k) \text{ for } w \text{ in the sum,} \\ \text{since the map } \mathbf{m}\circ\operatorname{mult}^{-1} : \mathfrak{N}\to\mathfrak{A}^* \text{ is a bijection}\end{array}\right) \\ &= \sum_{(k_w)_{w\in\mathfrak{L}}\in\mathfrak{N}} \underbrace{\mathbf{x}_{\left(\mathbf{m}\circ\operatorname{mult}^{-1}\right)\left((k_w)_{w\in\mathfrak{L}}\right)} t^{\ell\left(\left(\mathbf{m}\circ\operatorname{mult}^{-1}\right)\left((k_w)_{w\in\mathfrak{L}}\right)\right)}}_{\substack{=\prod_{w\in\mathfrak{L}}\left(\mathbf{x}_w t^{\ell(w)}\right)^{k_w} \\ \text{(by Claim 1)}}} \\ &\qquad \left(\begin{array}{c}\text{here, we have renamed the summation index } k \text{ as } (k_w)_{w\in\mathfrak{L}}, \\ \text{since each } k\in\mathfrak{N} \text{ is a family indexed by elements of } \mathfrak{L}\end{array}\right) \\ &= \sum_{(k_w)_{w\in\mathfrak{L}}\in\mathfrak{N}} \prod_{w\in\mathfrak{L}} \left(\mathbf{x}_w t^{\ell(w)}\right)^{k_w} \\ &= \prod_{w\in\mathfrak{L}} \frac{1}{1-\mathbf{x}_w t^{\ell(w)}} \qquad \text{(by (13.192.3))}.\end{aligned}$$

This proves Lemma 13.192.5. □

Proof of Proposition 6.6.39. Lemma 13.192.5 yields

$$\prod_{w\in\mathfrak{L}} \frac{1}{1-\mathbf{x}_w t^{\ell(w)}} = \sum_{w\in\mathfrak{A}^*} \mathbf{x}_w t^{\ell(w)} = \frac{1}{1-p_1 t}$$

(by Lemma 13.192.4). This proves Proposition 6.6.39. □

Our next step is to show the following lemma (similar to Lemma 13.192.5):

Lemma 13.192.6. *In the power series ring* $\mathbf{k}[[\mathbf{x},\mathbf{y}]]$, *we have*

$$\prod_{w\in\mathfrak{L}} \frac{1}{1-\mathbf{x}_w p_{\ell(w)}(\mathbf{y})} = \sum_{w\in\mathfrak{A}^*} \mathbf{x}_w p_{\operatorname{CFLtype} w}(\mathbf{y}).$$

Proof of Lemma 13.192.6. Let $\mathfrak{M}$ denote the set of all finite multisets of Lyndon words. Define two maps $\mathbf{m}:\mathfrak{M}\to\mathfrak{A}^*$ and $\mathbf{n}:\mathfrak{A}^*\to\mathfrak{M}$ as in Proposition 13.143.1. Then, Proposition 13.143.1 shows that the maps $\mathbf{m}$ and $\mathbf{n}$ are mutually inverse bijections.

Recall that $\mathfrak{L}$ is the set of all Lyndon words. Thus, the Lyndon words are precisely the elements of $\mathfrak{L}$. But the definition of $\mathfrak{M}$ says that $\mathfrak{M}$ is the set of all finite multisets of Lyndon words. In other words, $\mathfrak{M}$ is the set of all finite multisets of elements of $\mathfrak{L}$ (since the Lyndon words are precisely the elements of $\mathfrak{L}$).

Let $\mathfrak{N}$ be the set of all families $(k_w)_{w\in\mathfrak{L}}\in\mathbb{N}^{\mathfrak{L}}$ of nonnegative integers (indexed by the Lyndon words) such that all but finitely many $w\in\mathfrak{L}$ satisfy $k_w=0$. Thus,

$$(13.192.7) \qquad \sum_{\substack{(k_w)_{w\in\mathfrak{L}}\in\mathbb{N}^{\mathfrak{L}}; \\ \text{all but finitely many } w\in\mathfrak{L} \\ \text{satisfy } k_w=0}} = \sum_{(k_w)_{w\in\mathfrak{L}}\in\mathfrak{N}}$$

(an equality of summation signs). Proposition 13.143.2 (applied to $S=\mathfrak{L}$) shows that the map $\operatorname{mult}:\mathfrak{M}\to\mathfrak{N}$ that sends each multiset $M\in\mathfrak{M}$ to the family

$$\left((\text{multiplicity of } w \text{ in the multiset } M)\right)_{w\in\mathfrak{L}}\in\mathfrak{N}$$

is well-defined and is a bijection. Consider this map mult. Since $\operatorname{mult}:\mathfrak{M}\to\mathfrak{N}$ is a bijection, its inverse $\operatorname{mult}^{-1}:\mathfrak{N}\to\mathfrak{M}$ is well-defined and also a bijection.

For any $w \in \mathfrak{L}$, the power series $\mathbf{x}_w p_{\ell(w)}(\mathbf{y})$ is homogeneous of positive degree[1206]. Hence, for any $w \in \mathfrak{L}$, the formal power series $\dfrac{1}{1 - \mathbf{x}_w p_{\ell(w)}(\mathbf{y})}$ is well-defined and satisfies the equality

$$\frac{1}{1 - \mathbf{x}_w p_{\ell(w)}(\mathbf{y})} = \sum_{k \in \mathbb{N}} \left(\mathbf{x}_w p_{\ell(w)}(\mathbf{y})\right)^k$$

(since $\dfrac{1}{1-q} = \sum_{k \in \mathbb{N}} q^k$ in the ring $\mathbf{k}[[q]]$ of formal power series).

Multiplying these equalities over all $w \in \mathfrak{L}$, we obtain

$$\begin{aligned} \prod_{w \in \mathfrak{L}} \frac{1}{1 - \mathbf{x}_w p_{\ell(w)}(\mathbf{y})} &= \prod_{w \in \mathfrak{L}} \sum_{k \in \mathbb{N}} \left(\mathbf{x}_w p_{\ell(w)}(\mathbf{y})\right)^k \\ &= \sum_{\substack{(k_w)_{w \in \mathfrak{L}} \in \mathbb{N}^{\mathfrak{L}};\\ \text{all but finitely many } w \in \mathfrak{L} \\ \text{satisfy } k_w = 0}} \prod_{w \in \mathfrak{L}} \left(\mathbf{x}_w p_{\ell(w)}(\mathbf{y})\right)^{k_w} \qquad \text{(by the product rule)} \\ &= \sum_{(k_w)_{w \in \mathfrak{L}} \in \mathfrak{N}} \prod_{w \in \mathfrak{L}} \left(\mathbf{x}_w p_{\ell(w)}(\mathbf{y})\right)^{k_w} \qquad \text{(by (13.192.7))}. \end{aligned} \tag{13.192.8}$$

The composition $\mathbf{m} \circ \operatorname{mult}^{-1} : \mathfrak{N} \to \mathfrak{A}^*$ of the bijections $\mathbf{m}$ and mult^{-1} is clearly a bijection. We now claim the following:

Claim 1: Let $w \in \mathfrak{A}^*$. For each $v \in \mathfrak{L}$, let $k_v \in \mathbb{N}$ be such that

$$k_v = (\text{multiplicity of } v \text{ in the multiset } \mathbf{n}(w)). \tag{13.192.9}$$

Then,

$$p_{\operatorname{CFLtype} w} = \prod_{v \in \mathfrak{L}} p_{\ell(v)}^{k_v}.$$

[*Proof of Claim 1:* Let $(a_1, a_2, \ldots, a_k)$ be the CFL factorization of w. Then, the definition of the map $\mathbf{n}$ yields $\mathbf{n}(w) = \{a_1, a_2, \ldots, a_k\}_{\text{multiset}}$. Now, for each $v \in \mathfrak{L}$, we have

$$\begin{aligned} k_v &= (\text{multiplicity of } v \text{ in the multiset } \mathbf{n}(w)) \qquad \text{(by (13.192.9))} \\ &= (\text{multiplicity of } v \text{ in the multiset } \{a_1, a_2, \ldots, a_k\}_{\text{multiset}}) \\ &\qquad (\text{since } \mathbf{n}(w) = \{a_1, a_2, \ldots, a_k\}_{\text{multiset}}) \\ &= (\text{the number of all } i \in \{1, 2, \ldots, k\} \text{ such that } a_i = v). \end{aligned} \tag{13.192.10}$$

Recall that $(a_1, a_2, \ldots, a_k)$ is the CFL factorization of w. In other words, $(a_1, a_2, \ldots, a_k)$ is a tuple of Lyndon words satisfying $w = a_1 a_2 \cdots a_k$ and $a_1 \geq a_2 \geq \cdots \geq a_k$ (by the definition of "CFL factorization"). Thus, in particular, $(a_1, a_2, \ldots, a_k)$ is a tuple of Lyndon words, i.e., a tuple of elements of $\mathfrak{L}$ (since the Lyndon words are precisely the elements of $\mathfrak{L}$). In other words, $a_i \in \mathfrak{L}$ for each $i \in \{1, 2, \ldots, k\}$.

Let us make a general observation about partitions: If $b_1, b_2, \ldots, b_k$ are any k positive integers, and if μ is the partition whose parts are the positive integers $b_1, b_2, \ldots, b_k$ (listed in decreasing order), then

$$p_\lambda = p_{b_1} p_{b_2} \cdots p_{b_k}. \tag{13.192.11}$$

(Indeed, this follows from the definition of p_λ in Definition 2.2.1, since multiplication in Λ is commutative.)

[1206] *Proof.* Let $w \in \mathfrak{L}$. Thus, w is a Lyndon word (since $\mathfrak{L}$ is the set of all Lyndon words), therefore a nonempty word (since any Lyndon word is nonempty by definition). In other words, $w \in \mathfrak{A}^n$ for some positive integer n. Consider this n. From $w \in \mathfrak{A}^n$, we obtain $w = (w_1, w_2, \ldots, w_n)$, so that $\mathbf{x}_w = x_{w_1} x_{w_2} \cdots x_{w_n}$ (by the definition of $\mathbf{x}_w$). Hence, $\mathbf{x}_w$ is a monomial of degree n. Hence, the power series $\mathbf{x}_w$ is homogeneous of degree n. Moreover, $w \in \mathfrak{A}^n$ entails $\ell(w) = n$. Hence, $p_{\ell(w)} = p_n = x_1^n + x_2^n + x_3^n + \cdots$ (by the definition of p_n, since n is a positive integer). Substituting $y_1, y_2, y_3, \ldots$ for $x_1, x_2, x_3, \ldots$ in this equality, we find $p_{\ell(w)}(\mathbf{y}) = y_1^n + y_2^n + y_3^n + \cdots$. Thus, the power series $p_{\ell(w)}(\mathbf{y})$ is homogeneous of degree n.

We now know that the power series $\mathbf{x}_w$ and $p_{\ell(w)}(\mathbf{y})$ are both homogeneous of degree n. Hence, their product $\mathbf{x}_w p_{\ell(w)}(\mathbf{y})$ is homogeneous of degree $n + n$. Since $n + n = 2n$ is positive (because n is positive), this shows that $\mathbf{x}_w p_{\ell(w)}(\mathbf{y})$ is homogeneous of positive degree. Qed.

Now, recall again that $(a_1, a_2, \ldots, a_k)$ is the CFL factorization of w. Hence, $\operatorname{CFLtype} w$ is the partition whose parts are the positive integers $\ell(a_1), \ell(a_2), \ldots, \ell(a_k)$ (listed in decreasing order)[1207]. Hence, (13.192.11) (applied to $b_i = a_i$ and $\mu = \operatorname{CFLtype} w$) yields

$$\begin{aligned} p_{\operatorname{CFLtype} w} &= p_{\ell(a_1)} p_{\ell(a_2)} \cdots p_{\ell(a_k)} = \underbrace{\prod_{i\in\{1,2,\ldots,k\}}}_{\substack{=\prod_{v\in\mathfrak{L}} \prod_{i\in\{1,2,\ldots,k\};\ a_i=v} \\ (\text{since } a_i\in\mathfrak{L} \text{ for each } i\in\{1,2,\ldots,k\})}} p_{\ell(a_i)} \\ &= \prod_{v\in\mathfrak{L}} \prod_{\substack{i\in\{1,2,\ldots,k\};\\ a_i=v}} \underbrace{p_{\ell(a_i)}}_{\substack{=p_{\ell(v)} \\ (\text{since } a_i=v)}} = \prod_{v\in\mathfrak{L}} \underbrace{\prod_{\substack{i\in\{1,2,\ldots,k\};\\ a_i=v}} p_{\ell(v)}}_{\substack{=p_{\ell(v)}^{(\text{the number of all } i\in\{1,2,\ldots,k\} \text{ such that } a_i=v)} = p_{\ell(v)}^{k_v} \\ (\text{since (the number of all } i\in\{1,2,\ldots,k\} \text{ such that } a_i=v)=k_v \\ (\text{by } (13.192.10)))}} \\ &= \prod_{v\in\mathfrak{L}} p_{\ell(v)}^{k_v}. \end{aligned}$$

This proves Claim 1.]

Claim 2: Let $(k_w)_{w\in\mathfrak{L}} \in \mathfrak{N}$. Then,

$$\mathbf{x}_{(\mathbf{m}\circ\operatorname{mult}^{-1})((k_w)_{w\in\mathfrak{L}})} p_{\operatorname{CFLtype}((\mathbf{m}\circ\operatorname{mult}^{-1})((k_w)_{w\in\mathfrak{L}}))}(\mathbf{y}) = \prod_{w\in\mathfrak{L}} \left(\mathbf{x}_w p_{\ell(w)}(\mathbf{y})\right)^{k_w}.$$

[*Proof of Claim 2:* Let $M = \operatorname{mult}^{-1}\left((k_w)_{w\in\mathfrak{L}}\right)$. Then,

$$(k_w)_{w\in\mathfrak{L}} = \operatorname{mult} M = ((\text{multiplicity of } w \text{ in the multiset } M))_{w\in\mathfrak{L}}$$

(by the definition of the map mult). In other words, every $w \in \mathfrak{L}$ satisfies

$$k_w = (\text{multiplicity of } w \text{ in the multiset } M).$$

Renaming w as v in this result, we obtain the following: Every $v \in \mathfrak{L}$ satisfies

$$k_v = (\text{multiplicity of } v \text{ in the multiset } M). \tag{13.192.12}$$

Moreover, $M = \operatorname{mult}^{-1}\left((k_w)_{w\in\mathfrak{L}}\right) \in \mathfrak{M}$. In other words, M is a finite multiset of elements of $\mathfrak{L}$ (since $\mathfrak{M}$ is the set of all finite multisets of elements of $\mathfrak{L}$). Let $a_1, a_2, \ldots, a_k$ denote the elements of this multiset M listed in decreasing order. Then, the definition of $\mathbf{m}$ yields $\mathbf{m}(M) = a_1 a_2 \cdots a_k$. Also, $a_1, a_2, \ldots, a_k$ are elements of M (by the definition of $a_1, a_2, \ldots, a_k$), and thus belong to $\mathfrak{L}$ (since M is a finite multiset of elements of $\mathfrak{L}$). In other words, every a_i belongs to $\mathfrak{L}$.

Each $w \in \mathfrak{L}$ satisfies

$$(\text{the number of } i \in \{1, 2, \ldots, k\} \text{ satisfying } a_i = w) = k_w. \tag{13.192.13}$$

(Indeed, this is precisely the equality (13.192.5) from the proof of Claim 1 in the proof of Lemma 13.192.5; and it can be proved in the exact same way as the latter equality.)

Now, recall that the maps $\mathbf{m}$ and $\mathbf{n}$ are mutually inverse. Hence, $\mathbf{n} \circ \mathbf{m} = \operatorname{id}$. Thus, $\mathbf{n}(\mathbf{m}(M)) = \underbrace{(\mathbf{n} \circ \mathbf{m})}_{=\operatorname{id}}(M) = \operatorname{id}(M) = M$. In other words, $M = \mathbf{n}(\mathbf{m}(M))$. Now, every $v \in \mathfrak{L}$ satisfies

$$\begin{aligned} k_v &= \left(\text{multiplicity of } v \text{ in the multiset } \underbrace{M}_{=\mathbf{n}(\mathbf{m}(M))}\right) \qquad (\text{by } (13.192.12)) \\ &= (\text{multiplicity of } v \text{ in the multiset } \mathbf{n}(\mathbf{m}(M))). \end{aligned}$$

Thus, Claim 1 (applied to $w = \mathbf{m}(M)$) yields

$$p_{\operatorname{CFLtype}(\mathbf{m}(M))} = \prod_{v\in\mathfrak{L}} p_{\ell(v)}^{k_v} = \prod_{w\in\mathfrak{L}} p_{\ell(w)}^{k_w}$$

[1207] by the definition of $\operatorname{CFLtype} w$

(here, we have renamed the index v as w in the product). Substituting the variables $\mathbf{y} = (y_1, y_2, y_3, \ldots)$ for $\mathbf{x} = (x_1, x_2, x_3, \ldots)$ on both sides of this equality, we obtain

$$p_{\mathrm{CFLtype}(\mathbf{m}(M))}(\mathbf{y}) = \left(\prod_{w\in\mathfrak{L}} p_{\ell(w)}^{k_w}\right)(\mathbf{y}) = \prod_{w\in\mathfrak{L}} \left(p_{\ell(w)}(\mathbf{y})\right)^{k_w}. \tag{13.192.14}$$

Moreover, from $\mathbf{m}(M) = a_1 a_2 \cdots a_k$, we obtain

$$\begin{aligned}
\mathbf{x}_{\mathbf{m}(M)} &= \mathbf{x}_{a_1 a_2 \cdots a_k} = \mathbf{x}_{a_1}\mathbf{x}_{a_2}\cdots\mathbf{x}_{a_k} \qquad (\text{by Lemma 13.192.2})\\
&= \underbrace{\prod_{i\in\{1,2,\ldots,k\}}}_{\substack{=\prod_{w\in\mathfrak{L}}\prod_{i\in\{1,2,\ldots,k\};\ a_i=w}\\ (\text{since every } a_i \text{ belongs to } \mathfrak{L})}} \mathbf{x}_{a_i} = \prod_{w\in\mathfrak{L}}\prod_{\substack{i\in\{1,2,\ldots,k\};\\ a_i=w}} \underbrace{\mathbf{x}_{a_i}}_{\substack{=\mathbf{x}_w\\ (\text{since } a_i=w)}}\\
&= \prod_{w\in\mathfrak{L}} \underbrace{\prod_{\substack{i\in\{1,2,\ldots,k\};\\ a_i=w}} \mathbf{x}_w}_{\substack{=\mathbf{x}_w^{(\text{the number of } i\in\{1,2,\ldots,k\} \text{ satisfying } a_i=w)}=\mathbf{x}_w^{k_w}\\ (\text{since (the number of } i\in\{1,2,\ldots,k\} \text{ satisfying } a_i=w)=k_w\\ (\text{by (13.192.13)}))}}\\
&= \prod_{w\in\mathfrak{L}} \mathbf{x}_w^{k_w}.
\end{aligned}$$

Multiplying this equality by (13.192.14), we obtain

$$\begin{aligned}
\mathbf{x}_{\mathbf{m}(M)} p_{\mathrm{CFLtype}(\mathbf{m}(M))}(\mathbf{y}) &= \left(\prod_{w\in\mathfrak{L}} \mathbf{x}_w^{k_w}\right) \prod_{w\in\mathfrak{L}} \left(p_{\ell(w)}(\mathbf{y})\right)^{k_w} = \prod_{w\in\mathfrak{L}} \underbrace{\left(\mathbf{x}_w^{k_w}\left(p_{\ell(w)}(\mathbf{y})\right)^{k_w}\right)}_{=\left(\mathbf{x}_w p_{\ell(w)}(\mathbf{y})\right)^{k_w}}\\
&= \prod_{w\in\mathfrak{L}} \left(\mathbf{x}_w p_{\ell(w)}(\mathbf{y})\right)^{k_w}.
\end{aligned}$$

In view of

$$\mathbf{m}\left(\underbrace{M}_{=\mathrm{mult}^{-1}\left((k_w)_{w\in\mathfrak{L}}\right)}\right) = \mathbf{m}\left(\mathrm{mult}^{-1}\left((k_w)_{w\in\mathfrak{L}}\right)\right) = \left(\mathbf{m}\circ\mathrm{mult}^{-1}\right)\left((k_w)_{w\in\mathfrak{L}}\right),$$

this rewrites as

$$\mathbf{x}_{\left(\mathbf{m}\circ\mathrm{mult}^{-1}\right)\left((k_w)_{w\in\mathfrak{L}}\right)} p_{\mathrm{CFLtype}\left(\left(\mathbf{m}\circ\mathrm{mult}^{-1}\right)\left((k_w)_{w\in\mathfrak{L}}\right)\right)}(\mathbf{y}) = \prod_{w\in\mathfrak{L}} \left(\mathbf{x}_w p_{\ell(w)}(\mathbf{y})\right)^{k_w}.$$

This proves Claim 2.]

Now, in the ring $\mathbf{k}[[\mathbf{x},\mathbf{y}]]$ of formal power series, we have

$$\begin{aligned}
\sum_{w\in\mathfrak{A}^*} \mathbf{x}_w p_{\operatorname{CFLtype} w}(\mathbf{y}) &= \sum_{k\in\mathfrak{N}} \mathbf{x}_{\left(\mathbf{m}\circ\operatorname{mult}^{-1}\right)(k)} p_{\operatorname{CFLtype}\left(\left(\mathbf{m}\circ\operatorname{mult}^{-1}\right)(k)\right)}(\mathbf{y}) \\
&\qquad \left(\begin{array}{c}\text{here, we have substituted } \left(\mathbf{m}\circ\operatorname{mult}^{-1}\right)(k) \text{ for } w \text{ in the sum,} \\ \text{since the map } \mathbf{m}\circ\operatorname{mult}^{-1} : \mathfrak{N}\to\mathfrak{A}^* \text{ is a bijection}\end{array}\right) \\
&= \sum_{(k_w)_{w\in\mathfrak{L}}\in\mathfrak{N}} \underbrace{\mathbf{x}_{\left(\mathbf{m}\circ\operatorname{mult}^{-1}\right)\left((k_w)_{w\in\mathfrak{L}}\right)} p_{\operatorname{CFLtype}\left(\left(\mathbf{m}\circ\operatorname{mult}^{-1}\right)\left((k_w)_{w\in\mathfrak{L}}\right)\right)}(\mathbf{y})}_{\substack{=\prod_{w\in\mathfrak{L}}\left(\mathbf{x}_w p_{\ell(w)}(\mathbf{y})\right)^{k_w} \\ \text{(by Claim 2)}}} \\
&\qquad \left(\begin{array}{c}\text{here, we have renamed the summation index } k \text{ as } (k_w)_{w\in\mathfrak{L}}, \\ \text{since each } k\in\mathfrak{N} \text{ is a family indexed by elements of } \mathfrak{L}\end{array}\right) \\
&= \sum_{(k_w)_{w\in\mathfrak{L}}\in\mathfrak{N}} \prod_{w\in\mathfrak{L}} \left(\mathbf{x}_w p_{\ell(w)}(\mathbf{y})\right)^{k_w} \\
&= \prod_{w\in\mathfrak{L}} \frac{1}{1-\mathbf{x}_w p_{\ell(w)}(\mathbf{y})} \qquad \text{(by (13.192.8))}.
\end{aligned}$$

This proves Lemma 13.192.6. □

The following lemma is a slight restatement of Lemma 13.192.6:

Lemma 13.192.7. *Consider the power series ring $\mathbf{k}[[\mathbf{x},\mathbf{y}]]$. For each word $w=(w_1,w_2,\ldots,w_n)\in\mathfrak{A}^n$, we define a monomial $\mathbf{y}_w$ in $\mathbf{k}[[\mathbf{y}]]$ by $\mathbf{y}_w = y_{w_1}y_{w_2}\cdots y_{w_n}$. Then,*

$$\prod_{w\in\mathfrak{L}}\prod_{u\in\mathfrak{L}} \frac{1}{1-\mathbf{x}_w^{\ell(u)}\mathbf{y}_u^{\ell(w)}} = \sum_{w\in\mathfrak{A}^*} \mathbf{x}_w p_{\operatorname{CFLtype} w}(\mathbf{y}).$$

Proof of Lemma 13.192.7. Lemma 13.192.6 yields

$$\prod_{w\in\mathfrak{L}} \frac{1}{1-\mathbf{x}_w p_{\ell(w)}(\mathbf{y})} = \sum_{w\in\mathfrak{A}^*} \mathbf{x}_w p_{\operatorname{CFLtype} w}(\mathbf{y}). \tag{13.192.15}$$

Proposition 6.6.39 yields

$$\frac{1}{1-p_1 t} = \prod_{w\in\mathfrak{L}} \frac{1}{1-\mathbf{x}_w t^{\ell(w)}}. \tag{13.192.16}$$

Now, fix $v\in\mathfrak{L}$. Then, v is a Lyndon word (since $\mathfrak{L}$ is the set of all Lyndon words), and thus is nonempty (since any Lyndon word is nonempty). Hence, $\ell(v)\geq 1$. Thus, the definition of $p_{\ell(v)}$ yields $p_{\ell(v)} = x_1^{\ell(v)} + x_2^{\ell(v)} + x_3^{\ell(v)} + \cdots$. Substituting the variables $\mathbf{y}=(y_1,y_2,y_3,\ldots)$ for $\mathbf{x}=(x_1,x_2,x_3,\ldots)$ on both sides of this equality, we obtain

$$p_{\ell(v)}(\mathbf{y}) = y_1^{\ell(v)} + y_2^{\ell(v)} + y_3^{\ell(v)} + \cdots. \tag{13.192.17}$$

On the other hand, the definition of p_1 yields $p_1 = x_1^1 + x_2^1 + x_3^1 + \cdots = x_1 + x_2 + x_3 + \cdots$. Substituting $y_1^{\ell(v)}, y_2^{\ell(v)}, y_3^{\ell(v)}, \ldots$ for $x_1,x_2,x_3,\ldots$ on both sides of this equality, we obtain

$$p_1\left(y_1^{\ell(v)}, y_2^{\ell(v)}, y_3^{\ell(v)}, \ldots\right) = y_1^{\ell(v)} + y_2^{\ell(v)} + y_3^{\ell(v)} + \cdots.$$

Comparing this with (13.192.17), we obtain

$$p_1\left(y_1^{\ell(v)}, y_2^{\ell(v)}, y_3^{\ell(v)}, \ldots\right) = p_{\ell(v)}(\mathbf{y}). \tag{13.192.18}$$

Every $k\in\mathbb{N}$ and every word $w\in\mathfrak{A}^*$ satisfy

$$\mathbf{x}_w\left(y_1^k, y_2^k, y_3^k, \ldots\right) = \mathbf{y}_w^k \tag{13.192.19}$$

(where, of course, $\mathbf{x}_w\left(y_1^k, y_2^k, y_3^k, \ldots\right)$ denotes the result of substituting $y_1^k, y_2^k, y_3^k, \ldots$ for $x_1,x_2,x_3,\ldots$ in the polynomial $\mathbf{x}_w$).

[*Proof of (13.192.19):* Let $k\in\mathbb{N}$. Let $w\in\mathfrak{A}^*$ be a word. Write the word w in the form $w=(w_1,w_2,\ldots,w_n)$. Then, $\mathbf{x}_w = x_{w_1}x_{w_2}\cdots x_{w_n}$ (by the definition of $\mathbf{x}_w$) and $\mathbf{y}_w = y_{w_1}y_{w_2}\cdots y_{w_n}$ (by the definition of $\mathbf{y}_w$), so that $y_{w_1}y_{w_2}\cdots y_{w_n} = \mathbf{y}_w$. Now, substituting $y_1^k, y_2^k, y_3^k, \ldots$ for $x_1, x_2, x_3, \ldots$ on both sides of the equality $\mathbf{x}_w = x_{w_1}x_{w_2}\cdots x_{w_n}$, we obtain

$$\mathbf{x}_w\left(y_1^k, y_2^k, y_3^k, \ldots\right) = y_{w_1}^k y_{w_2}^k \cdots y_{w_n}^k = \left(\underbrace{y_{w_1}y_{w_2}\cdots y_{w_n}}_{=\mathbf{y}_w}\right)^k = \mathbf{y}_w^k.$$

This proves (13.192.19).]

Furthermore, the monomial $\mathbf{x}_v$ has positive degree[1208].

The equality (13.192.16) is an equality of formal power series in $(\mathbf{k}[[\mathbf{x}]])[[t]] = \mathbf{k}[[x_1, x_2, x_3, \ldots, t]]$. Substituting $y_1^{\ell(v)}, y_2^{\ell(v)}, y_3^{\ell(v)}, \ldots$ for $x_1, x_2, x_3, \ldots$ on both sides of this equality (while keeping t unchanged for now), we obtain

$$\begin{aligned}\frac{1}{1-p_1\left(y_1^{\ell(v)}, y_2^{\ell(v)}, y_3^{\ell(v)}, \ldots\right)t} &= \prod_{w\in\mathfrak{L}} \underbrace{\frac{1}{1-\mathbf{x}_w\left(y_1^{\ell(v)}, y_2^{\ell(v)}, y_3^{\ell(v)}, \ldots\right)t^{\ell(w)}}}_{\substack{=\dfrac{1}{1-\mathbf{y}_w^{\ell(v)}t^{\ell(w)}} \\ \text{(since } \mathbf{x}_w\left(y_1^{\ell(v)}, y_2^{\ell(v)}, y_3^{\ell(v)}, \ldots\right)=\mathbf{y}_w^{\ell(v)} \\ \text{(by (13.192.19), applied to } k=\ell(v)\text{)})}} \\ &= \prod_{w\in\mathfrak{L}} \frac{1}{1-\mathbf{y}_w^{\ell(v)}t^{\ell(w)}} = \prod_{u\in\mathfrak{L}} \frac{1}{1-\mathbf{y}_u^{\ell(v)}t^{\ell(u)}}\end{aligned}$$

(here, we have renamed the index w as u in the product). In view of (13.192.18), this rewrites as

$$\frac{1}{1-p_{\ell(v)}(\mathbf{y})\,t} = \prod_{u\in\mathfrak{L}} \frac{1}{1-\mathbf{y}_u^{\ell(v)}t^{\ell(u)}}.$$

This is an equality of formal power series in $(\mathbf{k}[[\mathbf{y}]])[[t]]$. We can substitute $\mathbf{x}_v$ for t in this equality (since the monomial $\mathbf{x}_v$ has positive degree), and thus obtain the equality

$$\frac{1}{1-p_{\ell(v)}(\mathbf{y})\,\mathbf{x}_v} = \prod_{u\in\mathfrak{L}} \frac{1}{1-\mathbf{y}_u^{\ell(v)}\mathbf{x}_v^{\ell(u)}} \tag{13.192.20}$$

in the ring $\mathbf{k}[[\mathbf{x}, \mathbf{y}]]$.

Now, forget that we fixed v. We thus have proved the equality (13.192.20) for each $v \in \mathfrak{L}$.

Now,

$$\begin{aligned}\prod_{w\in\mathfrak{L}} \underbrace{\frac{1}{1-\mathbf{x}_w p_{\ell(w)}(\mathbf{y})}}_{\substack{=\dfrac{1}{1-p_{\ell(w)}(\mathbf{y})\,\mathbf{x}_w} \\ \text{(since } \mathbf{x}_w p_{\ell(w)}(\mathbf{y}) = p_{\ell(w)}(\mathbf{y})\mathbf{x}_w\text{)}}} &= \prod_{w\in\mathfrak{L}} \underbrace{\frac{1}{1-p_{\ell(w)}(\mathbf{y})\,\mathbf{x}_w}}_{\substack{=\prod_{u\in\mathfrak{L}}\dfrac{1}{1-\mathbf{y}_u^{\ell(w)}\mathbf{x}_w^{\ell(u)}} \\ \text{(by (13.192.20), applied to } v=w\text{)}}} \\ &= \prod_{w\in\mathfrak{L}} \prod_{u\in\mathfrak{L}} \underbrace{\frac{1}{1-\mathbf{y}_u^{\ell(w)}\mathbf{x}_w^{\ell(u)}}}_{\substack{=\dfrac{1}{1-\mathbf{x}_w^{\ell(u)}\mathbf{y}_u^{\ell(w)}} \\ \text{(since } \mathbf{y}_u^{\ell(w)}\mathbf{x}_w^{\ell(u)} = \mathbf{x}_w^{\ell(u)}\mathbf{y}_u^{\ell(w)}\text{)}}} = \prod_{w\in\mathfrak{L}} \prod_{u\in\mathfrak{L}} \frac{1}{1-\mathbf{x}_w^{\ell(u)}\mathbf{y}_u^{\ell(w)}}.\end{aligned}$$

Comparing this with (13.192.15), we obtain

$$\prod_{w\in\mathfrak{L}} \prod_{u\in\mathfrak{L}} \frac{1}{1-\mathbf{x}_w^{\ell(u)}\mathbf{y}_u^{\ell(w)}} = \sum_{w\in\mathfrak{A}^*} \mathbf{x}_w p_{\operatorname{CFLtype} w}(\mathbf{y}).$$

This proves Lemma 13.192.7. □

[1208] *Proof.* Write the word v in the form $v = (v_1, v_2, \ldots, v_n)$. Then, $n = \ell(v) \geq 1$. Hence, n is positive. But the definition of $\mathbf{x}_v$ yields $\mathbf{x}_v = x_{v_1}x_{v_2}\cdots x_{v_n}$ (since $v = (v_1, v_2, \ldots, v_n)$). Hence, the monomial $\mathbf{x}_v$ has degree n. Thus, the monomial $\mathbf{x}_v$ has positive degree (since n is positive).

Now, we can prove Proposition 6.6.38:

Proof of Proposition 6.6.38. Define a monomial $\mathbf{y}_w$ for each word w as in Proposition 6.6.38(b). We have

$$\sum_{\lambda\in\operatorname{Par}} \underbrace{\mathbf{GR}_\lambda(\mathbf{x})}_{\substack{=\sum_{w\in\mathfrak{A}^*;\ \operatorname{CFLtype} w=\lambda} \mathbf{x}_w \\ \text{(by the definition of } \mathbf{GR}_\lambda)}} p_\lambda(\mathbf{y}) = \sum_{\lambda\in\operatorname{Par}} \sum_{\substack{w\in\mathfrak{A}^*;\\ \operatorname{CFLtype} w=\lambda}} \mathbf{x}_w \underbrace{p_\lambda(\mathbf{y})}_{\substack{=p_{\operatorname{CFLtype} w}(\mathbf{y}) \\ (\text{since } \lambda=\operatorname{CFLtype} w)}}$$

$$= \underbrace{\sum_{\lambda\in\operatorname{Par}} \sum_{\substack{w\in\mathfrak{A}^*;\\ \operatorname{CFLtype} w=\lambda}}}_{\substack{=\sum_{w\in\mathfrak{A}^*} \\ (\text{since } \operatorname{CFLtype} w\in\operatorname{Par} \\ \text{for each } w\in\mathfrak{A}^*)}} \mathbf{x}_w p_{\operatorname{CFLtype} w}(\mathbf{y})$$

$$= \sum_{w\in\mathfrak{A}^*} \mathbf{x}_w p_{\operatorname{CFLtype} w}(\mathbf{y}) \tag{13.192.21}$$

$$= \prod_{w\in\mathfrak{L}} \prod_{u\in\mathfrak{L}} \frac{1}{1-\mathbf{x}_w^{\ell(u)}\mathbf{y}_u^{\ell(w)}} \tag{13.192.22}$$

(by Lemma 13.192.7).

The same argument (but with the roles of the sets of variables $\mathbf{x}=(x_1,x_2,x_3,\ldots)$ and $\mathbf{y}=(y_1,y_2,y_3,\ldots)$ interchanged) yields

$$\sum_{\lambda\in\operatorname{Par}} \mathbf{GR}_\lambda(\mathbf{y})\, p_\lambda(\mathbf{x}) = \prod_{w\in\mathfrak{L}} \prod_{u\in\mathfrak{L}} \frac{1}{1-\mathbf{y}_w^{\ell(u)}\mathbf{x}_u^{\ell(w)}}.$$

Hence,

$$\sum_{\lambda\in\operatorname{Par}} \underbrace{p_\lambda(\mathbf{x})\,\mathbf{GR}_\lambda(\mathbf{y})}_{=\mathbf{GR}_\lambda(\mathbf{y})p_\lambda(\mathbf{x})}$$

$$= \sum_{\lambda\in\operatorname{Par}} \mathbf{GR}_\lambda(\mathbf{y})\, p_\lambda(\mathbf{x}) = \prod_{w\in\mathfrak{L}} \prod_{u\in\mathfrak{L}} \frac{1}{1-\mathbf{y}_w^{\ell(u)}\mathbf{x}_u^{\ell(w)}} = \underbrace{\prod_{u\in\mathfrak{L}} \prod_{w\in\mathfrak{L}}}_{=\prod_{w\in\mathfrak{L}}\prod_{u\in\mathfrak{L}}} \underbrace{\frac{1}{1-\mathbf{y}_u^{\ell(w)}\mathbf{x}_w^{\ell(u)}}}_{=\frac{1}{1-\mathbf{x}_w^{\ell(u)}\mathbf{y}_u^{\ell(w)}}}$$

(here, we have renamed the indices w and u as u and w)

$$= \prod_{w\in\mathfrak{L}} \prod_{u\in\mathfrak{L}} \frac{1}{1-\mathbf{x}_w^{\ell(u)}\mathbf{y}_u^{\ell(w)}}.$$

Combining this with (13.192.21) and (13.192.22), we obtain

$$\sum_{\lambda\in\operatorname{Par}} \mathbf{GR}_\lambda(\mathbf{x})\, p_\lambda(\mathbf{y}) = \sum_{w\in\mathfrak{A}^*} \mathbf{x}_w p_{\operatorname{CFLtype} w}(\mathbf{y}) = \prod_{w\in\mathfrak{L}} \prod_{u\in\mathfrak{L}} \frac{1}{1-\mathbf{x}_w^{\ell(u)}\mathbf{y}_u^{\ell(w)}}$$
$$= \sum_{\lambda\in\operatorname{Par}} p_\lambda(\mathbf{x})\,\mathbf{GR}_\lambda(\mathbf{y}).$$

This proves both parts (a) and (b) of Proposition 6.6.38. □

Next, let us prove Proposition 6.6.43:

Proof of Proposition 6.6.43. (a) The set N is a necklace. In other words, N is an n-necklace for some positive integer n. Consider this n.

Recall that $c\in C$ acts on $\mathfrak{A}^n$ by cyclically rotating n-tuples one step to the left (that is, it acts by the formula $c\cdot(a_1,a_2,\ldots,a_n)=(a_2,a_3,\ldots,a_n,a_1)$). Thus, the element c^n acts on $\mathfrak{A}^n$ by cyclically rotating n-tuples n steps to the left. In other words, the element c^n acts trivially on $\mathfrak{A}^n$ (since cyclically rotating an n-tuple n steps to the left does nothing). Hence, the whole subgroup $\langle c^n\rangle$ of C acts trivially on $\mathfrak{A}^n$.

The words w and w' belong to the same n-necklace (namely, to N). In other words, the words w and w' belong to the same orbit of the C-action on $\mathfrak{A}^n$ (since an n-necklace was defined to be an orbit of the

C-action on $\mathfrak{A}^n$). In other words, $C \cdot w = C \cdot w'$. Hence, $w' \in C \cdot w' = C \cdot w = \{c^p \cdot w \mid p \in \mathbb{Z}\}$ (since $C = \{c^p \mid p \in \mathbb{Z}\}$). In other words, there exists some $p \in \mathbb{Z}$ such that $w' = c^p \cdot w$. Consider this p.

Recall that n is a positive integer. Hence, we can divide p by n with remainder. Let q and r be the quotient and the remainder obtained when we divide p by n. Thus, $q \in \mathbb{Z}$ and $r \in \{0, 1, \ldots, n-1\}$ and $p = qn + r$. We have $c^{qn} = (c^n)^q \in \langle c^n \rangle$. Thus, c^{qn} acts trivially on $\mathfrak{A}^n$ (since the whole subgroup $\langle c^n \rangle$ of C acts trivially on $\mathfrak{A}^n$). Therefore, $c^{qn} \cdot w = w$.

But $p = qn + r = r + qn$, so that $c^p = c^{r+qn} = c^r c^{qn}$. Now,

$$w' = \underbrace{c^p}_{=c^r c^{qn}} \cdot w = (c^r c^{qn}) \cdot w = c^r \cdot \underbrace{(c^{qn} \cdot w)}_{=w} = c^r \cdot w.$$

Let us write the word w in the form $w = (w_1, w_2, \ldots, w_n)$. Recall that $c \in C$ acts on $\mathfrak{A}^n$ by cyclically rotating n-tuples one step to the left. Hence, c^r acts on $\mathfrak{A}^n$ by rotating n-tuples r steps to the left (since $r \in \{0, 1, \ldots, n-1\} \subset \mathbb{N}$). In other words,

$$c^r \cdot (a_1, a_2, \ldots, a_n) = (a_{r+1}, a_{r+2}, \ldots, a_n, a_1, a_2, \ldots, a_r) \qquad (13.192.23)$$

for each $(a_1, a_2, \ldots, a_n) \in \mathfrak{A}^n$ (since $r \in \{0, 1, \ldots, n-1\}$). Now,

$$\begin{aligned} w' &= c^r \cdot \underbrace{w}_{=(w_1, w_2, \ldots, w_n)} = c^r \cdot (w_1, w_2, \ldots, w_n) \\ &= (w_{r+1}, w_{r+2}, \ldots, w_n, w_1, w_2, \ldots, w_r) \qquad \text{(by (13.192.23), applied to } a_i = w_i\text{)}. \end{aligned}$$

Now, the two words $(w_1, w_2, \ldots, w_r)$ and $(w_{r+1}, w_{r+2}, \ldots, w_n)$ satisfy

$$w = (w_1, w_2, \ldots, w_n) = (w_1, w_2, \ldots, w_r)(w_{r+1}, w_{r+2}, \ldots, w_n)$$

and

$$w' = (w_{r+1}, w_{r+2}, \ldots, w_n, w_1, w_2, \ldots, w_r) = (w_{r+1}, w_{r+2}, \ldots, w_n)(w_1, w_2, \ldots, w_r).$$

Hence, there exist words u and v such that $w = uv$ and $w' = vu$ (namely, $u = (w_1, w_2, \ldots, w_r)$ and $v = (w_{r+1}, w_{r+2}, \ldots, w_n)$). This proves Proposition 6.6.43(a).

(b) Proposition 6.6.43(a) shows that there exist words u and v such that $w = uv$ and $w' = vu$. Consider these u and v. From $w = uv$, we obtain $\mathbf{x}_w = \mathbf{x}_{uv} = \mathbf{x}_u \mathbf{x}_v$ (by Lemma 13.192.1). From $w' = vu$, we obtain $\mathbf{x}_{w'} = \mathbf{x}_{vu} = \mathbf{x}_v \mathbf{x}_u$ (by Lemma 13.192.1, applied to v and u instead of u and v). Comparing this with $\mathbf{x}_w = \mathbf{x}_u \mathbf{x}_v = \mathbf{x}_v \mathbf{x}_u$, we obtain $\mathbf{x}_w = \mathbf{x}_{w'}$. Thus, Proposition 6.6.43(b) is proved. □

Next, we shall prove Proposition 6.6.48 using the following two almost trivial lemmas:

Lemma 13.192.8. *Let $a \in \mathfrak{A}^*$ be a nonempty word. Then, $\mathbf{x}_{[a]} = \mathbf{x}_a$.*

Proof of Lemma 13.192.8. Lemma 13.190.1 (applied to $u = a$) yields $a \in [a]$. In other words, a is an element of $[a]$.

Now, recall how the monomial $\mathbf{x}_N$ was defined for a necklace N: It was defined by setting $\mathbf{x}_N = \mathbf{x}_w$, where w is any element of N. Thus, if N is a necklace, then $\mathbf{x}_N = \mathbf{x}_w$ for any element w of N. Applying this to $N = [a]$ and $w = a$, we conclude that $\mathbf{x}_{[a]} = \mathbf{x}_a$ (since a is an element of $[a]$). This proves Lemma 13.192.8. □

Lemma 13.192.9. *Let λ be a partition. Let $w \in \mathfrak{A}^*$ be such that* $\operatorname{CFLtype} w = \lambda$. *Then:*

(a) *We have $\ell(w) = |\lambda|$.*

(b) *Let $(a_1, a_2, \ldots, a_k)$ be the CFL factorization of w. Then, $k = \ell(\lambda)$.*

Proof of Lemma 13.192.9. Let $(a_1, a_2, \ldots, a_k)$ be the CFL factorization of w. In other words, $(a_1, a_2, \ldots, a_k)$ is a tuple of Lyndon words satisfying $w = a_1 a_2 \cdots a_k$ and $a_1 \geq a_2 \geq \cdots \geq a_k$ (because this is how a "CFL factorization of w" is defined). From $\operatorname{CFLtype} w = \lambda$, we obtain

$$\begin{aligned} \lambda &= \operatorname{CFLtype} w \\ &= (\text{the partition obtained by listing the numbers } \ell(a_1), \ell(a_2), \ldots, \ell(a_k) \text{ in decreasing order}) \end{aligned}$$

(by the definition of $\operatorname{CFLtype} w$, since $(a_1, a_2, \ldots, a_k)$ is the CFL factorization of w). Hence, the partition λ is a rearrangement of the k-tuple $(\ell(a_1), \ell(a_2), \ldots, \ell(a_k))$ (of course, with trailing zeroes appended to it to form an infinite sequence). Thus,

$$|\lambda| = \ell(a_1) + \ell(a_2) + \cdots + \ell(a_k).$$

Comparing this with

$$\ell\left(a_1 a_2 \cdots a_k\right) = \ell\left(a_1\right) + \ell\left(a_2\right) + \cdots + \ell\left(a_k\right),$$

we obtain $\ell\left(a_1 a_2 \cdots a_k\right) = |\lambda|$. This rewrites as $\ell\left(w\right) = |\lambda|$ (since $w = a_1 a_2 \cdots a_k$). Lemma 13.192.9(a) is now proved.

(b) The words $a_1, a_2, \ldots, a_k$ are Lyndon words (since $\left(a_1, a_2, \ldots, a_k\right)$ is a tuple of Lyndon words) and thus are nonempty. Hence, their lengths $\ell\left(a_1\right), \ell\left(a_2\right), \ldots, \ell\left(a_k\right)$ are positive integers.

But recall that the partition λ is a rearrangement of the k-tuple $\left(\ell\left(a_1\right), \ell\left(a_2\right), \ldots, \ell\left(a_k\right)\right)$. Since the numbers $\ell\left(a_1\right), \ell\left(a_2\right), \ldots, \ell\left(a_k\right)$ are positive integers, we thus conclude that the partition λ has exactly k parts (i.e., exactly k nonzero entries). In other words, $\ell\left(\lambda\right) = k$. This proves Lemma 13.192.9(b). □

Proof of Proposition 6.6.48. Let us recall Definition 13.191.3 (in which we defined a nonnegative integer $\operatorname{sum} M \in \mathbb{N}$ for each finite multiset M of finite sets). Let $\mathfrak{B} = \mathfrak{A}$. Then, $\mathfrak{B}$ is a subset of $\mathfrak{A}$. Hence, Definition 13.191.5 applies.[1209] Note that $\mathfrak{B}^* = \mathfrak{A}^*$ (since $\mathfrak{B} = \mathfrak{A}$).

Let $n = |\lambda|$. Then, $n \in \mathbb{N}$. Hence, Proposition 13.191.23 yields that the maps $\operatorname{CFL} : \mathfrak{B}^n \to \mathfrak{M}\mathfrak{N}^{\mathfrak{a}}_{\mathfrak{B},n}$ and $\operatorname{LFC} : \mathfrak{M}\mathfrak{N}^{\mathfrak{a}}_{\mathfrak{B},n} \to \mathfrak{B}^n$ (introduced in Definition 13.191.10 and in Definition 13.191.21) are mutually inverse. Thus, these two maps CFL and LFC are invertible, i.e., they are bijections.

We shall now prove a sequence of simple claims:

Claim 1: Let $w \in \mathfrak{B}^n$. Then, $\operatorname{CFLtype} w = \operatorname{type}\left(\operatorname{CFL} w\right)$.

[*Proof of Claim 1:* We have $w \in \mathfrak{B}^n \subset \mathfrak{B}^* = \mathfrak{A}^*$.

Let $\left(a_1, a_2, \ldots, a_k\right)$ be the CFL factorization of w. Then, the definition of CFL yields

$$\operatorname{CFL} w = \left\{\left[a_1\right], \left[a_2\right], \ldots, \left[a_k\right]\right\}_{\text{multiset}}.$$

Hence, the sizes of the necklaces in $\operatorname{CFL} w$ are the numbers $\left|\left[a_1\right]\right|, \left|\left[a_2\right]\right|, \ldots, \left|\left[a_k\right]\right|$.

We have $\left|\left[a_i\right]\right| = \ell\left(a_i\right)$ for each $i \in \{1, 2, \ldots, k\}$. [1210]

Recall that $\left(a_1, a_2, \ldots, a_k\right)$ is the CFL factorization of w. Hence, the definition of $\operatorname{CFLtype} w$ shows that

$$\begin{aligned}&\operatorname{CFLtype} w\\&= \left(\text{the partition obtained by listing the numbers } \ell\left(a_1\right), \ell\left(a_2\right), \ldots, \ell\left(a_k\right) \text{ in decreasing order}\right).\end{aligned}$$

Comparing this with

$$\begin{aligned}&\operatorname{type}\left(\operatorname{CFL} w\right)\\&= \left(\text{the partition obtained by listing the sizes of the necklaces in } \operatorname{CFL} w \text{ in decreasing order}\right)\\&\qquad \left(\text{by the definition of } \operatorname{type}\left(\operatorname{CFL} w\right)\right)\\&= \left(\text{the partition obtained by listing the numbers } \left|\left[a_1\right]\right|, \left|\left[a_2\right]\right|, \ldots, \left|\left[a_k\right]\right| \text{ in decreasing order}\right)\\&\qquad \left(\text{since the sizes of the necklaces in } \operatorname{CFL} w \text{ are the numbers } \left|\left[a_1\right]\right|, \left|\left[a_2\right]\right|, \ldots, \left|\left[a_k\right]\right|\right)\\&= \left(\text{the partition obtained by listing the numbers } \ell\left(a_1\right), \ell\left(a_2\right), \ldots, \ell\left(a_k\right) \text{ in decreasing order}\right)\\&\qquad \left(\text{because } \left|\left[a_i\right]\right| = \ell\left(a_i\right) \text{ for each } i \in \{1, 2, \ldots, k\}\right),\end{aligned}$$

we obtain $\operatorname{CFLtype} w = \operatorname{type}\left(\operatorname{CFL} w\right)$. This proves Claim 1.]

Claim 2: We have $\{w \in \mathfrak{A}^* \mid \operatorname{CFLtype} w = \lambda\} = \{w \in \mathfrak{B}^n \mid \operatorname{CFLtype} w = \lambda\}$.

[*Proof of Claim 2:* Each $w \in \mathfrak{A}^*$ satisfying $\operatorname{CFLtype} w = \lambda$ must also satisfy $w \in \mathfrak{A}^n$ (since Lemma 13.192.9(a) yields $\ell\left(w\right) = |\lambda| = n$). Hence,

$$\{w \in \mathfrak{A}^* \mid \operatorname{CFLtype} w = \lambda\} \subset \{w \in \mathfrak{A}^n \mid \operatorname{CFLtype} w = \lambda\}.$$

On the other hand,

$$\{w \in \mathfrak{A}^n \mid \operatorname{CFLtype} w = \lambda\} \subset \{w \in \mathfrak{A}^* \mid \operatorname{CFLtype} w = \lambda\}$$

(since $\mathfrak{A}^n \subset \mathfrak{A}^*$). Combining these two inclusions, we obtain

$$\{w \in \mathfrak{A}^* \mid \operatorname{CFLtype} w = \lambda\} = \{w \in \mathfrak{A}^n \mid \operatorname{CFLtype} w = \lambda\}.$$

[1209] This is the reason why we introduced $\mathfrak{B}$ to begin with – we wanted to apply Definition 13.191.5 (and, later, Proposition 13.191.23).

[1210] Indeed, this is precisely the statement of Claim 1 **(b)** in the proof of Proposition 13.191.9. We refer to the latter proof for a proof of this statement.

Since $\mathfrak{B} = \mathfrak{A}$, this rewrites as $\{w \in \mathfrak{A}^* \mid \operatorname{CFLtype} w = \lambda\} = \{w \in \mathfrak{B}^n \mid \operatorname{CFLtype} w = \lambda\}$. This proves Claim 2.]

Claim 3: Let $w \in \mathfrak{B}^n$. Then, $\mathbf{x}_w = \mathbf{x}_{\operatorname{CFL} w}$.

[*Proof of Claim 3:* Set $M = \operatorname{CFL} w$. The definition of $\mathbf{x}_M$ yields

$$\mathbf{x}_M = \mathbf{x}_{N_1}\mathbf{x}_{N_2}\cdots\mathbf{x}_{N_k}, \tag{13.192.24}$$

where M is written in the form $M = \{N_1, N_2, \ldots, N_k\}_{\text{multiset}}$.

Let $(a_1, a_2, \ldots, a_k)$ be the CFL factorization of w. Then, the definition of CFL yields $\operatorname{CFL} w = \{[a_1], [a_2], \ldots, [a_k]\}_{\text{multiset}}$. Thus, $M = \operatorname{CFL} w = \{[a_1], [a_2], \ldots, [a_k]\}_{\text{multiset}}$. Hence, (13.192.24) (applied to $N_i = [a_i]$) yields

$$\mathbf{x}_M = \mathbf{x}_{[a_1]}\mathbf{x}_{[a_2]}\cdots\mathbf{x}_{[a_k]}. \tag{13.192.25}$$

But we know that $(a_1, a_2, \ldots, a_k)$ is a CFL factorization of w. In other words, $(a_1, a_2, \ldots, a_k)$ is a tuple of Lyndon words satisfying $w = a_1a_2\cdots a_k$ and $a_1 \geq a_2 \geq \cdots \geq a_k$ (because this is how a "CFL factorization of w" is defined). From $w = a_1a_2\cdots a_k$, we obtain

$$\mathbf{x}_w = \mathbf{x}_{a_1a_2\cdots a_k} = \mathbf{x}_{a_1}\mathbf{x}_{a_2}\cdots\mathbf{x}_{a_k} \qquad \text{(by Lemma 13.192.2)}. \tag{13.192.26}$$

On the other hand, each $i \in \{1, 2, \ldots, k\}$ satisfies $\mathbf{x}_{[a_i]} = \mathbf{x}_{a_i}$ [1211]. Hence, $\underbrace{\mathbf{x}_{[a_1]}}_{=\mathbf{x}_{a_1}}\underbrace{\mathbf{x}_{[a_2]}}_{=\mathbf{x}_{a_2}}\cdots\underbrace{\mathbf{x}_{[a_k]}}_{=\mathbf{x}_{a_k}} = \mathbf{x}_{a_1}\mathbf{x}_{a_2}\cdots\mathbf{x}_{a_k}$. This shows that the right hand sides of the equalities (13.192.25) and (13.192.26) are equal. Hence, their left hand sides are equal as well. In other words, $\mathbf{x}_M = \mathbf{x}_w$. Hence, $\mathbf{x}_w = \mathbf{x}_M = \mathbf{x}_{\operatorname{CFL} w}$ (since $M = \operatorname{CFL} w$). This proves Claim 3.]

Claim 4: Let $M \in \mathfrak{MN}^{\mathfrak{a}}$ be such that $\operatorname{type} M = \lambda$. Then, $\operatorname{sum} M = n$.

[*Proof of Claim 4:* We have $M \in \mathfrak{MN}^{\mathfrak{a}}$. In other words, M is a finite multiset of aperiodic necklaces. Let us thus write the multiset M in the form $M = \{N_1, N_2, \ldots, N_k\}_{\text{multiset}}$ for some aperiodic necklaces $N_1, N_2, \ldots, N_k$. Hence, the sizes of the necklaces are in M are the numbers $|N_1|, |N_2|, \ldots, |N_k|$.

But $\operatorname{type} M = \lambda$, so that

$$\begin{aligned}
\lambda &= \operatorname{type} M \\
&= (\text{the partition obtained by listing the sizes of the necklaces in } M \text{ in decreasing order}) \\
&\qquad (\text{by the definition of } \operatorname{type} M) \\
&= (\text{the partition obtained by listing the numbers } |N_1|, |N_2|, \ldots, |N_k| \text{ in decreasing order}) \\
&\qquad (\text{since the sizes of the necklaces in } M \text{ are the numbers } |N_1|, |N_2|, \ldots, |N_k|).
\end{aligned}$$

Hence,

$$|\lambda| = |N_1| + |N_2| + \cdots + |N_k| = \operatorname{sum} M \qquad (\text{by } (13.191.5)).$$

Hence, $n = |\lambda| = \operatorname{sum} M$. This proves Claim 4.]

Claim 5: We have

$$\left\{M \in \mathfrak{MN}^{\mathfrak{a}}_{\mathfrak{B},n} \mid \operatorname{type} M = \lambda\right\} = \left\{M \in \mathfrak{MN}^{\mathfrak{a}} \mid \operatorname{type} M = \lambda\right\}.$$

[*Proof of Claim 5:* First, we observe that $\mathfrak{B}$-necklaces and necklaces are the same thing[1212]. Moreover, $\mathfrak{MN}^{\mathfrak{a}}_{\mathfrak{B},n} \subset \mathfrak{MN}^{\mathfrak{a}}$ (by the definitions of $\mathfrak{MN}^{\mathfrak{a}}_{\mathfrak{B},n}$ and $\mathfrak{MN}^{\mathfrak{a}}$, since each $\mathfrak{B}$-necklace is a necklace).

Now, if $M \in \mathfrak{MN}^{\mathfrak{a}}$ satisfies $\operatorname{type} M = \lambda$, then $M \in \mathfrak{MN}^{\mathfrak{a}}_{\mathfrak{B},n}$ [1213]. Hence,

$$\left\{M \in \mathfrak{MN}^{\mathfrak{a}} \mid \operatorname{type} M = \lambda\right\} \subset \left\{M \in \mathfrak{MN}^{\mathfrak{a}}_{\mathfrak{B},n} \mid \operatorname{type} M = \lambda\right\}.$$

[1211] *Proof.* Let $i \in \{1, 2, \ldots, k\}$. Then, a_i is a Lyndon word (since $(a_1, a_2, \ldots, a_k)$ is a tuple of Lyndon words), and thus is nonempty (since any Lyndon word is nonempty). Hence, Lemma 13.192.8 (applied to $a = a_i$) yields $\mathbf{x}_{[a_i]} = \mathbf{x}_{a_i}$. Qed.

[1212] *Proof.* Every necklace is a subset of $\mathfrak{A}^*$. In other words, every necklace is a subset of $\mathfrak{B}^*$ (since $\mathfrak{A} = \mathfrak{B}$). Hence, every necklace is a $\mathfrak{B}$-necklace. Conversely, every $\mathfrak{B}$-necklace is a necklace. Combining the previous two sentences, we conclude that $\mathfrak{B}$-necklaces and necklaces are the same thing.

[1213] *Proof.* Let $M \in \mathfrak{MN}^{\mathfrak{a}}$ satisfy $\operatorname{type} M = \lambda$. Then, Claim 4 yields $\operatorname{sum} M = n$. Moreover, $M \in \mathfrak{MN}^{\mathfrak{a}}$ shows that M is a finite multiset of aperiodic necklaces (by the definition of $\mathfrak{MN}^{\mathfrak{a}}$). In other words, M is a finite multiset of aperiodic $\mathfrak{B}$-necklaces (since $\mathfrak{B}$-necklaces and necklaces are the same thing). Since M furthermore satisfies $\operatorname{sum} M = n$, we thus conclude that $M \in \mathfrak{MN}^{\mathfrak{a}}_{\mathfrak{B},n}$ (by the definition of $\mathfrak{MN}^{\mathfrak{a}}_{\mathfrak{B},n}$). Qed.

Combining this with

$$\left\{ M \in \underbrace{\mathfrak{MN}^{\mathfrak{a}}_{\mathfrak{B},n}}_{\subset \mathfrak{MN}^{a}} \mid \operatorname{type} M = \lambda \right\} \subset \left\{ M \in \mathfrak{MN}^{\mathfrak{a}} \mid \operatorname{type} M = \lambda \right\},$$

we obtain $\left\{ M \in \mathfrak{MN}^{\mathfrak{a}}_{\mathfrak{B},n} \mid \operatorname{type} M = \lambda \right\} = \left\{ M \in \mathfrak{MN}^{\mathfrak{a}} \mid \operatorname{type} M = \lambda \right\}$. This proves Claim 5.]

Now, Claim 2 shows that we have the following equality between summation signs:

$$\sum_{\substack{w \in \mathfrak{A}^*; \\ \operatorname{CFLtype} w = \lambda}} = \sum_{\substack{w \in \mathfrak{B}^n; \\ \operatorname{CFLtype} w = \lambda}} = \sum_{\substack{w \in \mathfrak{B}^n; \\ \operatorname{type}(\operatorname{CFL} w) = \lambda}}$$

(since Claim 1 shows that $\operatorname{CFLtype} w = \operatorname{type}(\operatorname{CFL} w)$ for each $w \in \mathfrak{B}^n$). But the definition of $\mathbf{GR}_\lambda$ yields

$$\mathbf{GR}_\lambda = \underbrace{\sum_{\substack{w \in \mathfrak{A}^*; \\ \operatorname{CFLtype} w = \lambda}}}_{= \sum_{\substack{w \in \mathfrak{B}^n; \\ \operatorname{type}(\operatorname{CFL} w) = \lambda}}} \mathbf{x}_w = \sum_{\substack{w \in \mathfrak{B}^n; \\ \operatorname{type}(\operatorname{CFL} w) = \lambda}} \underbrace{\mathbf{x}_w}_{\substack{= \mathbf{x}_{\operatorname{CFL} w} \\ \text{(by Claim 3)}}} = \sum_{\substack{w \in \mathfrak{B}^n; \\ \operatorname{type}(\operatorname{CFL} w) = \lambda}} \mathbf{x}_{\operatorname{CFL} w}$$

$$= \sum_{\substack{M \in \mathfrak{MN}^{\mathfrak{a}}_{\mathfrak{B},n}; \\ \operatorname{type} M = \lambda}} \mathbf{x}_M \qquad (13.192.27)$$

(here, we have substituted M for $\operatorname{CFL} w$ in the sum, since the map $\operatorname{CFL} : \mathfrak{B}^n \to \mathfrak{MN}^{\mathfrak{a}}_{\mathfrak{B},n}$ is a bijection). On the other hand, Claim 5 leads to the following equality between summation signs:

$$\sum_{\substack{M \in \mathfrak{MN}^{\mathfrak{a}}_{\mathfrak{B},n}; \\ \operatorname{type} M = \lambda}} = \sum_{\substack{M \in \mathfrak{MN}^{a}; \\ \operatorname{type} M = \lambda}}.$$

Using this equality, we can rewrite (13.192.27) as

$$\mathbf{GR}_\lambda = \sum_{\substack{M \in \mathfrak{MN}^{a}; \\ \operatorname{type} M = \lambda}} \mathbf{x}_M.$$

This proves Proposition 6.6.48. □

Next, we shall prove Proposition 6.6.49 using the following lemma:

Lemma 13.192.10. *Let $w \in \mathfrak{A}^*$. Then, $\mathbf{x}_{\operatorname{GR} w} = \mathbf{x}_w$.*

Proof of Lemma 13.192.10. Let $n = \ell(w)$. Thus, $w \in \mathfrak{A}^n$, so that $w = (w_1, w_2, \ldots, w_n)$. In particular, the n letters $w_1, w_2, \ldots, w_n$ are well-defined. For each $i \in \{1, 2, \ldots, n\}$, we set

$$q_i := x_{w_i}. \qquad (13.192.28)$$

(This notation will help us avoid towers of subscripts. For example, we shall soon work with expressions like q_{h_j}; without this notation, we would have to write $x_{w_{h_j}}$ for them.)

We have $w = (w_1, w_2, \ldots, w_n)$. Thus, the definition of $\mathbf{x}_w$ yields

$$\mathbf{x}_w = x_{w_1} x_{w_2} \cdots x_{w_n} = \prod_{i \in \{1,2,\ldots,n\}} \underbrace{x_{w_i}}_{\substack{= q_i \\ \text{(by (13.192.28))}}}$$

$$= \prod_{i \in \{1,2,\ldots,n\}} q_i. \qquad (13.192.29)$$

Recall that $w \in \mathfrak{A}^n$. Let τ be the permutation $(\operatorname{std} w)^{-1} \in \mathfrak{S}_n$. Then, the definition of the multiset $\operatorname{GR} w \in \mathfrak{MN}^{a}$ shows that

$$\operatorname{GR} w = \left\{ [w]_z \mid z \text{ is a cycle of } \tau \right\}_{\text{multiset}}. \qquad (13.192.30)$$

Let us now show the following two claims:

Claim 1: Let $u \in \mathfrak{A}^*$ be a word. Let $h_1, h_2, \ldots, h_k$ be some elements of $\{1, 2, \ldots, n\}$ such that $u = (w_{h_1}, w_{h_2}, \ldots, w_{h_k})$. Then,

$$\mathbf{x}_u = q_{h_1} q_{h_2} \cdots q_{h_k}.$$

[*Proof of Claim 1:* We have $u = (w_{h_1}, w_{h_2}, \ldots, w_{h_k})$. Thus, the definition of $\mathbf{x}_u$ yields

$$\mathbf{x}_u = x_{w_{h_1}} x_{w_{h_2}} \cdots x_{w_{h_k}} = \prod_{j \in \{1,2,\ldots,k\}} x_{w_{h_j}}.$$

Comparing this with

$$q_{h_1} q_{h_2} \cdots q_{h_k} = \prod_{j \in \{1,2,\ldots,k\}} \underbrace{q_{h_j}}_{\substack{= x_{w_{h_j}} \\ \text{(by (13.192.28), applied to } i = h_j)}} = \prod_{j \in \{1,2,\ldots,k\}} x_{w_{h_j}},$$

we obtain $\mathbf{x}_u = q_{h_1} q_{h_2} \cdots q_{h_k}$. This proves Claim 1.]

Claim 2: Let z be a cycle of τ. Then,

$$\mathbf{x}_{[w]_z} = \prod_{i \in z} q_i.$$

[*Proof of Claim 2:* The definition of $[w]_z$ yields $[w]_z = \{w_{\tau,i} \mid i \in z\}$. The set z is a cycle of τ, and thus is nonempty (since any cycle of τ is nonempty). Hence, there exists some $h \in z$. Consider this h. Proposition 6.6.7(a) yields $[w_{\tau,h}] = \{w_{\tau,i} \mid i \in z\}$. Comparing this with $[w]_z = \{w_{\tau,i} \mid i \in z\}$, we obtain $[w]_z = [w_{\tau,h}]$.

The definition of $w_{\tau,h}$ (given in Definition 6.6.3(b)) yields

$$w_{\tau,h} = w_{\tau^1(h)} w_{\tau^2(h)} \cdots w_{\tau^k(h)}, \qquad \text{where } k = \operatorname{ord}_\tau(h). \tag{13.192.31}$$

Let k be the size of z. Thus, $k = |z| \geq 1$ (since z is nonempty).

For each $u \in \mathbb{N}$, we set $h_u := \tau^u(h)$. Then, Lemma 13.189.1(e) (applied to $p = h$ and $p_i = h_i$ and $u = 0$) yields $h_{0+k} = h_0$. In other words, $h_k = h_0$ (since $0 + k = k$). Furthermore, Lemma 13.189.1(c) (applied to $p = h$ and $p_i = h_i$) yields $z = \{h_0, h_1, \ldots, h_{k-1}\}$. Also, Lemma 13.189.1(g) (applied to $p = h$ and $p_i = h_i$) yields that the k elements $h_0, h_1, \ldots, h_{k-1}$ are distinct. Finally, Lemma 13.189.1(i) (applied to $p = h$ and $p_i = h_i$) yields $\operatorname{ord}_\tau(h) = k$. Thus, $k = \operatorname{ord}_\tau(h)$; hence, from (13.192.31), we obtain

$$w_{\tau,h} = w_{\tau^1(h)} w_{\tau^2(h)} \cdots w_{\tau^k(h)} = \left(w_{\tau^1(h)}, w_{\tau^2(h)}, \ldots, w_{\tau^k(h)}\right) \tag{13.192.32}$$

(since $w_{\tau^1(h)}, w_{\tau^2(h)}, \ldots, w_{\tau^k(h)}$ are single letters). On the other hand, each $u \in \{1, 2, \ldots, k\}$ satisfies $w_{h_u} = w_{\tau^u(h)}$ (since $h_u = \tau^u(h)$ (by the definition of h_u)). In other words, we have

$$(w_{h_1}, w_{h_2}, \ldots, w_{h_k}) = \left(w_{\tau^1(h)}, w_{\tau^2(h)}, \ldots, w_{\tau^k(h)}\right).$$

Comparing this with (13.192.32), we obtain

$$w_{\tau,h} = (w_{h_1}, w_{h_2}, \ldots, w_{h_k}).$$

Hence, Claim 1 (applied to $u = w_{\tau,h}$) yields

$$\begin{aligned}
\mathbf{x}_{w_{\tau,h}} &= q_{h_1} q_{h_2} \cdots q_{h_k} = \left(q_{h_1} q_{h_2} \cdots q_{h_{k-1}}\right) \underbrace{q_{h_k}}_{\substack{= q_{h_0} \\ \text{(since } h_k = h_0)}} \qquad \text{(since } k \geq 1) \\
&= \left(q_{h_1} q_{h_2} \cdots q_{h_{k-1}}\right) q_{h_0} = q_{h_0} \left(q_{h_1} q_{h_2} \cdots q_{h_{k-1}}\right) \\
&= q_{h_0} q_{h_1} \cdots q_{h_{k-1}}.
\end{aligned} \tag{13.192.33}$$

But the list $(h_0, h_1, \ldots, h_{k-1})$ is a list of all elements of z without repetitions (since $z = \{h_0, h_1, \ldots, h_{k-1}\}$ and because the k elements $h_0, h_1, \ldots, h_{k-1}$ are distinct). Thus,

$$\prod_{i \in z} q_i = q_{h_0} q_{h_1} \cdots q_{h_{k-1}}.$$

Comparing this with (13.192.33), we obtain

$$\prod_{i \in z} q_i = \mathbf{x}_{w_{\tau,h}}. \tag{13.192.34}$$

The word $w_{\tau,h}$ has length k (since $w_{\tau,h} = (w_{h_1}, w_{h_2}, \ldots, w_{h_k})$), and thus is nonempty (since $k \geq 1 > 0$). Hence, Lemma 13.192.8 (applied to $a = w_{\tau,h}$) yields $\mathbf{x}_{[w_{\tau,h}]} = \mathbf{x}_{w_{\tau,h}}$. In view of $[w]_z = [w_{\tau,h}]$, this rewrites as $\mathbf{x}_{[w]_z} = \mathbf{x}_{w_{\tau,h}}$. Comparing this with (13.192.34), we obtain $\mathbf{x}_{[w]_z} = \prod_{i \in z} q_i$. This proves Claim 2.]

We know that $\operatorname{GR} w$ is a finite multiset of necklaces. Thus, the definition of the monomial $\mathbf{x}_{\operatorname{GR} w}$ shows that $\mathbf{x}_{\operatorname{GR} w} = \mathbf{x}_{N_1} \mathbf{x}_{N_2} \cdots \mathbf{x}_{N_k}$, where $\operatorname{GR} w$ is written in the form $\operatorname{GR} w = \{N_1, N_2, \ldots, N_k\}_{\text{multiset}}$. In other words, if $N_1, N_2, \ldots, N_k$ are necklaces such that $\operatorname{GR} w = \{N_1, N_2, \ldots, N_k\}_{\text{multiset}}$, then

$$\mathbf{x}_{\operatorname{GR} w} = \mathbf{x}_{N_1} \mathbf{x}_{N_2} \cdots \mathbf{x}_{N_k}. \tag{13.192.35}$$

Recall that $\tau \in \mathfrak{S}_n$. Hence, the cycles of τ are disjoint subsets of $\{1, 2, \ldots, n\}$, and each element of $\{1, 2, \ldots, n\}$ lies in exactly one of these cycles.

Let $z_1, z_2, \ldots, z_k$ be all the cycles of τ (listed without repetitions). The equality (13.192.30) becomes

$$\operatorname{GR} w = \left\{ [w]_z \ \mid \ z \text{ is a cycle of } \tau \right\}_{\text{multiset}} = \left\{ [w]_{z_1}, [w]_{z_2}, \ldots, [w]_{z_k} \right\}_{\text{multiset}}$$

(since $z_1, z_2, \ldots, z_k$ are all the cycles of τ, listed without repetitions).
Hence, (13.192.35) (applied to $N_i = [w]_{z_i}$) yields

$$\begin{aligned}
\mathbf{x}_{\operatorname{GR} w} = \mathbf{x}_{[w]_{z_1}} \mathbf{x}_{[w]_{z_2}} \cdots \mathbf{x}_{[w]_{z_k}} &= \prod_{z \text{ is a cycle of } \tau} \underbrace{\mathbf{x}_{[w]_z}}_{\substack{= \prod_{i \in z} q_i \\ \text{(by Claim 2)}}} \\
&\qquad \left(\begin{array}{c} \text{since } z_1, z_2, \ldots, z_k \text{ are all the cycles of } \tau \\ \text{(listed without repetitions)} \end{array} \right) \\
&= \underbrace{\prod_{z \text{ is a cycle of } \tau} \prod_{i \in z}}_{\substack{= \prod_{i \in \{1,2,\ldots,n\}} \\ \text{(since the cycles of } \tau \text{ are subsets of } \{1,2,\ldots,n\}, \\ \text{and since each element of } \{1,2,\ldots,n\} \text{ lies in} \\ \text{exactly one of these cycles)}}} q_i = \prod_{i \in \{1,2,\ldots,n\}} q_i = \mathbf{x}_w \qquad \text{(by (13.192.29))}.
\end{aligned}$$

This proves Lemma 13.192.10. □

Proof of Proposition 6.6.49. Theorem 6.6.29 says that the maps $\operatorname{GR} : \mathfrak{A}^* \to \mathfrak{MN}^{\mathfrak{a}}$ and $\operatorname{RG} : \mathfrak{MN}^{\mathfrak{a}} \to \mathfrak{A}^*$ are mutually inverse bijections. But Proposition 6.6.48 yields

$$\begin{aligned}
\mathbf{GR}_\lambda = \sum_{\substack{M \in \mathfrak{MN}^{\mathfrak{a}}; \\ \operatorname{type} M = \lambda}} \mathbf{x}_M &= \sum_{\substack{w \in \mathfrak{A}^*; \\ \operatorname{type}(\operatorname{GR} w) = \lambda}} \underbrace{\mathbf{x}_{\operatorname{GR} w}}_{\substack{= \mathbf{x}_w \\ \text{(by Lemma 13.192.10)}}} \\
&\qquad \left(\begin{array}{c} \text{here, we have substituted } \operatorname{GR} w \text{ for } M \text{ in the sum,} \\ \text{since the map } \operatorname{GR} : \mathfrak{A}^* \to \mathfrak{MN}^{\mathfrak{a}} \text{ is a bijection} \end{array} \right) \\
&= \sum_{\substack{w \in \mathfrak{A}^*; \\ \operatorname{type}(\operatorname{GR} w) = \lambda}} \mathbf{x}_w .
\end{aligned}$$

This proves Proposition 6.6.49. □

In order to prove Proposition 6.6.50, we will need the following two facts:

Proposition 13.192.11. *Let $n \in \mathbb{N}$. Let $\tau \in \mathfrak{S}_n$. Then,* $\operatorname{type}\left(\tau^{-1}\right) = \operatorname{type} \tau$.

Proof of Proposition 13.192.11. It is well-known that the permutations τ and τ^{-1} have the same cycle type[1214]. In other words, $\operatorname{type}\left(\tau^{-1}\right) = \operatorname{type} \tau$ (because $\operatorname{type} \sigma$ denotes the cycle type of a permutation σ). This proves Proposition 13.192.11. □

Proposition 13.192.12. *Let $w \in \mathfrak{A}^*$. Then,* $\operatorname{type}(\operatorname{GR} w) = \operatorname{type}(\operatorname{std} w)$.

[1214] In fact, they have the same cycles, if we regard the cycles of a permutation as sets.

Proof of Proposition 13.192.12. Let $n = \ell(w)$. Then, $w \in \mathfrak{A}^n$. Let τ be the permutation $(\operatorname{std} w)^{-1} \in \mathfrak{S}_n$. Then, the definition of the multiset $\operatorname{GR} w \in \mathfrak{MN}^{\mathfrak{a}}$ shows that

$$\operatorname{GR} w = \{[w]_z \mid z \text{ is a cycle of } \tau\}_{\text{multiset}} . \tag{13.192.36}$$

Let $z_1, z_2, \ldots, z_k$ be the cycles of τ, listed in some order (with no repetitions). Thus, (13.192.36) becomes

$$\begin{aligned}\operatorname{GR} w &= \{[w]_z \mid z \text{ is a cycle of } \tau\}_{\text{multiset}} \\ &= \left\{[w]_{z_1}, [w]_{z_2}, \ldots, [w]_{z_k}\right\}_{\text{multiset}} \qquad (13.192.37)\end{aligned}$$

(since $z_1, z_2, \ldots, z_k$ are the cycles of τ, listed with no repetitions).

Hence, the necklaces in $\operatorname{GR} w$ are $[w]_{z_1}, [w]_{z_2}, \ldots, [w]_{z_k}$. Therefore, the sizes of the necklaces in $\operatorname{GR} w$ are the numbers $\left|[w]_{z_1}\right|, \left|[w]_{z_2}\right|, \ldots, \left|[w]_{z_k}\right|$.

Now, it is easy to see that

$$\left|[w]_{z_m}\right| = |z_m| \qquad \text{for each } m \in \{1, 2, \ldots, k\} . \tag{13.192.38}$$

[*Proof of (13.192.38):* Let $m \in \{1, 2, \ldots, k\}$. Then, z_m is a cycle of τ (since $z_1, z_2, \ldots, z_k$ are the cycles of τ). Thus, the definition of $[w]_{z_m}$ yields $[w]_{z_m} = \{w_{\tau,i} \mid i \in z_m\}$. But Proposition 6.6.7(c) (applied to $z = z_m$) shows that $|\{w_{\tau,i} \mid i \in z_m\}| = |z_m|$. Since $[w]_{z_m} = \{w_{\tau,i} \mid i \in z_m\}$, this rewrites as $\left|[w]_{z_m}\right| = |z_m|$. This proves (13.192.38).]

But the definition of $\operatorname{type}(\operatorname{GR} w)$ yields

$$\begin{aligned}&\operatorname{type}(\operatorname{GR} w) \\ &= (\text{the partition obtained by listing the sizes of the necklaces in } \operatorname{GR} w \text{ in decreasing order}) \\ &= \left(\text{the partition obtained by listing the numbers } \left|[w]_{z_1}\right|, \left|[w]_{z_2}\right|, \ldots, \left|[w]_{z_k}\right| \text{ in decreasing order}\right) \\ &\qquad \left(\text{since the sizes of the necklaces in } \operatorname{GR} w \text{ are the numbers } \left|[w]_{z_1}\right|, \left|[w]_{z_2}\right|, \ldots, \left|[w]_{z_k}\right|\right) \\ &= (\text{the partition obtained by listing the numbers } |z_1|, |z_2|, \ldots, |z_k| \text{ in decreasing order}) \\ &\qquad (\text{by } (13.192.38)) .\end{aligned}$$

On the other hand, $\tau = (\operatorname{std} w)^{-1}$, so that $\tau^{-1} = \operatorname{std} w$. But Proposition 13.192.11 yields $\operatorname{type}\left(\tau^{-1}\right) = \operatorname{type} \tau$, so that $\operatorname{type} \tau = \operatorname{type}\left(\tau^{-1}\right) = \operatorname{type}(\operatorname{std} w)$ (since $\tau^{-1} = \operatorname{std} w$).

But $\operatorname{type} \tau$ is the cycle type of τ. In other words, $\operatorname{type} \tau$ is the partition obtained by listing the sizes of the cycles of τ in decreasing order (by the definition of the cycle type of τ). Hence,

$$\begin{aligned}&\operatorname{type} \tau \\ &= (\text{the partition obtained by listing the sizes of the cycles of } \tau \text{ in decreasing order}) \\ &= (\text{the partition obtained by listing the sizes of } z_1, z_2, \ldots, z_k \text{ in decreasing order}) \\ &\qquad (\text{since the cycles of } \tau \text{ are } z_1, z_2, \ldots, z_k \text{ (listed with no repetitions)}) \\ &= (\text{the partition obtained by listing the numbers } |z_1|, |z_2|, \ldots, |z_k| \text{ in decreasing order}) .\end{aligned}$$

Comparing this with

$$\begin{aligned}&\operatorname{type}(\operatorname{GR} w) \\ &= (\text{the partition obtained by listing the numbers } |z_1|, |z_2|, \ldots, |z_k| \text{ in decreasing order}) ,\end{aligned}$$

we obtain $\operatorname{type}(\operatorname{GR} w) = \operatorname{type} \tau = \operatorname{type}(\operatorname{std} w)$. This proves Proposition 13.192.12. □

Proof of Proposition 6.6.50. Proposition 6.6.49 yields

$$\mathbf{GR}_\lambda = \underbrace{\sum_{\substack{w \in \mathfrak{A}^*; \\ \operatorname{type}(\operatorname{GR} w) = \lambda}}}_{\substack{= \sum_{\substack{w \in \mathfrak{A}^*; \\ \operatorname{type}(\operatorname{std} w) = \lambda}} \\ \text{(since Proposition 13.192.12} \\ \text{yields } \operatorname{type}(\operatorname{GR} w) = \operatorname{type}(\operatorname{std} w) \text{ for each } w \in \mathfrak{A}^*)}} \mathbf{x}_w = \sum_{\substack{w \in \mathfrak{A}^*; \\ \operatorname{type}(\operatorname{std} w) = \lambda}} \mathbf{x}_w .$$

This proves Proposition 6.6.50. □

Next, we can easily prove Proposition 6.6.40:

Proof of Proposition 6.6.40. We have the following:

Claim 1: Let $w \in \mathfrak{A}^*$ satisfy $\operatorname{type}(\operatorname{std} w) = \lambda$. Then, $w \in \mathfrak{A}^n$.

[*Proof of Claim 1:* Let $m = \ell(w)$. Thus, $w \in \mathfrak{A}^m$, so that $\operatorname{std} w \in \mathfrak{S}_m$. But every permutation $\tau \in \mathfrak{S}_m$ satisfies $\operatorname{type} \tau \in \operatorname{Par}_m$ (since $\operatorname{type} \tau$ denotes the cycle type of τ, which is always a partition of m). Applying this to $\tau = \operatorname{std} w$, we conclude that $\operatorname{type}(\operatorname{std} w) \in \operatorname{Par}_m$. In view of $\operatorname{type}(\operatorname{std} w) = \lambda$, this rewrites as $\lambda \in \operatorname{Par}_m$. Thus, $|\lambda| = m$, so that $n = |\lambda| = m$. Hence, $m = n$. Now, $w \in \mathfrak{A}^m = \mathfrak{A}^n$ (since $m = n$). This proves Claim 1.]

Claim 2: We have $\{w \in \mathfrak{A}^* \mid \operatorname{type}(\operatorname{std} w) = \lambda\} = \{w \in \mathfrak{A}^n \mid \operatorname{type}(\operatorname{std} w) = \lambda\}$.

[*Proof of Claim 2:* Each $w \in \mathfrak{A}^*$ that satisfies $\operatorname{type}(\operatorname{std} w) = \lambda$ must also satisfy $w \in \mathfrak{A}^n$ (by Claim 1). Hence,

$$\{w \in \mathfrak{A}^* \mid \operatorname{type}(\operatorname{std} w) = \lambda\} \subset \{w \in \mathfrak{A}^n \mid \operatorname{type}(\operatorname{std} w) = \lambda\}.$$

Combining this with

$$\left\{w \in \underbrace{\mathfrak{A}^n}_{\subset \mathfrak{A}^*} \mid \operatorname{type}(\operatorname{std} w) = \lambda\right\} \subset \{w \in \mathfrak{A}^* \mid \operatorname{type}(\operatorname{std} w) = \lambda\},$$

we obtain $\{w \in \mathfrak{A}^* \mid \operatorname{type}(\operatorname{std} w) = \lambda\} = \{w \in \mathfrak{A}^n \mid \operatorname{type}(\operatorname{std} w) = \lambda\}$. This proves Claim 2.]

Claim 2 yields the following equality between summation signs:

$$\sum_{\substack{w \in \mathfrak{A}^*;\\ \operatorname{type}(\operatorname{std} w) = \lambda}} = \sum_{\substack{w \in \mathfrak{A}^n;\\ \operatorname{type}(\operatorname{std} w) = \lambda}}.$$

But Proposition 6.6.50 yields

$$\mathbf{GR}_\lambda = \underbrace{\sum_{\substack{w \in \mathfrak{A}^*;\\ \operatorname{type}(\operatorname{std} w) = \lambda}}}_{= \sum_{\substack{w \in \mathfrak{A}^n;\\ \operatorname{type}(\operatorname{std} w) = \lambda}}} \mathbf{x}_w = \underbrace{\sum_{\substack{w \in \mathfrak{A}^n;\\ \operatorname{type}(\operatorname{std} w) = \lambda}}}_{\substack{= \sum_{\substack{\sigma \in \mathfrak{S}_n;\\ \operatorname{type} \sigma = \lambda}} \sum_{\substack{w \in \mathfrak{A}^n;\\ \operatorname{std} w = \sigma}} \\ \text{(here, we have split the sum} \\ \text{according to the value of } \operatorname{std} w, \\ \text{because } \operatorname{std} w \in \mathfrak{S}_n \text{ for each } w \in \mathfrak{A}^n)}} \mathbf{x}_w$$

$$= \sum_{\substack{\sigma \in \mathfrak{S}_n;\\ \operatorname{type} \sigma = \lambda}} \sum_{\substack{w \in \mathfrak{A}^n;\\ \operatorname{std} w = \sigma}} \mathbf{x}_w = \underbrace{\sum_{\substack{\sigma \in \mathfrak{S}_n;\\ \operatorname{type}(\sigma^{-1}) = \lambda}}}_{\substack{= \sum_{\substack{\sigma \in \mathfrak{S}_n;\\ \operatorname{type} \sigma = \lambda}} \\ \text{(since each } \sigma \in \mathfrak{S}_n \text{ satisfies } \operatorname{type}(\sigma^{-1}) = \operatorname{type} \sigma \\ \text{(by Proposition 13.192.11, applied to } \tau = \sigma))}} \sum_{\substack{w \in \mathfrak{A}^n;\\ \operatorname{std} w = \sigma^{-1}}} \mathbf{x}_w$$

$$\left(\begin{array}{c}\text{here, we have substituted } \sigma^{-1} \text{ for } \sigma \text{ in the outer sum,} \\ \text{since the map } \mathfrak{S}_n \to \mathfrak{S}_n,\ \sigma \mapsto \sigma^{-1} \text{ is a bijection}\end{array}\right)$$

$$= \sum_{\substack{\sigma \in \mathfrak{S}_n;\\ \operatorname{type} \sigma = \lambda}} \sum_{\substack{w \in \mathfrak{A}^n;\\ \operatorname{std} w = \sigma^{-1}}} \mathbf{x}_w.$$

Comparing this with

$$\underbrace{\sum_{\substack{\sigma \in \mathfrak{S}_n;\\ \sigma \text{ has cycle type } \lambda}}}_{\substack{= \sum_{\substack{\sigma \in \mathfrak{S}_n;\\ \operatorname{type} \sigma = \lambda}} \\ \text{(since } \operatorname{type} \sigma \text{ means} \\ \text{the cycle type of } \sigma)}} \underbrace{L_{\gamma(\sigma)}}_{\substack{= \sum_{\substack{w \in \mathfrak{A}^n;\\ \operatorname{std} w = \sigma^{-1}}} \mathbf{x}_w \\ \text{(by Lemma 5.3.6)}}} = \sum_{\substack{\sigma \in \mathfrak{S}_n;\\ \operatorname{type} \sigma = \lambda}} \sum_{\substack{w \in \mathfrak{A}^n;\\ \operatorname{std} w = \sigma^{-1}}} \mathbf{x}_w,$$

we obtain

$$\mathbf{GR}_\lambda = \sum_{\substack{\sigma \in \mathfrak{S}_n; \\ \sigma \text{ has cycle type } \lambda}} L_{\gamma(\sigma)}.$$

This proves Proposition 6.6.40. □

Next, we can prove Proposition 6.6.42:

Proof of Proposition 6.6.42. Let $n \in \mathbb{N}$. Then, the cycle type of any permutation $\sigma \in \mathfrak{S}_n$ is a partition of n. Hence, we can split up the sum

$$\sum_{\sigma \in \mathfrak{S}_n} L_{\gamma(\sigma)}(\mathbf{x})\, p_{\operatorname{type}\sigma}(\mathbf{y})$$

according to the cycle type of σ. We thus obtain

$$\sum_{\sigma \in \mathfrak{S}_n} L_{\gamma(\sigma)}(\mathbf{x})\, p_{\operatorname{type}\sigma}(\mathbf{y}) = \sum_{\lambda \text{ is a partition of } n} \sum_{\substack{\sigma \in \mathfrak{S}_n; \\ \sigma \text{ has cycle type } \lambda}} L_{\gamma(\sigma)}(\mathbf{x})\, p_{\operatorname{type}\sigma}(\mathbf{y}). \tag{13.192.39}$$

But if λ is a partition satisfying $|\lambda| = n$, then

$$\mathbf{GR}_\lambda(\mathbf{x}) = \mathbf{GR}_\lambda = \sum_{\substack{\sigma \in \mathfrak{S}_n; \\ \sigma \text{ has cycle type } \lambda}} \underbrace{L_{\gamma(\sigma)}}_{= L_{\gamma(\sigma)}(\mathbf{x})} \qquad \text{(by Proposition 6.6.40)}$$

$$= \sum_{\substack{\sigma \in \mathfrak{S}_n; \\ \sigma \text{ has cycle type } \lambda}} L_{\gamma(\sigma)}(\mathbf{x}). \tag{13.192.40}$$

Hence, (13.192.39) becomes

$$\sum_{\sigma \in \mathfrak{S}_n} L_{\gamma(\sigma)}(\mathbf{x})\, p_{\operatorname{type}\sigma}(\mathbf{y})$$

$$= \sum_{\lambda \text{ is a partition of } n} \sum_{\substack{\sigma \in \mathfrak{S}_n; \\ \sigma \text{ has cycle type } \lambda}} L_{\gamma(\sigma)}(\mathbf{x}) \underbrace{p_{\operatorname{type}\sigma}(\mathbf{y})}_{\substack{= p_\lambda(\mathbf{y}) \\ \text{(since } \operatorname{type}\sigma = \lambda \\ \text{(since } \sigma \text{ has cycle type } \lambda\text{))}}}$$

$$= \sum_{\lambda \text{ is a partition of } n} \underbrace{\sum_{\substack{\sigma \in \mathfrak{S}_n; \\ \sigma \text{ has cycle type } \lambda}} L_{\gamma(\sigma)}(\mathbf{x})}_{\substack{= \mathbf{GR}_\lambda(\mathbf{x}) \\ \text{(by (13.192.40))}}} p_\lambda(\mathbf{y})$$

$$= \sum_{\lambda \text{ is a partition of } n} \mathbf{GR}_\lambda(\mathbf{x})\, p_\lambda(\mathbf{y}). \tag{13.192.41}$$

Now, forget that we fixed n. We thus have proved the equality (13.192.41) for each $n \in \mathbb{N}$. Now,

$$\underbrace{\sum_{\lambda \in \operatorname{Par}}}_{= \sum_{n \in \mathbb{N}} \sum_{\lambda \text{ is a partition of } n}} \mathbf{GR}_\lambda(\mathbf{x})\, p_\lambda(\mathbf{y}) = \sum_{n \in \mathbb{N}} \underbrace{\sum_{\lambda \text{ is a partition of } n} \mathbf{GR}_\lambda(\mathbf{x})\, p_\lambda(\mathbf{y})}_{\substack{= \sum_{\sigma \in \mathfrak{S}_n} L_{\gamma(\sigma)}(\mathbf{x}) p_{\operatorname{type}\sigma}(\mathbf{y}) \\ \text{(by (13.192.41))}}}$$

$$= \underbrace{\sum_{n \in \mathbb{N}} \sum_{\sigma \in \mathfrak{S}_n}}_{\substack{= \sum_{\sigma \in \bigsqcup_{n \in \mathbb{N}} \mathfrak{S}_n} = \sum_{\sigma \in \mathfrak{S}} \\ \text{(since } \bigsqcup_{n \in \mathbb{N}} \mathfrak{S}_n = \mathfrak{S}\text{)}}} L_{\gamma(\sigma)}(\mathbf{x})\, p_{\operatorname{type}\sigma}(\mathbf{y}) = \sum_{\sigma \in \mathfrak{S}} L_{\gamma(\sigma)}(\mathbf{x})\, p_{\operatorname{type}\sigma}(\mathbf{y}).$$

Combining this with the result of Proposition 6.6.38(a), we obtain

$$\sum_{\lambda\in\operatorname{Par}} \mathbf{GR}_\lambda\left(\mathbf{x}\right) p_\lambda\left(\mathbf{y}\right) = \sum_{\lambda\in\operatorname{Par}} p_\lambda\left(\mathbf{x}\right) \mathbf{GR}_\lambda\left(\mathbf{y}\right) = \sum_{\sigma\in\mathfrak{S}} L_{\gamma(\sigma)}\left(\mathbf{x}\right) p_{\operatorname{type}\sigma}\left(\mathbf{y}\right).$$

This proves Proposition 6.6.42. □

We shall prove Proposition 6.6.37 in two different ways. Both of these proofs will rely on the following simple lemma:

Lemma 13.192.13. *Let λ be a partition. Then, the power series $\mathbf{GR}_\lambda$ is homogeneous of degree $|\lambda|$.*

Proof of Lemma 13.192.13. The definition of $\mathbf{GR}_\lambda$ yields

$$\mathbf{GR}_\lambda = \sum_{\substack{w\in\mathfrak{A}^*;\\ \operatorname{CFLtype} w=\lambda}} \mathbf{x}_w. \qquad (13.192.42)$$

Set $n = |\lambda|$.

Let $w \in \mathfrak{A}^*$ be such that $\operatorname{CFLtype} w = \lambda$. Then, Lemma 13.192.9(a) yields $\ell\left(w\right) = |\lambda| = n$, so that $w \in \mathfrak{A}^n$ and thus $w = \left(w_1, w_2, \ldots, w_n\right)$. Hence, the definition of $\mathbf{x}_w$ yields $\mathbf{x}_w = x_{w_1}x_{w_2}\cdots x_{w_n}$. This shows that $\mathbf{x}_w$ is a monomial of degree n.

Forget that we fixed w. We thus have showed that $\mathbf{x}_w$ is a monomial of degree n whenever $w \in \mathfrak{A}^*$ satisfies $\operatorname{CFLtype} w = \lambda$. Hence, $\sum\limits_{\substack{w\in\mathfrak{A}^*;\\ \operatorname{CFLtype} w=\lambda}} \mathbf{x}_w$ is a sum of monomials of degree n. In other words, $\mathbf{GR}_\lambda$ is a sum of monomials of degree n (because of (13.192.42)). Thus, $\mathbf{GR}_\lambda$ is a homogeneous formal power series of degree n. In other words, $\mathbf{GR}_\lambda$ is a homogeneous formal power series of degree $|\lambda|$ (since $n = |\lambda|$). This proves Lemma 13.192.13. □

We can now step to the first proof of Proposition 6.6.37:

First proof of Proposition 6.6.37 (sketched). Lemma 13.192.13 shows that the power series $\mathbf{GR}_\lambda$ is homogeneous of degree $|\lambda|$. Hence, this power series $\mathbf{GR}_\lambda$ is of bounded degree (since any homogeneous power series is of bounded degree). In other words, $\mathbf{GR}_\lambda \in R\left(\mathbf{x}\right)$ (since $R\left(\mathbf{x}\right)$ is the set of all formal power series of bounded degree).

Recall the finitary symmetric group $\mathfrak{S}_{(\infty)}$ defined in Section 2.1; it acts on the ring $R\left(\mathbf{x}\right)$. The ring Λ is the invariant ring $\left\{f \in R\left(\mathbf{x}\right) : \sigma\left(f\right) = f \text{ for all } \sigma \in \mathfrak{S}_{(\infty)}\right\}$ of this action.

We shall now prove that $\sigma\left(f\right) = f$ for each $\sigma \in \mathfrak{S}_{(\infty)}$.

Indeed, let $\sigma \in \mathfrak{S}_{(\infty)}$. Thus, σ is a permutation of the set $\{1, 2, 3, \ldots\}$. In other words, σ is a permutation of the set $\mathfrak{A}$ (since $\mathfrak{A} = \{1, 2, 3, \ldots\}$). In other words, σ is a bijection $\mathfrak{A} \to \mathfrak{A}$.

For any word $w \in \mathfrak{A}^*$, we define a new word $\sigma_{\text{word}}\left(w\right)$ to be the result of applying σ to each letter of w. That is, if $w = \left(w_1, w_2, \ldots, w_n\right)$, then $\sigma_{\text{word}}\left(w\right) = \left(\sigma\left(w_1\right), \sigma\left(w_2\right), \ldots, \sigma\left(w_n\right)\right)$. This defines a map $\mathfrak{A}^* \to \mathfrak{A}^*,\ w \mapsto \sigma_{\text{word}}\left(w\right)$. This map is a bijection (since σ is a bijection). We shall denote this bijection by σ_{word}.

For any necklace N, we define a new necklace $\sigma_{\text{neck}}\left(N\right)$ to be the result of applying this bijection σ_{word} to every word in N. That is, $\sigma_{\text{neck}}\left(N\right) = \left\{\sigma_{\text{word}}\left(w\right) \mid w \in N\right\}$. It is straightforward to see that this set $\sigma_{\text{neck}}\left(N\right)$ is indeed a necklace[1215], and furthermore is aperiodic if and only if N is aperiodic. Thus, we obtain a map $\mathfrak{N}^{\mathfrak{a}} \to \mathfrak{N}^{\mathfrak{a}},\ N \mapsto \sigma_{\text{neck}}\left(N\right)$ (since $\mathfrak{N}^{\mathfrak{a}}$ denotes the set of all aperiodic necklaces). This map is a bijection (since σ_{word} is a bijection). We shall denote this bijection by σ_{neck}.

Finally, for any finite multiset M of aperiodic necklaces, we define a new finite multiset $\sigma_{\text{mul}}\left(M\right)$ of aperiodic necklaces to be the result of applying this bijection σ_{neck} to each necklace in M. That is, $\sigma_{\text{mul}}\left(M\right) = \left(\sigma_{\text{neck}}\right)_* M$, using the notation of Definition 13.191.16. This defines a map $\mathfrak{MN}^{\mathfrak{a}} \to \mathfrak{MN}^{\mathfrak{a}},\ M \mapsto \sigma_{\text{mul}}\left(M\right)$ (since $\mathfrak{MN}^{\mathfrak{a}}$ denotes the set of all finite multisets of aperiodic necklaces). Again, this map is a bijection (since σ_{neck} is a bijection). We shall denote this bijection by σ_{mul}.

It is easy to describe directly what this bijection σ_{mul} does to any given multiset: If $M \in \mathfrak{MN}^{\mathfrak{a}}$, then $\sigma_{\text{mul}}\left(M\right)$ is obtained from M by applying σ to each letter of each word in each necklace in M. Hence, it is

[1215] Indeed, the bijection σ_{word} commutes with the actions of C on $\mathfrak{A}^1, \mathfrak{A}^2, \mathfrak{A}^3, \ldots$.

straightforward to see that each $M \in \mathfrak{MN}^{\mathfrak{a}}$ satisfies

$$\mathbf{x}_{\sigma_{\mathrm{mul}}(M)} = \sigma\left(\mathbf{x}_M\right) \tag{13.192.43}$$

and

$$\operatorname{type}\left(\sigma_{\mathrm{mul}}\left(M\right)\right) = \operatorname{type} M. \tag{13.192.44}$$

But Proposition 6.6.48 yields

$$\begin{aligned}
\mathbf{GR}_\lambda &= \sum_{\substack{M\in\mathfrak{MN}^{\mathfrak{a}};\\ \operatorname{type} M=\lambda}} \mathbf{x}_M = \underbrace{\sum_{\substack{M\in\mathfrak{MN}^{\mathfrak{a}};\\ \operatorname{type}(\sigma_{\mathrm{mul}}(M))=\lambda}}}_{\substack{=\sum_{\substack{M\in\mathfrak{MN}^{\mathfrak{a}};\\ \operatorname{type} M=\lambda}}\\ \text{(since each } M\in\mathfrak{MN}^{\mathfrak{a}} \text{ satisfies (13.192.44))}}} \underbrace{\mathbf{x}_{\sigma_{\mathrm{mul}}(M)}}_{\substack{=\sigma(\mathbf{x}_M)\\ \text{(by (13.192.43))}}} \\
&\qquad \left(\begin{array}{c}\text{here, we have substituted } \sigma_{\mathrm{mul}}\left(M\right) \text{ for } M \text{ in the sum,}\\ \text{since the map } \sigma_{\mathrm{mul}} : \mathfrak{MN}^{\mathfrak{a}} \to \mathfrak{MN}^{\mathfrak{a}} \text{ is a bijection}\end{array}\right) \\
&= \sum_{\substack{M\in\mathfrak{MN}^{\mathfrak{a}};\\ \operatorname{type} M=\lambda}} \sigma\left(\mathbf{x}_M\right) = \sigma\left(\underbrace{\sum_{\substack{M\in\mathfrak{MN}^{\mathfrak{a}};\\ \operatorname{type} M=\lambda}} \mathbf{x}_M}_{\substack{=\mathbf{GR}_\lambda\\ \text{(by Proposition 6.6.48)}}}\right) = \sigma\left(\mathbf{GR}_\lambda\right).
\end{aligned}$$

In other words, $\sigma\left(\mathbf{GR}_\lambda\right) = \mathbf{GR}_\lambda$.

Now, forget that we fixed σ. We thus have showed that $\sigma\left(\mathbf{GR}_\lambda\right) = \mathbf{GR}_\lambda$ for all $\sigma \in \mathfrak{S}_{(\infty)}$. Hence, $\mathbf{GR}_\lambda$ is an $f \in R\left(\mathbf{x}\right)$ that satisfies $\sigma\left(f\right) = f$ for all $\sigma \in \mathfrak{S}_{(\infty)}$ (since we already know that $\mathbf{GR}_\lambda \in R\left(\mathbf{x}\right)$). In other words, $\mathbf{GR}_\lambda \in \left\{f \in R\left(\mathbf{x}\right) : \sigma\left(f\right) = f \text{ for all } \sigma \in \mathfrak{S}_{(\infty)}\right\}$. In other words, $\mathbf{GR}_\lambda \in \Lambda$ (since $\Lambda = \left\{f \in R\left(\mathbf{x}\right) : \sigma\left(f\right) = f \text{ for all } \sigma \in \mathfrak{S}_{(\infty)}\right\}$). This proves Proposition 6.6.37. □

Our next proof of Proposition 6.6.37 will rely on two lemmas:

Lemma 13.192.14. *Let λ be a partition. Then, the power series p_λ is homogeneous of degree $\left|\lambda\right|$.*

Proof of Lemma 13.192.14. Let $\ell = \ell\left(\lambda\right)$; thus, $\lambda = \left(\lambda_1, \lambda_2, \ldots, \lambda_\ell\right)$ and $p_\lambda = p_{\lambda_1} p_{\lambda_2} \cdots p_{\lambda_\ell}$ (by the definition of p_λ). From $\lambda = \left(\lambda_1, \lambda_2, \ldots, \lambda_\ell\right)$, we obtain $\left|\lambda\right| = \lambda_1 + \lambda_2 + \cdots + \lambda_\ell$.

But the power-sum symmetric functions $p_{\lambda_1}, p_{\lambda_2}, \ldots, p_{\lambda_\ell}$ are homogeneous of degrees $\lambda_1, \lambda_2, \ldots, \lambda_\ell$, respectively (since each power-sum symmetric function p_n is homogeneous of degree n). Hence, their product $p_{\lambda_1} p_{\lambda_2} \cdots p_{\lambda_\ell}$ is homogeneous of degree $\lambda_1 + \lambda_2 + \cdots + \lambda_\ell$ (since Λ is a graded $\mathbf{k}$-algebra). In view of $p_\lambda = p_{\lambda_1} p_{\lambda_2} \cdots p_{\lambda_\ell}$ and $\left|\lambda\right| = \lambda_1 + \lambda_2 + \cdots + \lambda_\ell$, this rewrites as follows:

$$p_\lambda \text{ is homogeneous of degree } \left|\lambda\right|.$$

This proves Lemma 13.192.14. □

Lemma 13.192.15. *Let A be a commutative $\mathbf{k}$-algebra. Let $\left(a_\lambda\right)_{\lambda\in\operatorname{Par}} \in \left(\mathbf{k}\left[\left[\mathbf{x}\right]\right]\right)^{\operatorname{Par}}$ and $\left(b_\lambda\right)_{\lambda\in\operatorname{Par}} \in \left(\mathbf{k}\left[\left[\mathbf{x}\right]\right]\right)^{\operatorname{Par}}$ be two families of power series in $\mathbf{k}\left[\left[\mathbf{x}\right]\right]$, and let $\left(u_\lambda\right)_{\lambda\in\operatorname{Par}} \in A^{\operatorname{Par}}$ and $\left(v_\lambda\right)_{\lambda\in\operatorname{Par}} \in A^{\operatorname{Par}}$ be two families of elements of A. Assume that the following four conditions are satisfied:*

Assumption 1: If $\lambda \in \operatorname{Par}$, then both power series a_λ and b_λ are homogeneous of degree $\left|\lambda\right|$.

Assumption 2: The family $\left(u_\lambda\right)_{\lambda\in\operatorname{Par}}$ is $\mathbf{k}$-linearly independent.

Assumption 3: We have[1216]

$$\sum_{\lambda\in\operatorname{Par}} u_\lambda a_\lambda = \sum_{\lambda\in\operatorname{Par}} v_\lambda b_\lambda \tag{13.192.45}$$

[1216] The products $u_\lambda a_\lambda$ and $v_\lambda b_\lambda$ in the following equation are taken in the power series ring $A\left[\left[\mathbf{x}\right]\right] = A\left[\left[x_1, x_2, x_3, \ldots\right]\right]$. (Indeed, u_λ and v_λ belong to A and thus to $A\left[\left[\mathbf{x}\right]\right]$ as well, whereas a_λ and b_λ are elements of $\mathbf{k}\left[\left[\mathbf{x}\right]\right]$. Hence, the products $u_\lambda a_\lambda$ and $v_\lambda b_\lambda$ make sense because $A\left[\left[\mathbf{x}\right]\right]$ is canonically a $\mathbf{k}\left[\left[\mathbf{x}\right]\right]$-module.)

[1217] *in the power series ring* $A[[\mathbf{x}]] = A[[x_1, x_2, x_3, \ldots]]$.

Assumption 4: We have $b_\lambda \in \Lambda$ *for each* $\lambda \in \operatorname{Par}$.

Then, $a_\lambda \in \Lambda$ *for each* $\lambda \in \operatorname{Par}$.

Proof of Lemma 13.192.15. In this proof, the word "monomial" always means a monomial in the indeterminates $x_1, x_2, x_3, \ldots$. Any formal power series in $\mathbf{k}[[\mathbf{x}]]$ is an infinite $\mathbf{k}$-linear combination of monomials, whereas any formal power series in $A[[\mathbf{x}]]$ is an infinite A-linear combination of monomials.

Each monomial $\mathfrak{m} = x_1^{r_1} x_2^{r_2} x_3^{r_3} \cdots$ has a well-defined degree $\deg \mathfrak{m}$, namely $\deg \mathfrak{m} = r_1 + r_2 + r_3 + \cdots$. Recall that a formal power series $f \in A[[\mathbf{x}]]$ is said to be *homogeneous of degree* n (for some $n \in \mathbb{N}$) if it is an infinite A-linear combination of monomials that have degree n. Thus, in particular, an element of A (when regarded as a formal power series in $A[[\mathbf{x}]]$) is always homogeneous of degree 0.

Consider the (unique) $\mathbf{k}$-algebra homomorphism $\iota : \mathbf{k} \to A$ (which comes from the fact that A is a $\mathbf{k}$-algebra). This homomorphism $\iota : \mathbf{k} \to A$ is injective[1218]. Hence, the canonical $\mathbf{k}[[\mathbf{x}]]$-algebra homomorphism $\iota[[\mathbf{x}]] : \mathbf{k}[[\mathbf{x}]] \to A[[\mathbf{x}]]$ induced by it (which simply applies ι to each coefficient of a power series) must also be injective. We shall thus identify $\mathbf{k}[[\mathbf{x}]]$ with a subring of $A[[\mathbf{x}]]$ (via the latter $\mathbf{k}[[\mathbf{x}]]$-algebra homomorphism $\iota[[\mathbf{x}]]$).

We shall now prove a few auxiliary claims:

Claim 1: Let $n \in \mathbb{N}$. Then,

$$\sum_{\lambda \in \operatorname{Par}_n} u_\lambda a_\lambda = \sum_{\lambda \in \operatorname{Par}_n} v_\lambda b_\lambda.$$

[*Proof of Claim 1:* Let π_n denote the map $A[[\mathbf{x}]] \to A[[\mathbf{x}]]$ that sends each formal power series $f \in A[[\mathbf{x}]]$ to its n-th homogeneous component. Explicitly, this map π_n is given by

$$\pi_n \left(\sum_{\mathfrak{m} \text{ is a monomial}} \alpha_{\mathfrak{m}} \mathfrak{m} \right) = \sum_{\substack{\mathfrak{m} \text{ is a monomial;} \\ \deg \mathfrak{m} = n}} \alpha_{\mathfrak{m}} \mathfrak{m} \qquad \text{for all } \sum_{\mathfrak{m} \text{ is a monomial}} \alpha_{\mathfrak{m}} \mathfrak{m} \in A[[\mathbf{x}]] \text{ (with } \alpha_m \in A).$$

This map π_n is A-linear and continuous (with respect to the topology on $A[[\mathbf{x}]]$); thus, it respects infinite sums. Moreover, it sends each a_λ (with $\lambda \in \operatorname{Par}$) to $\begin{cases} a_\lambda, & \text{if } |\lambda| = n; \\ 0, & \text{if } |\lambda| \neq n \end{cases}$ (since Assumption 1 shows that a_λ is

[1217] Both sums $\sum_{\lambda \in \operatorname{Par}} u_\lambda a_\lambda$ and $\sum_{\lambda \in \operatorname{Par}} v_\lambda b_\lambda$ converge (with respect to the standard topology on $A[[\mathbf{x}]]$). Here is why:

Fix $n \in \mathbb{N}$. For each $\lambda \in \operatorname{Par}$, the power series $a_\lambda \in \mathbf{k}[[\mathbf{x}]]$ is homogeneous of degree $|\lambda|$ (by Assumption 1), and thus the power series $u_\lambda a_\lambda \in A[[\mathbf{x}]]$ is also homogeneous of degree $|\lambda|$ (since the factor u_λ belongs to A and therefore does not affect the degree). Hence, for each $\lambda \in \operatorname{Par}$, the power series $u_\lambda a_\lambda$ contains no degree-n monomials unless $|\lambda| = n$. Thus, there are only finitely many $\lambda \in \operatorname{Par}$ such that the power series $u_\lambda a_\lambda$ contains degree-n monomials (because there are only finitely many $\lambda \in \operatorname{Par}$ such that $|\lambda| = n$).

Forget that we fixed n. We thus have shown that, for each $n \in \mathbb{N}$, there are only finitely many $\lambda \in \operatorname{Par}$ such that the power series $u_\lambda a_\lambda$ contains degree-n monomials. In other words, for each $n \in \mathbb{N}$, there are only finitely many addends in the sum $\sum_{\lambda \in \operatorname{Par}} u_\lambda a_\lambda$ that contain degree-n monomials. Therefore, this sum $\sum_{\lambda \in \operatorname{Par}} u_\lambda a_\lambda$ converges. Similarly, the sum $\sum_{\lambda \in \operatorname{Par}} v_\lambda b_\lambda$ converges. Qed.

[1218] *Proof.* Let $c \in \operatorname{Ker} \iota$. Thus, $c \in \mathbf{k}$ and $\iota(c) = 0$. The definition of ι yields $\iota(c) = c \cdot 1_A$. Hence, $c \cdot 1_A = \iota(c) = 0$.

Consider the empty partition $\varnothing \in \operatorname{Par}$. We have $u_\varnothing \in A$ and $c \underbrace{u_\varnothing}_{= 1_A \cdot u_\varnothing} = \underbrace{c \cdot 1_A}_{= 0} \cdot u_\varnothing = 0$.

The empty partition $\varnothing$ satisfies $\{\varnothing\} \subset \operatorname{Par}$. Hence, the family $(u_\lambda)_{\lambda \in \{\varnothing\}}$ (which consists of the single element $u_\varnothing$) is a subfamily of the family $(u_\lambda)_{\lambda \in \operatorname{Par}}$. Thus, the former family is $\mathbf{k}$-linearly independent (since Assumption 2 says that the latter family is $\mathbf{k}$-linearly independent). In other words, if $(j_\lambda)_{\lambda \in \{\varnothing\}} \in \mathbf{k}^{\{\varnothing\}}$ is a family of elements of $\mathbf{k}$ such that $\sum_{\lambda \in \{\varnothing\}} j_\lambda u_\lambda = 0$, then

$$(j_\lambda = 0 \text{ for each } \lambda \in \{\varnothing\}).$$

We can apply this to $(j_\lambda)_{\lambda \in \{\varnothing\}} = (c)_{\lambda \in \{\varnothing\}}$ (since $\sum_{\lambda \in \{\varnothing\}} c u_\lambda = c u_\varnothing = 0$), and thus conclude that

$$(c = 0 \text{ for each } \lambda \in \{\varnothing\}).$$

Applying this to $\lambda = \varnothing$, we obtain $c = 0$ (since $\varnothing \in \{\varnothing\}$).

Now, forget that we fixed c. We thus have shown that $c = 0$ for each $c \in \operatorname{Ker} \iota$. In other words, $\operatorname{Ker} \iota = 0$. Thus, ι is injective (since ι is a ring homomorphism). Qed.

homogeneous of degree $|\lambda|$), and sends each b_λ to $\begin{cases} b_\lambda, & \text{if } |\lambda| = n; \\ 0, & \text{if } |\lambda| \neq n \end{cases}$ (for similar reasons). Thus, applying π_n to both sides of (13.192.45), we obtain

$$\sum_{\lambda \in \operatorname{Par}_n} u_\lambda a_\lambda = \sum_{\lambda \in \operatorname{Par}_n} v_\lambda b_\lambda.$$

This proves Claim 1.]

Claim 2: Let $n \in \mathbb{N}$. Let $(c_\lambda)_{\lambda \in \operatorname{Par}_n} \in (\mathbf{k}[[\mathbf{x}]])^{\operatorname{Par}_n}$ be a family of elements of $\mathbf{k}[[\mathbf{x}]]$ such that

$$\sum_{\lambda \in \operatorname{Par}_n} u_\lambda c_\lambda = 0. \tag{13.192.46}$$

Then, $c_\lambda = 0$ for each $\lambda \in \operatorname{Par}_n$.

[*Proof of Claim 2:* For each $\lambda \in \operatorname{Par}_n$, let us write the formal power series $c_\lambda \in \mathbf{k}[[\mathbf{x}]]$ in the form

$$c_\lambda = \sum_{\mathfrak{m} \text{ is a monomial}} c_{\lambda,\mathfrak{m}} \mathfrak{m} \qquad \text{for some } c_{\lambda,\mathfrak{m}} \in \mathbf{k}. \tag{13.192.47}$$

Now, fix any monomial $\mathfrak{n}$. Then, (13.192.46) yields

$$0 = \sum_{\lambda \in \operatorname{Par}_n} u_\lambda \underbrace{c_\lambda}_{\substack{= \sum_{\mathfrak{m} \text{ is a monomial}} c_{\lambda,\mathfrak{m}} \mathfrak{m} \\ \text{(by (13.192.47))}}} = \sum_{\lambda \in \operatorname{Par}_n} u_\lambda \sum_{\mathfrak{m} \text{ is a monomial}} c_{\lambda,\mathfrak{m}} \mathfrak{m} = \sum_{\mathfrak{m} \text{ is a monomial}} \sum_{\lambda \in \operatorname{Par}_n} c_{\lambda,\mathfrak{m}} u_\lambda \mathfrak{m}.$$

Comparing coefficients of $\mathfrak{n}$ on both sides of this equality, we obtain $0 = \sum_{\lambda \in \operatorname{Par}_n} c_{\lambda,\mathfrak{n}} u_\lambda$. In other words,

$$\sum_{\lambda \in \operatorname{Par}_n} c_{\lambda,\mathfrak{n}} u_\lambda = 0. \tag{13.192.48}$$

But the family $(u_\lambda)_{\lambda \in \operatorname{Par}}$ is $\mathbf{k}$-linearly independent (by Assumption 2). Hence, its subfamily $(u_\lambda)_{\lambda \in \operatorname{Par}_n}$ is $\mathbf{k}$-linearly independent as well. Thus, from (13.192.48), we conclude that

$$c_{\lambda,\mathfrak{n}} = 0 \text{ for each } \lambda \in \operatorname{Par}_n \tag{13.192.49}$$

(since $c_{\lambda,\mathfrak{n}} \in \mathbf{k}$ for each $\lambda \in \operatorname{Par}_n$).

Forget that we fixed $\mathfrak{n}$. We thus have proved (13.192.49) for each monomial $\mathfrak{n}$. Now, fix $\lambda \in \operatorname{Par}_n$. Then, (13.192.47) yields

$$c_\lambda = \sum_{\mathfrak{m} \text{ is a monomial}} \underbrace{c_{\lambda,\mathfrak{m}}}_{\substack{=0 \\ \text{(by (13.192.49), applied to } \mathfrak{n}=\mathfrak{m})}} \mathfrak{m} = 0.$$

This proves Claim 2.]

Recall the finitary symmetric group $\mathfrak{S}_{(\infty)}$ defined in Section 2.1. It acts on the rings $R(\mathbf{x})$ and $\mathbf{k}[[\mathbf{x}]]$ and $A[[\mathbf{x}]]$ in the same way (viz., by permuting the variables $x_1, x_2, x_3, \ldots$). The ring Λ is the invariant ring $\left\{ f \in R(\mathbf{x}) : \sigma(f) = f \text{ for all } \sigma \in \mathfrak{S}_{(\infty)} \right\}$ of its action on $R(\mathbf{x})$. Of course, $R(\mathbf{x})$ is an $\mathfrak{S}_{(\infty)}$-subset of $\mathbf{k}[[\mathbf{x}]]$. Moreover, the canonical $\mathbf{k}[[\mathbf{x}]]$-algebra homomorphism $\iota[[\mathbf{x}]] : \mathbf{k}[[\mathbf{x}]] \to A[[\mathbf{x}]]$ (which we are using to identify $\mathbf{k}[[\mathbf{x}]]$ with a subring of $A[[\mathbf{x}]]$) is $\mathfrak{S}_{(\infty)}$-equivariant. This ensures that expressions like $\sigma(a_\lambda)$ (for $\sigma \in \mathfrak{S}_{(\infty)}$ and $\lambda \in \operatorname{Par}$) are well-defined (i.e., they don't depend on whether we regard a_λ as an element of $\mathbf{k}[[\mathbf{x}]]$ or as an element of $A[[\mathbf{x}]]$).

We now claim:

Claim 3: Let $n \in \mathbb{N}$. Let $\lambda \in \operatorname{Par}_n$. Then, $\sigma(b_\lambda) = b_\lambda$ for all $\sigma \in \mathfrak{S}_{(\infty)}$.

[*Proof of Claim 3:* Assumption 4 yields $b_\lambda \in \Lambda = \left\{ f \in R(\mathbf{x}) : \sigma(f) = f \text{ for all } \sigma \in \mathfrak{S}_{(\infty)} \right\}$. Hence, $\sigma(b_\lambda) = b_\lambda$ for all $\sigma \in \mathfrak{S}_{(\infty)}$. This proves Claim 3.]

Claim 4: Let $\sigma \in \mathfrak{S}_{(\infty)}$. Then, $\sigma(a_\mu) = a_\mu$ for each $\mu \in \operatorname{Par}$.

[*Proof of Claim 4:* Let $\mu \in \operatorname{Par}$. Set $n = |\mu|$. Then, $\mu \in \operatorname{Par}_n$ (since $\mu \in \operatorname{Par}$ and $|\mu| = n$).

Recall that $(u_\lambda)_{\lambda \in \operatorname{Par}} \in A^{\operatorname{Par}}$. Hence, $u_\lambda \in A$ for each $\lambda \in \operatorname{Par}_n$. Likewise, $v_\lambda \in A$ for each $\lambda \in \operatorname{Par}_n$. Also, recall that $(a_\lambda)_{\lambda \in \operatorname{Par}} \in (\mathbf{k}[[\mathbf{x}]])^{\operatorname{Par}}$. Hence, each $\lambda \in \operatorname{Par}_n$ satisfies $a_\lambda \in \mathbf{k}[[\mathbf{x}]]$ and thus $\sigma(a_\lambda) - a_\lambda \in \mathbf{k}[[\mathbf{x}]]$. In other words, $(\sigma(a_\lambda) - a_\lambda)_{\lambda \in \operatorname{Par}_n} \in (\mathbf{k}[[\mathbf{x}]])^{\operatorname{Par}_n}$.

Claim 1 yields

$$\sum_{\lambda\in\operatorname{Par}_n} u_\lambda a_\lambda = \sum_{\lambda\in\operatorname{Par}_n} v_\lambda b_\lambda . \tag{13.192.50}$$

Note that we can regard the (finite) sum $\sum_{\lambda\in\operatorname{Par}_n} u_\lambda a_\lambda$ as an A-linear combination of the elements $a_\lambda \in A[[\mathbf{x}]]$ [1219] with coefficients $u_\lambda \in A$. Likewise, we can regard the (finite) sum $\sum_{\lambda\in\operatorname{Par}_n} v_\lambda b_\lambda$ as an A-linear combination of the elements $b_\lambda \in A[[\mathbf{x}]]$ with coefficients $v_\lambda \in A$.

But the group $\mathfrak{S}_{(\infty)}$ acts on $A[[\mathbf{x}]]$ by A-linear maps. Thus, the map $A[[\mathbf{x}]] \to A[[\mathbf{x}]]$, $f \mapsto \sigma(f)$ is A-linear. Hence, this map respects A-linear combinations. Consequently,

$$\sigma\left(\sum_{\lambda\in\operatorname{Par}_n} u_\lambda a_\lambda\right) = \sum_{\lambda\in\operatorname{Par}_n} u_\lambda \sigma(a_\lambda) \qquad \text{(since } u_\lambda \in A \text{ for each } \lambda \in \operatorname{Par}_n\text{)}$$

and

$$\begin{aligned}\sigma\left(\sum_{\lambda\in\operatorname{Par}_n} v_\lambda b_\lambda\right) &= \sum_{\lambda\in\operatorname{Par}_n} v_\lambda \underbrace{\sigma(b_\lambda)}_{\substack{=b_\lambda\\ \text{(by Claim 3)}}} \qquad \text{(since } v_\lambda \in A \text{ for each } \lambda \in \operatorname{Par}_n\text{)}\\ &= \sum_{\lambda\in\operatorname{Par}_n} v_\lambda b_\lambda = \sum_{\lambda\in\operatorname{Par}_n} u_\lambda a_\lambda \qquad \text{(by (13.192.50))}.\end{aligned}$$

Subtracting these two equalities, we find

$$\sigma\left(\sum_{\lambda\in\operatorname{Par}_n} u_\lambda a_\lambda\right) - \sigma\left(\sum_{\lambda\in\operatorname{Par}_n} v_\lambda b_\lambda\right) = \sum_{\lambda\in\operatorname{Par}_n} u_\lambda \sigma(a_\lambda) - \sum_{\lambda\in\operatorname{Par}_n} u_\lambda a_\lambda = \sum_{\lambda\in\operatorname{Par}_n} u_\lambda (\sigma(a_\lambda) - a_\lambda).$$

Hence,

$$\begin{aligned}\sum_{\lambda\in\operatorname{Par}_n} u_\lambda (\sigma(a_\lambda) - a_\lambda) &= \sigma\left(\underbrace{\sum_{\lambda\in\operatorname{Par}_n} u_\lambda a_\lambda}_{\substack{=\sum_{\lambda\in\operatorname{Par}_n} v_\lambda b_\lambda\\ \text{(by Claim 1)}}}\right) - \sigma\left(\sum_{\lambda\in\operatorname{Par}_n} v_\lambda b_\lambda\right)\\ &= \sigma\left(\sum_{\lambda\in\operatorname{Par}_n} v_\lambda b_\lambda\right) - \sigma\left(\sum_{\lambda\in\operatorname{Par}_n} v_\lambda b_\lambda\right) = 0.\end{aligned}$$

Hence, Claim 2 (applied to $c_\lambda = \sigma(a_\lambda) - a_\lambda$) yields that $\sigma(a_\lambda) - a_\lambda = 0$ for each $\lambda \in \operatorname{Par}_n$ (since $(\sigma(a_\lambda) - a_\lambda)_{\lambda\in\operatorname{Par}_n} \in (\mathbf{k}[[\mathbf{x}]])^{\operatorname{Par}_n}$). Applying this to $\lambda = \mu$, we obtain $\sigma(a_\mu) - a_\mu = 0$ (since $\mu \in \operatorname{Par}_n$). In other words, $\sigma(a_\mu) = a_\mu$. This proves Claim 4.]

Claim 5: We have $a_\mu \in \Lambda$ for each $\mu \in \operatorname{Par}$.

[*Proof of Claim 5:* Let $\mu \in \operatorname{Par}$. Thus, $a_\mu \in \mathbf{k}[[\mathbf{x}]]$ (since $(a_\lambda)_{\lambda\in\operatorname{Par}} \in (\mathbf{k}[[\mathbf{x}]])^{\operatorname{Par}}$). Also, Assumption 1 (applied to $\lambda = \mu$) shows that the power series a_μ is homogeneous of degree $|\mu|$. Hence, this power series a_μ is of bounded degree (since any homogeneous power series is of bounded degree). In other words, $a_\mu \in R(\mathbf{x})$. Claim 4 yields that $\sigma(a_\mu) = a_\mu$ for all $\sigma \in \mathfrak{S}_{(\infty)}$.

Thus, a_μ is an $f \in R(\mathbf{x})$ satisfying $\sigma(f) = f$ for all $\sigma \in \mathfrak{S}_{(\infty)}$. In other words,

$$a_\mu \in \left\{f \in R(\mathbf{x}) : \sigma(f) = f \text{ for all } \sigma \in \mathfrak{S}_{(\infty)}\right\}.$$

In view of $\Lambda = \left\{f \in R(\mathbf{x}) : \sigma(f) = f \text{ for all } \sigma \in \mathfrak{S}_{(\infty)}\right\}$, this rewrites as $a_\mu \in \Lambda$. This proves Claim 5.]

We have $a_\mu \in \Lambda$ for each $\mu \in \operatorname{Par}$ (by Claim 5). Renaming the index μ as λ in this statement, we can rewrite this as follows: We have $a_\lambda \in \Lambda$ for each $\lambda \in \operatorname{Par}$. This proves Lemma 13.192.15. □

[1219] Here, we are using the fact that the a_λ (for $\lambda \in \operatorname{Par}$) are elements of $A[[\mathbf{x}]]$. (This is because they are elements of $\mathbf{k}[[\mathbf{x}]]$, but we are identifying $\mathbf{k}[[\mathbf{x}]]$ with a subring of $A[[\mathbf{x}]]$.)

Second proof of Proposition 6.6.37. We begin with a trick. Recall that there is a canonical ring homomorphism $\varphi : \mathbb{Z} \to \mathbf{k}$. This homomorphism gives rise to a ring homomorphism $\varphi[[\mathbf{x}]] : \mathbb{Z}[[\mathbf{x}]] \to \mathbf{k}[[\mathbf{x}]]$, and this latter homomorphism $\varphi[[\mathbf{x}]]$ sends $\Lambda_{\mathbb{Z}}$ to $\Lambda_{\mathbf{k}}$; that is, we have $(\varphi[[\mathbf{x}]])(\Lambda_{\mathbb{Z}}) \subset \Lambda_{\mathbf{k}}$. Moreover, it is clear that the ring homomorphism $\varphi[[\mathbf{x}]]$ sends the element $\mathbf{GR}_\lambda$ of $\mathbb{Z}[[\mathbf{x}]]$ to the element $\mathbf{GR}_\lambda$ of $\mathbf{k}[[\mathbf{x}]]$ (because the definition of $\mathbf{GR}_\lambda$ is functorial in the base ring $\mathbf{k}$). Therefore, if we can prove that the element $\mathbf{GR}_\lambda$ of $\mathbb{Z}[[\mathbf{x}]]$ belongs to $\Lambda_{\mathbb{Z}}$, then it will automatically follow that the element $\mathbf{GR}_\lambda$ of $\mathbf{k}[[\mathbf{x}]]$ belongs to $(\varphi[[\mathbf{x}]])(\Lambda_{\mathbb{Z}}) \subset \Lambda_{\mathbf{k}}$; this will complete the proof of Proposition 6.6.37. Hence, in order to prove Proposition 6.6.37, it only remains to prove that the element $\mathbf{GR}_\lambda$ of $\mathbb{Z}[[\mathbf{x}]]$ belongs to $\Lambda_{\mathbb{Z}}$. In other words, it only remains to prove Proposition 6.6.37 in the case of $\mathbf{k} = \mathbb{Z}$. Hence, in proving Proposition 6.6.37, we can WLOG assume that $\mathbf{k} = \mathbb{Z}$. Assume this.

Forget that we fixed λ. Consider the countable set of indeterminates $\mathbf{y} = (y_1, y_2, y_3, \ldots)$.

It is easy to see that the family $(p_\lambda)_{\lambda\in\mathrm{Par}} \in (\mathbf{k}[[\mathbf{x}]])^{\mathrm{Par}}$ is $\mathbf{k}$-linearly independent[1220]. Hence, the family $(p_\lambda(\mathbf{y}))_{\lambda\in\mathrm{Par}} \in (\mathbf{k}[[\mathbf{y}]])^{\mathrm{Par}}$ is $\mathbf{k}$-linearly independent[1221]. If $\lambda \in \mathrm{Par}$, then both power series $\mathbf{GR}_\lambda$ and p_λ are homogeneous of degree $|\lambda|$ (by Lemma 13.192.13 and Lemma 13.192.14).

We identify the power series ring $\mathbf{k}[[\mathbf{x},\mathbf{y}]] = \mathbf{k}[[x_1, x_2, x_3, \ldots, y_1, y_2, y_3, \ldots]]$ with the power series ring $(\mathbf{k}[[\mathbf{y}]])[[\mathbf{x}]] = (\mathbf{k}[[y_1, y_2, y_3, \ldots]])[[x_1, x_2, x_3, \ldots]]$. In this power series ring, we have

$$\begin{aligned}\sum_{\lambda\in\mathrm{Par}} p_\lambda(\mathbf{y}) \underbrace{\mathbf{GR}_\lambda}_{=\mathbf{GR}_\lambda(\mathbf{x})} &= \sum_{\lambda\in\mathrm{Par}} p_\lambda(\mathbf{y})\,\mathbf{GR}_\lambda(\mathbf{x}) = \sum_{\lambda\in\mathrm{Par}} \mathbf{GR}_\lambda(\mathbf{x})\, p_\lambda(\mathbf{y}) \\ &= \sum_{\lambda\in\mathrm{Par}} p_\lambda(\mathbf{x})\,\mathbf{GR}_\lambda(\mathbf{y}) \qquad \text{(by Proposition 6.6.38(a))} \\ &= \sum_{\lambda\in\mathrm{Par}} \mathbf{GR}_\lambda(\mathbf{y}) \underbrace{p_\lambda(\mathbf{x})}_{=p_\lambda} = \sum_{\lambda\in\mathrm{Par}} \mathbf{GR}_\lambda(\mathbf{y})\, p_\lambda. \qquad (13.192.51)\end{aligned}$$

Hence, we can apply Lemma 13.192.15 to

$$A = \mathbf{k}[[\mathbf{y}]], \quad a_\lambda = \mathbf{GR}_\lambda, \quad b_\lambda = p_\lambda, \quad u_\lambda = p_\lambda(\mathbf{y}), \quad \text{and} \quad v_\lambda = \mathbf{GR}_\lambda(\mathbf{y}).$$

[1222] Thus, we conclude that $\mathbf{GR}_\lambda \in \Lambda$ for each $\lambda \in \mathrm{Par}$. In other words, the power series $\mathbf{GR}_\lambda$ belongs to Λ for each partition λ. This proves Proposition 6.6.37. □

Finally, let us prove Proposition 6.6.36:

Proof of Proposition 6.6.36. (a) We proceed by showing some auxiliary claims:

> *Claim 1:* Let $w \in \mathfrak{A}^*$ satisfy $\operatorname{CFLtype} w = (n)$. Then, the word w is Lyndon and satisfies $w \in \mathfrak{A}^n$.

[*Proof of Claim 1:* Lemma 13.192.9(a) (applied to $\lambda = (n)$) yields $\ell(w) = |(n)| = n$. Hence, $w \in \mathfrak{A}^n$.

Let $(a_1, a_2, \ldots, a_k)$ be the CFL factorization of w. Thus, $(a_1, a_2, \ldots, a_k)$ is a tuple of Lyndon words satisfying $w = a_1 a_2 \cdots a_k$ and $a_1 \geq a_2 \geq \cdots \geq a_k$ (by the definition of "CFL factorization"). But Lemma 13.192.9(b) (applied to $\lambda = (n)$) yields $k = \ell((n)) = 1$ (since n is a positive integer). Hence, $a_1 a_2 \cdots a_k = a_1$. Thus, $w = a_1 a_2 \cdots a_k = a_1$. But a_1 is a Lyndon word (since $(a_1, a_2, \ldots, a_k)$ is a tuple of Lyndon words). In other words, w is a Lyndon word (since $w = a_1$). Thus, Claim 1 is proven.]

> *Claim 2:* Let $w \in \mathfrak{A}^n$ be a Lyndon word. Then, $w \in \mathfrak{A}^*$ and $\operatorname{CFLtype} w = (n)$.

[1220] *Proof.* We have $\Lambda = \Lambda_{\mathbf{k}} = \Lambda_{\mathbb{Z}}$ (since $\mathbf{k} = \mathbb{Z}$). But $\Lambda_{\mathbb{Z}}$ is clearly a subring of $\Lambda_{\mathbb{Q}}$ (since the ring $\mathbb{Z}[[\mathbf{x}]]$ is a subring of the ring $\mathbb{Q}[[\mathbf{x}]]$, and the finitary symmetric group $\mathfrak{S}_{(\infty)}$ acts in the same way on both of these rings). Furthermore, the symmetric functions e_λ, s_λ and p_λ are defined in the same way over the ring $\mathbb{Z}$ as they are defined over the ring $\mathbb{Q}$.

Now, Proposition 2.2.10 (applied to $\mathbb{Q}$ instead of $\mathbf{k}$) shows that the families $(e_\lambda)_{\lambda\in\mathrm{Par}}$, $(s_\lambda)_{\lambda\in\mathrm{Par}}$ and $(p_\lambda)_{\lambda\in\mathrm{Par}}$ are bases of the $\mathbb{Q}$-module $\Lambda_{\mathbb{Q}}$. Hence, in particular, the family $(p_\lambda)_{\lambda\in\mathrm{Par}}$ is a basis of the $\mathbb{Q}$-module $\Lambda_{\mathbb{Q}}$. Thus, this family is $\mathbb{Q}$-linearly independent, and therefore $\mathbb{Z}$-linearly independent as well (since $\mathbb{Z}$ is a subring of $\mathbb{Q}$). In other words, it is $\mathbf{k}$-linearly independent (since $\mathbf{k} = \mathbb{Z}$). Qed.

[1221] Indeed, the family $(p_\lambda(\mathbf{y}))_{\lambda\in\mathrm{Par}}$ is obtained from the family $(p_\lambda)_{\lambda\in\mathrm{Par}}$ by renaming the indeterminates $x_1, x_2, x_3, \ldots$ as $y_1, y_2, y_3, \ldots$. Thus, the $\mathbf{k}$-linear independence of the latter family implies the $\mathbf{k}$-linear independence of the former.

[1222] Indeed, Assumption 1 of Lemma 13.192.15 is satisfied (because if $\lambda \in \mathrm{Par}$, then both power series $\mathbf{GR}_\lambda$ and p_λ are homogeneous of degree $|\lambda|$); Assumption 2 is satisfied as well (since the family $(p_\lambda(\mathbf{y}))_{\lambda\in\mathrm{Par}} \in (\mathbf{k}[[\mathbf{y}]])^{\mathrm{Par}}$ is $\mathbf{k}$-linearly independent); Assumption 3 is also satisfied (due to (13.192.51)); finally, Assumption 4 is satisfied (since we have $p_\lambda \in \Lambda$ for each $\lambda \in \mathrm{Par}$).

[*Proof of Claim 2:* We have $w \in \mathfrak{A}^n \subset \mathfrak{A}^*$. It remains to prove that $\operatorname{CFLtype} w = (n)$.

The word w is Lyndon. Hence, the definition of a "CFL factorization" shows that the 1-tuple (w) is a CFL factorization of w. We can replace "a CFL factorization" by "the CFL factorization" in this sentence (since Theorem 6.1.27 shows that the CFL factorization of w is unique). Thus, we conclude that the 1-tuple (w) is the CFL factorization of w. Hence, the definition of $\operatorname{CFLtype} w$ shows that

$$\operatorname{CFLtype} w = (\ell(w)) = (n) \qquad \left(\text{since } \ell(w) = n \text{ (because } w \in \mathfrak{A}^n)\right).$$

This concludes the proof of Claim 2.]

Claim 3: We have $\{w \in \mathfrak{A}^* \mid \operatorname{CFLtype} w = (n)\} = \{w \in \mathfrak{A}^n \mid w \text{ is Lyndon}\}$.

[*Proof of Claim 3:* Claim 1 shows that $\{w \in \mathfrak{A}^* \mid \operatorname{CFLtype} w = (n)\} \subset \{w \in \mathfrak{A}^n \mid w \text{ is Lyndon}\}$. Claim 2 shows that $\{w \in \mathfrak{A}^n \mid w \text{ is Lyndon}\} \subset \{w \in \mathfrak{A}^* \mid \operatorname{CFLtype} w = (n)\}$. Combining these two inclusions, we obtain precisely the equality claimed in Claim 3.]

The definition of $\mathbf{GR}_{(n)}$ yields

$$\mathbf{GR}_{(n)} = \underbrace{\sum_{\substack{w \in \mathfrak{A}^*; \\ \operatorname{CFLtype} w = (n)}}}_{\substack{= \sum_{\substack{w \in \mathfrak{A}^n; \\ w \text{ is Lyndon}}} \\ \text{(by Claim 3)}}} \mathbf{x}_w = \sum_{\substack{w \in \mathfrak{A}^n; \\ w \text{ is Lyndon}}} \mathbf{x}_w.$$

This proves Proposition 6.6.36(a).

(b) Forget that we fixed n. We shall first prove several auxiliary claims:

Claim 4: Let d be a positive integer. Then,

$$\{w \in \mathfrak{L} \mid \ell(w) = d\} = \left\{w \in \mathfrak{A}^d \mid w \text{ is Lyndon}\right\}.$$

[*Proof of Claim 4:* Since $\mathfrak{L}$ is the set of all Lyndon words in $\mathfrak{A}^*$, we have

$$\begin{aligned}\{w \in \mathfrak{L} \mid \ell(w) = d\} &= \{w \text{ is a Lyndon word in } \mathfrak{A}^* \mid \ell(w) = d\} \\ &= \{w \in \mathfrak{A}^* \mid w \text{ is Lyndon, and } \ell(w) = d\}.\end{aligned}$$

But since $\mathfrak{A}^d$ is the set of words $w \in \mathfrak{A}^*$ satisfying $\ell(w) = d$, we have

$$\begin{aligned}\left\{w \in \mathfrak{A}^d \mid w \text{ is Lyndon}\right\} &= \{w \in \mathfrak{A}^* \text{ satisfying } \ell(w) = d \mid w \text{ is Lyndon}\} \\ &= \{w \in \mathfrak{A}^* \mid w \text{ is Lyndon, and } \ell(w) = d\}.\end{aligned}$$

Comparing these two equalities, we obtain $\{w \in \mathfrak{L} \mid \ell(w) = d\} = \left\{w \in \mathfrak{A}^d \mid w \text{ is Lyndon}\right\}$. This proves Claim 4.]

Claim 5: Let $w \in \mathfrak{L}$. Then, $\ell(w) \in \{1, 2, 3, \ldots\}$.

[*Proof of Claim 5:* From $w \in \mathfrak{L}$, we conclude that w is a Lyndon word in $\mathfrak{A}^*$ (since $\mathfrak{L}$ is the set of Lyndon words in $\mathfrak{A}^*$). Hence, the word w is Lyndon, and thus nonempty. Therefore, $\ell(w) \in \{1, 2, 3, \ldots\}$. This proves Claim 5.]

The next two claims concern the results of substituting $x_1^k, x_2^k, x_3^k, \ldots$ for $x_1, x_2, x_3, \ldots$ in some power series $f \in \mathbf{k}[[\mathbf{x}]]$ (where k is a given positive integer). Recall that the result of substituting $x_1^k, x_2^k, x_3^k, \ldots$ for $x_1, x_2, x_3, \ldots$ in a power series $f \in \mathbf{k}[[\mathbf{x}]]$ is denoted by $f\left(x_1^k, x_2^k, x_3^k, \ldots\right)$.

Claim 6: Let k be a positive integer. Let $w \in \mathfrak{A}^*$. Then,

$$\mathbf{x}_w\left(x_1^k, x_2^k, x_3^k, \ldots\right) = \mathbf{x}_w^k$$

(where, of course, $\mathbf{x}_w\left(x_1^k, x_2^k, x_3^k, \ldots\right)$ denotes the result of substituting $x_1^k, x_2^k, x_3^k, \ldots$ for $x_1, x_2, x_3, \ldots$ in the polynomial $\mathbf{x}_w$).

[*Proof of Claim 6:* Write the word w in the form $w = (w_1, w_2, \ldots, w_n)$ (for some $n \in \mathbb{N}$). Then, $\mathbf{x}_w = x_{w_1} x_{w_2} \cdots x_{w_n}$ (by the definition of $\mathbf{x}_w$). Substituting $x_1^k, x_2^k, x_3^k, \ldots$ for $x_1, x_2, x_3, \ldots$ on both sides of this equality, we obtain

$$\mathbf{x}_w\left(x_1^k, x_2^k, x_3^k, \ldots\right) = x_{w_1}^k x_{w_2}^k \cdots x_{w_n}^k = \left(\underbrace{x_{w_1} x_{w_2} \cdots x_{w_n}}_{=\mathbf{x}_w}\right)^k = \mathbf{x}_w^k.$$

This proves Claim 6.]

Claim 7: Let k and m be positive integers. Then,

$$p_m\left(x_1^k, x_2^k, x_3^k, \ldots\right) = p_{km}$$

(where, of course, $p_m\left(x_1^k, x_2^k, x_3^k, \ldots\right)$ denotes the result of substituting $x_1^k, x_2^k, x_3^k, \ldots$ for $x_1, x_2, x_3, \ldots$ in the power series p_m).

[*Proof of Claim 7:* The definition of p_{km} yields $p_{km} = x_1^{km} + x_2^{km} + x_3^{km} + \cdots$.

On the other hand, the definition of p_m yields $p_m = x_1^m + x_2^m + x_3^m + \cdots = \sum_{i\geq 1} x_i^m$. Substituting $x_1^k, x_2^k, x_3^k, \ldots$ for $x_1, x_2, x_3, \ldots$ on both sides of this equality, we obtain

$$p_m\left(x_1^k, x_2^k, x_3^k, \ldots\right) = \sum_{i\geq 1} \underbrace{\left(x_i^k\right)^m}_{=x_i^{km}} = \sum_{i\geq 1} x_i^{km} = x_1^{km} + x_2^{km} + x_3^{km} + \cdots = p_{km}$$

(since $p_{km} = x_1^{km} + x_2^{km} + x_3^{km} + \cdots$). This proves Claim 7.]

Claim 8: Let k and d be positive integers. Then,

$$\mathbf{GR}_{(d)}\left(x_1^k, x_2^k, x_3^k, \ldots\right) = \sum_{\substack{w\in\mathfrak{L};\\ \ell(w)=d}} \mathbf{x}_w^k. \tag{13.192.52}$$

[*Proof of Claim 8:* Because of Claim 4, we have the following equality of summation signs:

$$\sum_{\substack{w\in\mathfrak{L};\\ \ell(w)=d}} = \sum_{\substack{w\in\mathfrak{A}^d;\\ w \text{ is Lyndon}}} . \tag{13.192.53}$$

But Proposition 6.6.36(a) (applied to $n = d$) yields

$$\mathbf{GR}_{(d)} = \underbrace{\sum_{\substack{w\in\mathfrak{A}^d;\\ w \text{ is Lyndon}}}}_{\substack{= \sum_{\substack{w\in\mathfrak{L};\\ \ell(w)=d}} \\ (\text{by } (13.192.53))}} \mathbf{x}_w = \sum_{\substack{w\in\mathfrak{L};\\ \ell(w)=d}} \mathbf{x}_w.$$

Substituting $x_1^k, x_2^k, x_3^k, \ldots$ for $x_1, x_2, x_3, \ldots$ on both sides of this equality, we obtain

$$\mathbf{GR}_{(d)}\left(x_1^k, x_2^k, x_3^k, \ldots\right) = \sum_{\substack{w\in\mathfrak{L};\\ \ell(w)=d}} \underbrace{\mathbf{x}_w\left(x_1^k, x_2^k, x_3^k, \ldots\right)}_{\substack{=\mathbf{x}_w^k \\ (\text{by Claim 6})}} = \sum_{\substack{w\in\mathfrak{L};\\ \ell(w)=d}} \mathbf{x}_w^k.$$

This proves Claim 8.]

For each positive integer n, we define a symmetric function $\widetilde{\mathbf{GR}}_n \in \Lambda$ by

$$\widetilde{\mathbf{GR}}_n = \frac{1}{n}\sum_{d\mid n} \mu(d)\, p_d^{n/d}.$$

Our goal will be to prove that $\mathbf{GR}_{(n)} = \widetilde{\mathbf{GR}}_n$ for each positive integer n.

The following two claims are key to our proof:

Claim 9: Every positive integer n satisfies

$$p_1^n = \sum_{d\mid n} d \cdot \mathbf{GR}_{(d)}\left(x_1^{n/d}, x_2^{n/d}, x_3^{n/d}, \ldots\right).$$

[*Proof of Claim 9:* It is well-known that

$$-\log\left(1-t\right)=\sum_{k\geq1}\frac{1}{k}t^{k} \tag{13.192.54}$$

in the ring $\mathbf{k}\left[\left[t\right]\right]$. (Indeed, this is the well-known Mercator series for the logarithm.)

Proposition 6.6.39 yields the equality

$$\frac{1}{1-p_1 t}=\prod_{w\in\mathfrak{L}}\frac{1}{1-\mathbf{x}_w t^{\ell(w)}}$$

in the power series ring $\left(\mathbf{k}\left[\left[\mathbf{x}\right]\right]\right)\left[\left[t\right]\right]$. Taking the logarithm of both sides of this identity, we obtain

$$\begin{aligned}
\log\frac{1}{1-p_1 t} &= \log\left(\prod_{w\in\mathfrak{L}}\frac{1}{1-\mathbf{x}_w t^{\ell(w)}}\right) = \underbrace{\sum_{w\in\mathfrak{L}}}_{\substack{=\sum_{d\geq1}\sum_{w\in\mathfrak{L};\ \ell(w)=d} \\ \text{(since each } w\in\mathfrak{L}\text{ satisfies } \ell(w)\in\{1,2,3,\ldots\} \\ \text{(by Claim 5))}}} \log\left(\frac{1}{1-\mathbf{x}_w t^{\ell(w)}}\right) \\
&= \sum_{d\geq1}\sum_{\substack{w\in\mathfrak{L};\\ \ell(w)=d}} \underbrace{\log\left(\frac{1}{1-\mathbf{x}_w t^{\ell(w)}}\right)}_{\substack{=\log\left(\dfrac{1}{1-\mathbf{x}_w t^{d}}\right) \\ \text{(since } \ell(w)=d\text{)}}} = \sum_{d\geq1}\sum_{\substack{w\in\mathfrak{L};\\ \ell(w)=d}} \underbrace{\log\left(\frac{1}{1-\mathbf{x}_w t^{d}}\right)}_{\substack{=-\log\left(1-\mathbf{x}_w t^d\right) \\ =\sum_{k\geq1}\dfrac{1}{k}\left(\mathbf{x}_w t^d\right)^k \\ \text{(by substituting } \mathbf{x}_w t^d \text{ for } t \\ \text{in (13.192.54))}}} \\
&= \sum_{d\geq1}\underbrace{\sum_{\substack{w\in\mathfrak{L};\\ \ell(w)=d}}\sum_{k\geq1}}_{=\sum_{k\geq1}\sum_{w\in\mathfrak{L};\ \ell(w)=d}} \frac{1}{k}\underbrace{\left(\mathbf{x}_w t^d\right)^k}_{=\mathbf{x}_w^k t^{kd}} = \sum_{d\geq1}\sum_{k\geq1}\sum_{\substack{w\in\mathfrak{L};\\ \ell(w)=d}}\frac{1}{k}\mathbf{x}_w^k t^{kd} = \sum_{d\geq1}\sum_{k\geq1}\frac{1}{k}\underbrace{\left(\sum_{\substack{w\in\mathfrak{L};\\ \ell(w)=d}}\mathbf{x}_w^k\right)}_{\substack{=\mathbf{GR}_{(d)}\left(x_1^k,x_2^k,x_3^k,\ldots\right) \\ \text{(by (13.192.52))}}} t^{kd} \\
&= \sum_{d\geq1}\underbrace{\sum_{k\geq1}\frac{1}{k}\mathbf{GR}_{(d)}\left(x_1^k,x_2^k,x_3^k,\ldots\right)t^{kd}}_{\substack{=\sum_{n\geq1;\ d\mid n}\dfrac{1}{n/d}\mathbf{GR}_{(d)}\left(x_1^{n/d},x_2^{n/d},x_3^{n/d},\ldots\right)t^{(n/d)d} \\ \text{(here, we have substituted } n/d \text{ for } k \text{ in the sum)}}} \\
&= \underbrace{\sum_{d\geq1}\sum_{\substack{n\geq1;\\ d\mid n}}}_{\substack{=\sum_{n\geq1}\sum_{d\mid n} \\ \text{(because the summation sign ``}\sum_{d\mid n}\text{''} \\ \text{ranges over all } \textbf{positive} \text{ divisors } d \text{ of } n\text{)}}} \frac{1}{n/d}\mathbf{GR}_{(d)}\left(x_1^{n/d},x_2^{n/d},x_3^{n/d},\ldots\right)\underbrace{t^{(n/d)d}}_{=t^n} \\
&= \sum_{n\geq1}\sum_{d\mid n}\frac{1}{n/d}\mathbf{GR}_{(d)}\left(x_1^{n/d},x_2^{n/d},x_3^{n/d},\ldots\right)t^n.
\end{aligned}$$

Comparing this with

$$\log \frac{1}{1-p_1 t} = -\log\left(1-p_1 t\right) = \sum_{k\geq 1} \frac{1}{k} \underbrace{\left(p_1 t\right)^k}_{=p_1^k t^k}$$

(this follows by substituting $p_1 t$ for t in (13.192.54))

$$= \sum_{k\geq 1} \frac{1}{k} p_1^k t^k = \sum_{n\geq 1} \frac{1}{n} p_1^n t^n,$$

we obtain

$$\sum_{n\geq 1} \frac{1}{n} p_1^n t^n = \sum_{n\geq 1} \sum_{d\mid n} \frac{1}{n/d} \mathbf{GR}_{(d)} \left(x_1^{n/d}, x_2^{n/d}, x_3^{n/d}, \ldots\right) t^n.$$

This is an equality between two formal power series in t. If we compare coefficients on both sides of this equality, then we obtain

$$\frac{1}{n} p_1^n = \sum_{d\mid n} \frac{1}{n/d} \mathbf{GR}_{(d)} \left(x_1^{n/d}, x_2^{n/d}, x_3^{n/d}, \ldots\right) \tag{13.192.55}$$

for each $n \in \{1,2,3,\ldots\}$.

Now, let n be a positive integer. Thus, $n \in \{1,2,3,\ldots\}$. Hence, the equality (13.192.55) holds. Multiplying both sides of this equality by n, we find

$$p_1^n = \sum_{d\mid n} \underbrace{n \cdot \frac{1}{n/d}}_{=d} \mathbf{GR}_{(d)} \left(x_1^{n/d}, x_2^{n/d}, x_3^{n/d}, \ldots\right) = \sum_{d\mid n} d \cdot \mathbf{GR}_{(d)} \left(x_1^{n/d}, x_2^{n/d}, x_3^{n/d}, \ldots\right).$$

This proves Claim 9.]

Claim 10: Every positive integer n satisfies

$$p_1^n = \sum_{d\mid n} d \cdot \widetilde{\mathbf{GR}}_d \left(x_1^{n/d}, x_2^{n/d}, x_3^{n/d}, \ldots\right).$$

[*Proof of Claim 10:* Let n be a positive integer.

Let f be a positive divisor of n. Thus, n/f is a positive integer. The definition of $\widetilde{\mathbf{GR}}_f$ yields

$$\widetilde{\mathbf{GR}}_f = \frac{1}{f} \sum_{d\mid f} \mu\left(d\right) p_d^{f/d}.$$

Substituting $x_1^{n/f}, x_2^{n/f}, x_3^{n/f}, \ldots$ for $x_1, x_2, x_3, \ldots$ on both sides of this equality, we obtain

$$\widetilde{\mathbf{GR}}_f \left(x_1^{n/f}, x_2^{n/f}, x_3^{n/f}, \ldots\right) = \frac{1}{f} \sum_{d\mid f} \mu\left(d\right) \left(\underbrace{p_d\left(x_1^{n/f}, x_2^{n/f}, x_3^{n/f}, \ldots\right)}_{\substack{=p_{(n/f)d} \\ \text{(by Claim 7, applied to } k=n/f \text{ and } m=d)}} \right)^{f/d}$$

$$= \frac{1}{f} \sum_{d\mid f} \mu\left(d\right) p_{(n/f)d}^{f/d}.$$

Multiplying both sides of this equality by f, we obtain

$$f \cdot \widetilde{\mathbf{GR}}_f \left(x_1^{n/f}, x_2^{n/f}, x_3^{n/f}, \ldots\right) = \sum_{d\mid f} \mu\left(d\right) p_{(n/f)d}^{f/d}. \tag{13.192.56}$$

Forget that we fixed f. We thus have proved the equality (13.192.56) for each positive divisor f of n.

Now,

$$\begin{aligned}
&\sum_{d\mid n} d\cdot\widetilde{\mathbf{GR}}_d\left(x_1^{n/d},x_2^{n/d},x_3^{n/d},\ldots\right)\\
&=\sum_{f\mid n}\underbrace{f\cdot\widetilde{\mathbf{GR}}_f\left(x_1^{n/f},x_2^{n/f},x_3^{n/f},\ldots\right)}_{\substack{=\sum_{d\mid f}\mu(d)p_{(n/f)d}^{f/d}\\ \text{(by (13.192.56))}}}\\
&\qquad\text{(here, we have renamed the summation index } d \text{ as } f\text{)}\\
&=\underbrace{\sum_{f\mid n}\sum_{d\mid f}}_{=\sum_{d\mid n}\sum_{\substack{f\mid n;\\ d\mid f}}}\mu\left(d\right)p_{(n/f)d}^{f/d}=\sum_{d\mid n}\sum_{\substack{f\mid n;\\ d\mid f}}\mu\left(d\right)p_{(n/f)d}^{f/d}. \qquad (13.192.57)
\end{aligned}$$

Now, fix a positive divisor d of n. Then, n/d is a positive integer. Hence, the map

$$\begin{aligned}
\{\text{positive divisors } g \text{ of } n/d\}&\to\{\text{positive divisors } f \text{ of } n \text{ satisfying } d\mid f\},\\
g&\mapsto dg
\end{aligned}$$

is a bijection[1223]. Thus, we can substitute dg for f in the sum $\sum_{\substack{f\mid n;\\ d\mid f}}\mu\left(d\right)p_{(n/f)d}^{f/d}$. We thus obtain

$$\sum_{\substack{f\mid n;\\ d\mid f}}\mu\left(d\right)p_{(n/f)d}^{f/d}=\sum_{g\mid n/d}\mu\left(d\right)\underbrace{p_{(n/(dg))d}^{dg/d}}_{=p_{n/g}^{g}}=\sum_{g\mid n/d}\mu\left(d\right)p_{n/g}^{g}. \qquad (13.192.58)$$

Forget that we fixed d. We thus have proved the equality (13.192.58) for each positive divisor d of n. Now, (13.192.57) becomes

$$\begin{aligned}
\sum_{d\mid n} d\cdot\widetilde{\mathbf{GR}}_d\left(x_1^{n/d},x_2^{n/d},x_3^{n/d},\ldots\right)&=\sum_{d\mid n}\underbrace{\sum_{\substack{f\mid n;\\ d\mid f}}\mu\left(d\right)p_{(n/f)d}^{f/d}}_{\substack{=\sum_{g\mid n/d}\mu(d)p_{n/g}^{g}\\ \text{(by (13.192.58))}}}\\
&=\sum_{d\mid n}\sum_{g\mid n/d}\mu\left(d\right)p_{n/g}^{g}. \qquad (13.192.59)
\end{aligned}$$

But if d and g are two positive integers, then the statement "$d\mid n$ and $g\mid n/d$" is equivalent to the statement "$g\mid n$ and $d\mid n/g$" [1224]. Hence, we have the following equality of summation signs:

$$\sum_{d\mid n}\sum_{g\mid n/d}=\sum_{g\mid n}\sum_{d\mid n/g}. \qquad (13.192.60)$$

[1223] Its inverse sends each e to e/d.

[1224] Indeed, it is easy to see that both of these statements are equivalent to "$dg\mid n$".

Thus, (13.192.59) becomes

$$\begin{aligned}
\sum_{d\mid n} d\cdot \widetilde{\mathbf{GR}}_d\left(x_1^{n/d},x_2^{n/d},x_3^{n/d},\ldots\right) &= \underbrace{\sum_{d\mid n}\sum_{g\mid n/d}}_{\substack{=\sum\limits_{g\mid n}\sum\limits_{d\mid n/g} \\ \text{(by (13.192.60))}}} \mu\left(d\right)p_{n/g}^{g} \\
&= \sum_{g\mid n} \underbrace{\sum_{d\mid n/g}\mu\left(d\right)}_{\substack{=\delta_{n/g,1} \\ \text{(by (13.84.3), applied} \\ \text{to } n/g \text{ instead of } n)}} p_{n/g}^{g} = \sum_{g\mid n}\delta_{n/g,1}p_{n/g}^{g} \\
&= \underbrace{\delta_{n/n,1}}_{\substack{=1 \\ \text{(since } n/n=1)}} \underbrace{p_{n/n}^{n}}_{=p_1^n} + \sum_{\substack{g\mid n;\\ g\neq n}} \underbrace{\delta_{n/g,1}}_{\substack{=0 \\ \text{(since } n/g\neq 1 \\ \text{(because } g\neq n))}} p_{n/g}^{g} \\
&\qquad \text{(here, we have split off the addend for } g=n \text{ from the sum)} \\
&= p_1^n + \underbrace{\sum_{\substack{g\mid n;\\ g\neq n}} 0p_{n/g}^{g}}_{=0} = p_1^n .
\end{aligned}$$

This proves Claim 10.]

Claim 11: We have $\mathbf{GR}_{(n)} = \widetilde{\mathbf{GR}}_n$ for each positive integer n.

[*Proof of Claim 11:* We shall prove Claim 11 by strong induction on n:

Induction step: Fix a positive integer m. Assume (as the induction hypothesis) that Claim 11 holds for all $n < m$. We must then show that Claim 11 holds for $n = m$.

As a consequence of our induction hypothesis, we can easily see that

$$\mathbf{GR}_{(d)} = \widetilde{\mathbf{GR}}_d \tag{13.192.61}$$

whenever d is a positive divisor of m satisfying $d \neq m$. [1225]

Now, Claim 9 (applied to $n = m$) yields

$$\begin{aligned}
p_1^m &= \sum_{d\mid m} d\cdot \mathbf{GR}_{(d)}\left(x_1^{m/d},x_2^{m/d},x_3^{m/d},\ldots\right) \\
&= m\cdot \underbrace{\mathbf{GR}_{(m)}\left(x_1^{m/m},x_2^{m/m},x_3^{m/m},\ldots\right)}_{\substack{=\mathbf{GR}_{(m)}(x_1,x_2,x_3,\ldots) \\ \text{(since } \left(x_1^{m/m},x_2^{m/m},x_3^{m/m},\ldots\right)=(x_1,x_2,x_3,\ldots) \\ \text{(because } x_i^{m/m}=x_i^1=x_i \text{ for each } i\in\{1,2,3,\ldots\}))}} + \sum_{\substack{d\mid m;\\ d\neq m}} d\cdot \underbrace{\mathbf{GR}_{(d)}}_{\substack{=\widetilde{\mathbf{GR}}_d \\ \text{(by (13.192.61))}}} \left(x_1^{m/d},x_2^{m/d},x_3^{m/d},\ldots\right) \\
&\qquad \left(\begin{array}{c}\text{here, we have split off the addend for } d=m \text{ from the sum} \\ \text{(since } m \text{ is a positive divisor of } m)\end{array}\right) \\
&= m\cdot \underbrace{\mathbf{GR}_{(m)}\left(x_1,x_2,x_3,\ldots\right)}_{\substack{=\mathbf{GR}_{(m)} \\ \text{(since } f(x_1,x_2,x_3,\ldots)=f \\ \text{for each } f\in\mathbf{k}[[\mathbf{x}]])}} + \sum_{\substack{d\mid m;\\ d\neq m}} d\cdot \widetilde{\mathbf{GR}}_d\left(x_1^{m/d},x_2^{m/d},x_3^{m/d},\ldots\right) \\
&= m\cdot \mathbf{GR}_{(m)} + \sum_{\substack{d\mid m;\\ d\neq m}} d\cdot \widetilde{\mathbf{GR}}_d\left(x_1^{m/d},x_2^{m/d},x_3^{m/d},\ldots\right).
\end{aligned}$$

[1225] *Proof of (13.192.61):* Let d be a positive divisor of m satisfying $d \neq m$. Then, $d \leq m$ (since d is a positive divisor of the positive integer m). Combining this with $d \neq m$, we obtain $d < m$.

But we have assumed (as the induction hypothesis) that Claim 11 holds for all $n < m$. Hence, Claim 11 holds for $n = d$ (since d is a positive integer satisfying $d < m$). In other words, we have $\mathbf{GR}_{(d)} = \widetilde{\mathbf{GR}}_d$. This proves (13.192.61).

Meanwhile, Claim 10 (applied to $n = m$) yields

$$\begin{aligned}
p_1^m &= \sum_{d \mid m} d \cdot \widetilde{\mathbf{GR}}_d \left(x_1^{m/d}, x_2^{m/d}, x_3^{m/d}, \ldots \right) \\
&= m \cdot \underbrace{\widetilde{\mathbf{GR}}_m \left(x_1^{m/m}, x_2^{m/m}, x_3^{m/m}, \ldots \right)}_{\substack{= \widetilde{\mathbf{GR}}_m (x_1, x_2, x_3, \ldots) \\ \text{(since } \left(x_1^{m/m}, x_2^{m/m}, x_3^{m/m}, \ldots \right) = (x_1, x_2, x_3, \ldots) \\ \text{(because } x_i^{m/m} = x_i^1 = x_i \text{ for each } i \in \{1,2,3,\ldots\}\text{))}}} + \sum_{\substack{d \mid m; \\ d \neq m}} d \cdot \widetilde{\mathbf{GR}}_d \left(x_1^{m/d}, x_2^{m/d}, x_3^{m/d}, \ldots \right) \\
&\qquad \left(\begin{array}{c} \text{here, we have split off the addend for } d = m \text{ from the sum} \\ \text{(since } m \text{ is a positive divisor of } m\text{)} \end{array} \right) \\
&= m \cdot \underbrace{\widetilde{\mathbf{GR}}_m (x_1, x_2, x_3, \ldots)}_{\substack{= \widetilde{\mathbf{GR}}_m \\ \text{(since } f(x_1, x_2, x_3, \ldots) = f \\ \text{for each } f \in \mathbf{k}[[\mathbf{x}]]\text{)}}} + \sum_{\substack{d \mid m; \\ d \neq m}} d \cdot \widetilde{\mathbf{GR}}_d \left(x_1^{m/d}, x_2^{m/d}, x_3^{m/d}, \ldots \right) \\
&= m \cdot \widetilde{\mathbf{GR}}_m + \sum_{\substack{d \mid m; \\ d \neq m}} d \cdot \widetilde{\mathbf{GR}}_d \left(x_1^{m/d}, x_2^{m/d}, x_3^{m/d}, \ldots \right).
\end{aligned}$$

Comparing these two equalities, we obtain

$$\begin{aligned}
& m \cdot \mathbf{GR}_{(m)} + \sum_{\substack{d \mid m; \\ d \neq m}} d \cdot \widetilde{\mathbf{GR}}_d \left(x_1^{m/d}, x_2^{m/d}, x_3^{m/d}, \ldots \right) \\
&= m \cdot \widetilde{\mathbf{GR}}_m + \sum_{\substack{d \mid m; \\ d \neq m}} d \cdot \widetilde{\mathbf{GR}}_d \left(x_1^{m/d}, x_2^{m/d}, x_3^{m/d}, \ldots \right).
\end{aligned}$$

Subtracting $\sum_{\substack{d \mid m; \\ d \neq m}} d \cdot \widetilde{\mathbf{GR}}_d \left(x_1^{m/d}, x_2^{m/d}, x_3^{m/d}, \ldots \right)$ from both sides of this equality, we find $m \cdot \mathbf{GR}_{(m)} = m \cdot \widetilde{\mathbf{GR}}_m$.

But m is a positive integer, and thus is invertible in $\mathbb{Q}$. Hence, m is invertible in $\mathbf{k}$ as well (since $\mathbf{k}$ is a $\mathbb{Q}$-algebra). Thus, we can divide both sides of the equality $m \cdot \mathbf{GR}_{(m)} = m \cdot \widetilde{\mathbf{GR}}_m$ by m. We thus obtain $\mathbf{GR}_{(m)} = \widetilde{\mathbf{GR}}_m$. In other words, Claim 11 holds for $n = m$. This completes the induction step. Hence, Claim 11 is proved by strong induction.]

Now, fix a positive integer n. Then, Claim 11 yields

$$\mathbf{GR}_{(n)} = \widetilde{\mathbf{GR}}_n = \frac{1}{n} \sum_{d \mid n} \mu(d) \, p_d^{n/d} \qquad \left(\text{by the definition of } \widetilde{\mathbf{GR}}_n \right).$$

This proves Proposition 6.6.36(b). □

We have now proved Proposition 6.6.39, Proposition 6.6.38, Proposition 6.6.43, Proposition 6.6.48, Proposition 6.6.49, Proposition 6.6.50, Proposition 6.6.40, Proposition 6.6.42, Proposition 6.6.37 and Proposition 6.6.36. In other words, we have proved all statements made in Subsection 6.6.2. This solves Exercise 6.6.51.

13.193. **Solution to Exercise 6.6.53.** *Solution to Exercise 6.6.53.* Let us first forget about Remark 6.6.52(b), and prove some basic lemmas. Our first lemma will be about roots of unity:

Lemma 13.193.1. *Let n be a positive integer. Let ω be a primitive n-th root of unity in $\mathbb{C}$ (for instance, $\exp(2\pi i/n)$). Then:*

(a) *If d is a positive divisor of n, then*

$$\sum_{\substack{i\in\{1,2,\ldots,n\};\\ d\mid i}} \omega^i = \delta_{n/d,1}.$$

(b) *If d is a positive divisor of n, then*

$$\sum_{\substack{i\in\{1,2,\ldots,n\};\\ \gcd(n,i)=d}} \omega^i = \mu\left(n/d\right).$$

Here, μ denotes the number-theoretical Möbius function (defined as in Exercise 2.9.6).

Proof of Lemma 13.193.1. We have required ω to be a primitive n-th root of unity in $\mathbb{C}$. Thus, $\omega^n = 1$, but if d is a positive integer satisfying $d < n$, then

$$\omega^d \neq 1. \tag{13.193.1}$$

(a) Let d be a positive divisor of n. Then, the elements $i \in \{1, 2, \ldots, n\}$ satisfying $d \mid i$ are precisely the n/d distinct elements $1d, 2d, \ldots, (n/d)\, d$. Hence,

$$\sum_{\substack{i\in\{1,2,\ldots,n\};\\ d\mid i}} \omega^i = \omega^{1d} + \omega^{2d} + \cdots + \omega^{(n/d)d} = \sum_{j=1}^{n/d} \underbrace{\omega^{jd}}_{=(\omega^d)^j} = \sum_{j=1}^{n/d} \left(\omega^d\right)^j. \tag{13.193.2}$$

Now, we are in one of the following two cases:

Case 1: We have $d = n$.

Case 2: We have $d \neq n$.

Let us first consider Case 1. In this case, we have $d = n$. Hence, $\omega^d = \omega^n = 1$ and $n/d = n/n = 1$. Thus, (13.193.2) becomes

$$\sum_{\substack{i\in\{1,2,\ldots,n\};\\ d\mid i}} \omega^i = \sum_{j=1}^{n/d} \left(\underbrace{\omega^d}_{=1}\right)^j = \sum_{j=1}^{n/d} \underbrace{1^j}_{=1} = \sum_{j=1}^{n/d} 1 = n/d = 1 = \delta_{n/d,1}.$$

(since $n/d = 1$ leads to $\delta_{n/d,1} = 1$). Hence, Lemma 13.193.1(a) is proved in Case 1.

Let us next consider Case 2. In this case, we have $d \neq n$. But d is a positive divisor of the positive integer n; hence, $d \leq n$. Combining this with $d \neq n$, we obtain $d < n$. Hence, (13.193.1) yields $\omega^d \neq 1$. Thus, $1 - \omega^d \neq 0$. Also, from $d \neq n$, we obtain $n/d \neq 1$.

But n/d is a positive integer (since d is a positive divisor of the positive integer n). Now,

$$\begin{aligned}
\sum_{\substack{i\in\{1,2,\ldots,n\};\\ d\mid i}} \omega^i &= \omega^{1d} + \omega^{2d} + \cdots + \omega^{(n/d)d} \\
&= \left(\omega^{1d} + \omega^{2d} + \cdots + \omega^{(n/d-1)d}\right) + \underbrace{\omega^{(n/d)d}}_{\substack{=\omega^n=1=\omega^{0d} \\ (\text{since } \omega^{0d}=\omega^0=1)}} \qquad (\text{since } n/d \text{ is a positive integer}) \\
&= \left(\omega^{1d} + \omega^{2d} + \cdots + \omega^{(n/d-1)d}\right) + \omega^{0d} \\
&= \omega^{0d} + \left(\omega^{1d} + \omega^{2d} + \cdots + \omega^{(n/d-1)d}\right) \\
&= \omega^{0d} + \omega^{1d} + \cdots + \omega^{(n/d-1)d} = \sum_{j=0}^{n/d-1} \underbrace{\omega^{jd}}_{=(\omega^d)^j} = \sum_{j=0}^{n/d-1} \left(\omega^d\right)^j.
\end{aligned}$$

Hence,

$$\sum_{j=0}^{n/d-1}\left(\omega^d\right)^j = \sum_{\substack{i\in\{1,2,\ldots,n\};\\ d\mid i}} \omega^i = \sum_{j=1}^{n/d} \underbrace{\left(\omega^d\right)^j}_{=\omega^d\left(\omega^d\right)^{j-1}} \qquad (\text{by } (13.193.2)) \tag{13.193.3}$$

$$= \omega^d \sum_{j=1}^{n/d}\left(\omega^d\right)^{j-1} = \omega^d \sum_{j=0}^{n/d-1}\left(\omega^d\right)^j \tag{13.193.4}$$

$$(\text{here, we have substituted } j \text{ for } j-1 \text{ in the sum}).$$

Hence,

$$\left(1-\omega^d\right)\sum_{j=0}^{n/d-1}\left(\omega^d\right)^j = \underbrace{\sum_{j=0}^{n/d-1}\left(\omega^d\right)^j}_{\substack{=\omega^d\sum_{j=0}^{n/d-1}\left(\omega^d\right)^j\\ (\text{by } (13.193.4))}} - \omega^d\sum_{j=0}^{n/d-1}\left(\omega^d\right)^j = \omega^d\sum_{j=0}^{n/d-1}\left(\omega^d\right)^j - \omega^d\sum_{j=0}^{n/d-1}\left(\omega^d\right)^j = 0.$$

We can divide both sides of this equality by $1-\omega^d$ (since $1-\omega^d \neq 0$), and thus obtain $\sum_{j=0}^{n/d-1}\left(\omega^d\right)^j = 0$. Comparing this with (13.193.3), we obtain $\sum_{j=1}^{n/d}\left(\omega^d\right)^j = 0$. Thus, (13.193.2) becomes

$$\sum_{\substack{i\in\{1,2,\ldots,n\};\\ d\mid i}} \omega^i = \sum_{j=1}^{n/d}\left(\omega^d\right)^j = 0 = \delta_{n/d,1} \qquad \left(\text{since } n/d \neq 1 \text{ leads to } \delta_{n/d,1} = 0\right).$$

Hence, Lemma 13.193.1(a) is proved in Case 2.

We have now proved Lemma 13.193.1(a) in both Cases 1 and 2. Thus, Lemma 13.193.1(a) always holds.

(b) For each positive integer m, we have[1226]

$$\sum_{d\mid m}\mu(d) = \delta_{m,1} \qquad (\text{by } (13.84.3), \text{ applied to } m \text{ instead of } n)$$

and thus

$$\delta_{m,1} = \sum_{d\mid m}\mu(d) = \sum_{f\mid m}\mu(f) \tag{13.193.5}$$

(here, we have renamed the summation index d as f).

Set

$$x_d = \sum_{\substack{i\in\{1,2,\ldots,n\};\\ \gcd(n,i)=d}} \omega^i \qquad \text{for each } d\in\mathbb{Z}. \tag{13.193.6}$$

We now claim that

$$x_d = \mu(n/d) \qquad \text{for each positive divisor } d \text{ of } n. \tag{13.193.7}$$

[*Proof of* (13.193.7)*:* If d is a positive divisor of n, then n/d is a positive integer (since n is a positive integer). Thus, we can prove (13.193.7) by strong induction on n/d. Let us do this:

Induction step: Fix a positive integer m. Assume (as the induction hypothesis) that (13.193.7) holds whenever $n/d < m$. We must prove that (13.193.7) holds whenever $n/d = m$.

If d is a positive divisor of n that satisfies $n/d < m$, then

$$x_d = \mu(n/e). \tag{13.193.8}$$

(Indeed, this is just a restatement of our induction hypothesis.)

Now, fix a positive divisor d of n that satisfies $n/d = m$. We shall prove that $x_d = \mu(n/d)$.

[1226] Here and in the following, the summation sign "$\sum_{d\mid m}$" will always be understood to range over all **positive** divisors d of m.

As a consequence of the induction hypothesis, we can easily see the following: If e is a positive divisor of n satisfying $d \mid e$ and $e \neq d$, then

$$x_e = \mu(n/e). \tag{13.193.9}$$

[*Proof of* (13.193.9)*:* Let e be a positive divisor of n satisfying $d \mid e$ and $e \neq d$. From $d \mid e$, we obtain $d \leq e$ (since d and e are positive integers). Combining this with $d \neq e$ (which follows from $e \neq d$), we obtain $d < e$. Since n is positive, this entails $n/e < n/d = m$. Hence, (13.193.8) (applied to e instead of d) yields $x_e = \mu(n/e)$. This proves (13.193.9).]

Recall that n is a positive integer, and d is a positive divisor of n. Hence, n/d is a positive integer, and we have $d \mid n$.

For any $i \in \{1, 2, \ldots, n\}$, we have the following chain of logical equivalences:

$$\begin{aligned}
&((\gcd(n,i) \text{ is a positive divisor of } n) \wedge (d \mid \gcd(n,i))) \\
\iff\ & (d \mid \gcd(n,i)) \qquad (\text{since ``}\gcd(n,i) \text{ is a positive divisor of } n\text{'' is always true}) \\
\iff\ & ((d \mid n) \wedge (d \mid i)) \\
& \left(\begin{array}{c}\text{because the logical equivalence } (x \mid \gcd(y,z)) \iff ((x \mid y) \wedge (x \mid z)) \\ \text{holds for any three integers } x,\ y \text{ and } z\end{array}\right) \\
\iff\ & (d \mid i) \qquad (\text{since ``}d \mid n\text{'' is always true}).
\end{aligned}$$

Hence, we have the following equality of summation signs:

$$\sum_{\substack{i\in\{1,2,\ldots,n\};\\ \gcd(n,i) \text{ is a positive divisor of } n;\\ d|\gcd(n,i)}} = \sum_{\substack{i\in\{1,2,\ldots,n\};\\ d|i}}. \tag{13.193.10}$$

We have[1227]

$$\begin{aligned}
\underbrace{\sum_{\substack{e|n;\\ d|e}}}_{=\sum_{\substack{e \text{ is a positive divisor of } n;\\ d|e}}} \underbrace{x_e}_{\substack{=\sum_{\substack{i\in\{1,2,\ldots,n\};\\ \gcd(n,i)=e}} \omega^i \\ (\text{by the definition of } x_e)}}
&= \underbrace{\sum_{\substack{e \text{ is a positive divisor of } n;\\ d|e}} \sum_{\substack{i\in\{1,2,\ldots,n\};\\ \gcd(n,i)=e}}}_{\substack{=\sum_{\substack{i\in\{1,2,\ldots,n\};\\ \gcd(n,i) \text{ is a positive divisor of } n;\\ d|\gcd(n,i)}} \\ =\sum_{\substack{i\in\{1,2,\ldots,n\};\\ d|i}} \\ (\text{by (13.193.10)})}} \omega^i \\
&= \sum_{\substack{i\in\{1,2,\ldots,n\};\\ d|i}} \omega^i = \delta_{n/d,1}
\end{aligned}$$

(by Lemma 13.193.1(a)). Hence,

$$\begin{aligned}
\delta_{n/d,1} &= \sum_{\substack{e|n;\\ d|e}} x_e = x_d + \sum_{\substack{e|n;\\ d|e;\\ e\neq d}} \underbrace{x_e}_{\substack{=\mu(n/e)\\ (\text{by (13.193.9)})}} \\
&\qquad (\text{here, we have split off the addend for } e = d \text{ from the sum}) \\
&= x_d + \sum_{\substack{e|n;\\ d|e;\\ e\neq d}} \mu(n/e). \qquad (13.193.11)
\end{aligned}$$

On the other hand, recall that n/d is a positive integer. Hence, (13.193.5) (applied to $m = n/d$) yields

$$\delta_{n/d,1} = \sum_{f|n/d} \mu(f). \tag{13.193.12}$$

[1227] Recall that the summation sign "$\sum_{\substack{e|n;\\ d|e}}$" stands for a sum over all **positive** divisors e of n satisfying $d \mid e$.

But it is straightforward to see that the map

$$\{\text{positive divisors } e \text{ of } n \text{ satisfying } d \mid e\} \to \{\text{positive divisors of } n/d\}, \qquad e \mapsto n/e$$

is well-defined and is a bijection[1228]. Thus, we can substitute n/e for f in the sum $\sum\limits_{f\mid n/d} \mu(f)$. We therefore obtain

$$\sum_{f\mid n/d} \mu(f) = \sum_{\substack{e\mid n;\\ d\mid e}} \mu(n/e).$$

Thus, (13.193.12) becomes

$$\delta_{n/d,1} = \sum_{f\mid n/d} \mu(f) = \sum_{\substack{e\mid n;\\ d\mid e}} \mu(n/e) = \mu(n/d) + \sum_{\substack{e\mid n;\\ d\mid e;\\ e\neq d}} \mu(n/e)$$

(here, we have split off the addend for $e = d$ from the sum). Comparing this with (13.193.11), we obtain

$$x_d + \sum_{\substack{e\mid n;\\ d\mid e;\\ e\neq d}} \mu(n/e) = \mu(n/d) + \sum_{\substack{e\mid n;\\ d\mid e;\\ e\neq d}} \mu(n/e).$$

Subtracting $\sum\limits_{\substack{e\mid n;\\ d\mid e;\\ e\neq d}} \mu(n/e)$ from both sides of this equality, we obtain $x_d = \mu(n/d)$. In other words, (13.193.7) holds.

Forget that we fixed d. We thus have proved that (13.193.7) holds whenever $n/d = m$. This completes the induction step. Thus, the induction proof of (13.193.7) is complete.]

Now, let d be a positive divisor of n. Then, the definition of x_d yields $x_d = \sum\limits_{\substack{i\in\{1,2,\ldots,n\};\\ \gcd(n,i)=d}} \omega^i$. Hence,

$$\sum_{\substack{i\in\{1,2,\ldots,n\};\\ \gcd(n,i)=d}} \omega^i = x_d = \mu(n/d) \qquad (\text{by } (13.193.7)).$$

This proves Lemma 13.193.1(b). □

Remark 13.193.2. Let n be a positive integer. If d is a positive divisor of n, then the complex numbers ω^i for all $i \in \{1, 2, \ldots, n\}$ satisfying $\gcd(n, i) = d$ are precisely the primitive (n/d)-th roots of unity in $\mathbb{C}$ (this is not hard to check). Thus, Lemma 13.193.1(b) states that the sum of all primitive (n/d)-th roots of unity in $\mathbb{C}$ is $\mu(n/d)$. Hence, Lemma 13.193.1(b) is an equivalent restatement of the following fact: For any positive integer m, the sum of all primitive m-th roots of unity in $\mathbb{C}$ is $\mu(m)$.

The next lemma we will need is an elementary fact from the theory of numbers, whose easy proof (e.g., using the Bezout theorem or the basic properties of greatest common divisors) we omit:

Lemma 13.193.3. *Let x, y and z be three integers such that $x \neq 0$. Let $d = \gcd(x, y)$. Then, we have the logical equivalence $(x \mid yz) \iff (x/d \mid z)$.*

The next lemma we will need is a mostly trivial property of permutations:

Lemma 13.193.4. *Let n be a positive integer. For every permutation $\sigma \in \mathfrak{S}_n$, we let $\operatorname{type}\sigma$ denote the cycle type of σ.*

Let $q \in \mathbb{Q}$ and $\sigma \in \mathfrak{S}_n$. Assume that every cycle of σ has size q. Then, q is a positive divisor of n, and we have $\operatorname{type}\sigma = \left(\underbrace{q, q, \ldots, q}_{n/q \text{ times}}\right)$.

[1228] Its inverse sends each $f \in \{\text{positive divisors of } n/d\}$ to n/f.

Proof of Lemma 13.193.4. Let k be the number of cycles of σ. Thus, σ has exactly k cycles, and each of these k cycles has size q (because we assumed that every cycle of σ has size q). But the cycles of σ are disjoint, and their union is the n-element set $\{1, 2, \ldots, n\}$. Hence, the sum of the sizes of all cycles of σ is n. Thus,

$$\begin{aligned} n &= (\text{the sum of the sizes of all cycles of } \sigma) = \underbrace{q + q + \cdots + q}_{k \text{ times}} \\ &\qquad (\text{since } \sigma \text{ has exactly } k \text{ cycles, and each of these } k \text{ cycles has size } q) \\ &= kq. \end{aligned}$$

Thus, $kq = n \neq 0$, so that $k \neq 0$ and $q \neq 0$. From $k \neq 0$, we conclude that σ has at least one cycle. The size of this cycle is q (since we assumed that every cycle of σ has size q); thus, q is the size of a cycle of σ, and therefore is a nonnegative integer. Combining this with $q \neq 0$, we conclude that q is a positive integer. Since $n = kq$, we conclude that $q \mid n$ (because $k \in \mathbb{N}$), so that q is a positive divisor of n. Furthermore, the definition of $\operatorname{type} \sigma$ shows that

$$\begin{aligned} \operatorname{type} \sigma &= (\text{the cycle type of } \sigma) = \Bigg(\underbrace{q, q, \ldots, q}_{k \text{ times}}\Bigg) \\ &\qquad (\text{since } \sigma \text{ has exactly } k \text{ cycles, and each of these } k \text{ cycles has size } q) \\ &= \Bigg(\underbrace{q, q, \ldots, q}_{n/q \text{ times}}\Bigg) \qquad (\text{since } k = n/q \text{ (because } n = kq)). \end{aligned}$$

This proves Lemma 13.193.4. □

Let us now step to the analysis of n-cycles in $\mathfrak{S}_n$:

Lemma 13.193.5. *Let n be a positive integer. For every permutation $\sigma \in \mathfrak{S}_n$, we let $\operatorname{type} \sigma$ denote the cycle type of σ.*

Let $z \in \mathfrak{S}_n$ be an n-cycle. Let $m \in \mathbb{Z}$. Let $d = \gcd(m, n)$. Then, $\operatorname{type}(z^m) = \Bigg(\underbrace{n/d, n/d, \ldots, n/d}_{d \text{ times}}\Bigg)$.

Proof of Lemma 13.193.5. Let us consider the cycles of the permutation $z^m \in \mathfrak{S}_n$. We shall show that every cycle of z^m has size n/d.

Indeed, let Z be a cycle of z^m (considered as a subset of $\{1, 2, \ldots, n\}$). Then, Z is a nonempty set; in other words, there exists some $x \in Z$. Consider this x.

Let $g = |Z|$. Then, g is a positive integer (since Z is nonempty), and the cycle Z has length g (since $g = |Z|$).

Now, Z is the cycle of z^m containing x (since Z is a cycle of z^m and since $x \in Z$). Thus, the cycle of z^m containing x has length g (since we know that the cycle Z has length g). Hence, the sequence $x, z^m(x), (z^m)^2(x), (z^m)^3(x), \ldots$ (obtained from x by applying z^m again and again) repeats every g elements, but not more frequently. Thus, for any $N \in \mathbb{N}$, we have the following equivalence of statements:

$$\left((z^m)^N(x) = x\right) \iff (g \mid N). \tag{13.193.13}$$

On the other hand, the permutation z is an n-cycle on an n-element set (namely, on $\{1, 2, \ldots, n\}$); thus, it has exactly one cycle, which has length n. The cycle of z which contains x must therefore be this unique cycle, and thus has length n. Hence, the sequence $x, z(x), z^2(x), z^3(x), \ldots$ (obtained from x by applying z again and again) repeats every n elements, but not more frequently. Thus, for any $N \in \mathbb{N}$, we have the following equivalence of statements:

$$\left(z^N(x) = x\right) \iff (n \mid N). \tag{13.193.14}$$

Note that $n \neq 0$. Hence, d is a positive integer (since $d = \gcd(m, n)$) and satisfies $d = \gcd(m, n) = \gcd(n, m) \mid n$. Hence, d is a positive divisor of n. Therefore, n/d is a positive integer (since n is also a positive integer), so that $n/d \in \mathbb{N}$.

Now, for any $N \in \mathbb{N}$, we have the following chain of logical equivalences:

$$\begin{aligned}(g \mid N) &\iff \left(\underbrace{(z^m)^N}_{=z^{mN}}(x) = x\right) && \text{(by (13.193.13))}\\ &\iff \left(z^{mN}(x) = x\right) \iff (n \mid mN) && \text{(by (13.193.14), applied to } mN \text{ instead of } N\text{)}\\ &\iff (n/d \mid N) \qquad (13.193.15)\end{aligned}$$

(by Lemma 13.193.3, applied to n, m and N instead of x, y and z).

Applying (13.193.15) to $N = n/d$, we obtain the equivalence $(g \mid n/d) \iff (n/d \mid n/d)$ (since $n/d \in \mathbb{N}$). Thus, we have $g \mid n/d$ (since we clearly have $n/d \mid n/d$). Hence, $g \leq n/d$ (since g and n/d are positive integers).

Applying (13.193.15) to $N = g$, we obtain the equivalence $(g \mid g) \iff (n/d \mid g)$ (since $g \in \mathbb{N}$). Thus, we have $n/d \mid g$ (since we clearly have $g \mid g$). Hence, $n/d \leq g$ (since n/d and g are positive integers).

Combining $g \leq n/d$ with $n/d \leq g$, we obtain $g = n/d$. Hence, $|Z| = g = n/d$. In other words, the set Z has size n/d.

Now, forget that we fixed Z. We thus have shown that if Z is a cycle of z^m (considered as a subset of $\{1, 2, \ldots, n\}$), then Z has size n/d. In other words, every cycle of z^m has size n/d.

Hence, Lemma 13.193.4 (applied to $q = n/d$ and $\sigma = z^m$) yields that n/d is a positive divisor of n, and that $\operatorname{type}(z^m) = \left(\underbrace{n/d, n/d, \ldots, n/d}_{n/(n/d) \text{ times}}\right)$. Hence,

$$\operatorname{type}(z^m) = \left(\underbrace{n/d, n/d, \ldots, n/d}_{n/(n/d) \text{ times}}\right) = \left(\underbrace{n/d, n/d, \ldots, n/d}_{d \text{ times}}\right) \qquad (\text{since } n/(n/d) = d).$$

This proves Lemma 13.193.5. □

At last, we can now solve Exercise 6.6.53 itself:

Let us use the notations from Remark 6.6.52(b). In particular, let n be a positive integer. We fix some n-cycle z in $\mathfrak{S}_n$, and we fix some generator g of the cyclic group C_n. We embed the cyclic group $C_n = \mathbb{Z}/n\mathbb{Z}$ as a subgroup in the symmetric group $\mathfrak{S}_n$ by identifying $g \in C_n$ with $z \in \mathfrak{S}_n$. (This is legitimate, since the n-cycle $z \in \mathfrak{S}_n$ is an element of order n.) Let ω be a primitive n-th root of unity in $\mathbb{C}$ (for instance, $\exp(2\pi i/n)$). Let $\gamma : C_n \to \mathbb{C}$ be the character of C_n that sends each $g^i \in C_n$ to ω^i.

We must prove the claim of Remark 6.6.52(b). In other words, we must prove that $\mathbf{GR}_{(n)} = \operatorname{ch}\left(\operatorname{Ind}_{C_n}^{\mathfrak{S}_n} \gamma\right)$.

In the following, μ denotes the number-theoretical Möbius function (defined as in Exercise 2.9.6), and the summation sign "$\sum_{d|n}$" will always be understood to range over all **positive** divisors d of n.

Extend the map $\operatorname{ch} : A = A(\mathfrak{S}) \to \Lambda$ to a $\mathbb{C}$-linear map $A_{\mathbb{C}} \to \Lambda_{\mathbb{C}}$. We shall call the latter map ch, too.

The cyclic group C_n is abelian; thus, every function from C_n to $\mathbb{C}$ is a class function of C_n (because any function from an abelian group to $\mathbb{C}$ is a class function). Hence, γ is a class function of C_n (since γ is a function from C_n to $\mathbb{C}$). In other words, $\gamma \in R_{\mathbb{C}}(C_n)$.

For every permutation $\sigma \in \mathfrak{S}_n$, we let $\operatorname{type}\sigma$ denote the cycle type of σ. Exercise 4.4.5(b) (applied to $H = C_n$ and $f = \gamma$) yields

$$\operatorname{ch}\left(\operatorname{Ind}_{C_n}^{\mathfrak{S}_n} \gamma\right) = \frac{1}{|C_n|} \sum_{h \in C_n} \gamma(h)\, p_{\operatorname{type} h} = \frac{1}{n} \sum_{h \in C_n} \gamma(h)\, p_{\operatorname{type} h} \qquad (13.193.16)$$

(since $|C_n| = n$).

If m and d are integers such that $\gcd(n,m) = d$, then

$$\begin{aligned}\operatorname{type}(g^m) &= \operatorname{type}(z^m) \qquad \text{(since } g = z \text{ (because we are identifying } g \text{ with } z\text{))} \\ &= \Bigg(\underbrace{n/d, n/d, \ldots, n/d}_{d \text{ times}}\Bigg) \\ &\qquad \text{(by Lemma 13.193.5 (since } d = \gcd(n,m) = \gcd(m,n)\text{))}\end{aligned}$$

and thus

$$\begin{aligned}p_{\operatorname{type}(g^m)} &= p_{\left(\underbrace{n/d, n/d, \ldots, n/d}_{d \text{ times}}\right)} = \underbrace{p_{n/d} p_{n/d} \cdots p_{n/d}}_{d \text{ times}} \\ &\qquad \left(\begin{array}{c}\text{by the definition of } p_\lambda \text{ for a partition } \lambda \\ \text{(since } \ell\left(\left(\underbrace{n/d, n/d, \ldots, n/d}_{d \text{ times}}\right)\right) = d\text{)}\end{array}\right) \\ &= p_{n/d}^d. \end{aligned} \tag{13.193.17}$$

But C_n is the cyclic group of size n, and g is a generator of C_n. Hence, the n elements of C_n are $g^1, g^2, \ldots, g^n$ (listed here without repetition). Hence,

$$\begin{aligned}\sum_{h \in C_n} \gamma(h) p_{\operatorname{type} h} &= \gamma(g^1) p_{\operatorname{type}(g^1)} + \gamma(g^2) p_{\operatorname{type}(g^2)} + \cdots + \gamma(g^n) p_{\operatorname{type}(g^n)} \\ &= \underbrace{\sum_{m \in \{1,2,\ldots,n\}}}_{\substack{= \sum_{d \mid n} \sum_{\substack{m \in \{1,2,\ldots,n\}; \\ \gcd(n,m) = d}} \\ \text{(here, we have split up the sum} \\ \text{according to the value of } \gcd(n,m)\text{,} \\ \text{because } \gcd(n,m) \text{ is a positive divisor of } n \\ \text{for each } m \in \{1,2,\ldots,n\}\text{)}}} \gamma(g^m) p_{\operatorname{type}(g^m)} \\ &= \sum_{d \mid n} \sum_{\substack{m \in \{1,2,\ldots,n\}; \\ \gcd(n,m) = d}} \underbrace{\gamma(g^m)}_{\substack{= \omega^m \\ \text{(by the definition of } \gamma\text{)}}} \underbrace{p_{\operatorname{type}(g^m)}}_{\substack{= p_{n/d}^d \\ \text{(by (13.193.17))}}} \\ &= \sum_{d \mid n} \underbrace{\sum_{\substack{m \in \{1,2,\ldots,n\}; \\ \gcd(n,m) = d}} \omega^m}_{\substack{= \sum_{\substack{i \in \{1,2,\ldots,n\}; \\ \gcd(n,i) = d}} \omega^i \\ \text{(here, we have renamed} \\ \text{the summation index } m \text{ as } i\text{)}}} p_{n/d}^d = \sum_{d \mid n} \underbrace{\sum_{\substack{i \in \{1,2,\ldots,n\}; \\ \gcd(n,i) = d}} \omega^i}_{\substack{= \mu(n/d) \\ \text{(by Lemma 13.193.1(b))}}} p_{n/d}^d \\ &= \sum_{d \mid n} \mu(n/d) p_{n/d}^d = \sum_{d \mid n} \mu\Bigg(\underbrace{n/(n/d)}_{= d}\Bigg) \underbrace{p_{n/(n/d)}^{n/d}}_{\substack{= p_d^{n/d} \\ \text{(since } n/(n/d) = d\text{)}}} \\ &\qquad \left(\begin{array}{c}\text{here, we have substituted } n/d \text{ for } d \text{ in the sum, since} \\ \text{the map } \{\text{positive divisors of } n\} \to \{\text{positive divisors of } n\}, \ d \mapsto n/d \\ \text{is a bijection}\end{array}\right) \\ &= \sum_{d \mid n} \mu(d) p_d^{n/d}.\end{aligned}$$

Hence, (13.193.16) becomes

$$\operatorname{ch}\left(\operatorname{Ind}_{C_n}^{\mathfrak{S}_n}\gamma\right)=\frac{1}{n}\underbrace{\sum_{h\in C_n}\gamma\left(h\right)p_{\operatorname{type}h}}_{=\sum_{d\mid n}\mu(d)p_d^{n/d}}=\frac{1}{n}\sum_{d\mid n}\mu\left(d\right)p_d^{n/d}.$$

But Proposition 6.6.36(b) yields

$$\mathbf{GR}_{(n)}=\frac{1}{n}\sum_{d\mid n}\mu\left(d\right)p_d^{n/d}.$$

Comparing these two equalities, we obtain $\mathbf{GR}_{(n)}=\operatorname{ch}\left(\operatorname{Ind}_{C_n}^{\mathfrak{S}_n}\gamma\right)$. This proves the claim of Remark 6.6.52(b). Thus, Exercise 6.6.53 is solved.

13.194. **Solution to Exercise 6.6.55.** *Solution to Exercise 6.6.55.* Before we prove Proposition 6.6.54, let us prove two simple lemmas:

Lemma 13.194.1. *Let $n\in\mathbb{N}$. Then:*

(a) *For any two compositions $\alpha,\beta\in\operatorname{Comp}_n$, we have the logical equivalence*

$$\left(\alpha=\beta\right)\iff\left(D\left(\alpha\right)=D\left(\beta\right)\right).$$

(b) *For any permutation $\sigma\in\mathfrak{S}_n$ and any composition $\beta\in\operatorname{Comp}_n$, we have the logical equivalence*

$$\left(\beta=\gamma\left(\sigma\right)\right)\iff\left(\operatorname{Des}\sigma=D\left(\beta\right)\right).$$

(c) *For any permutation $\sigma\in\mathfrak{S}_n$ and any composition $\beta\in\operatorname{Comp}_n$, we have the logical equivalence*

$$\left(\beta\text{ refines }\gamma\left(\sigma\right)\right)\iff\left(\operatorname{Des}\sigma\subset D\left(\beta\right)\right).$$

Proof of Lemma 13.194.1. If $\sigma\in\mathfrak{S}_n$ is any permutation, then $\gamma\left(\sigma\right)$ is a composition of n and satisfies

$$D\left(\gamma\left(\sigma\right)\right)=\operatorname{Des}\sigma \tag{13.194.1}$$

(since $\gamma\left(\sigma\right)$ was defined to be the unique composition α of n satisfying $D\left(\alpha\right)=\operatorname{Des}\sigma$).

(a) Set $[n-1]:=\{1,2,\ldots,n-1\}$. The map $\operatorname{Comp}_n\to 2^{[n-1]}$, $\alpha\mapsto D\left(\alpha\right)$ is a bijection. Thus, in particular, this map is injective. Hence, for any two compositions $\alpha,\beta\in\operatorname{Comp}_n$, we have the logical equivalence $\left(\alpha=\beta\right)\iff\left(D\left(\alpha\right)=D\left(\beta\right)\right)$. This proves Lemma 13.194.1(a).

(b) Let $\sigma\in\mathfrak{S}_n$ be a permutation. Let $\beta\in\operatorname{Comp}_n$ be a composition. Note that $\gamma\left(\sigma\right)\in\operatorname{Comp}_n$ (since $\sigma\in\mathfrak{S}_n$). Now, we have the following chain of logical equivalences:

$$\begin{aligned}\left(\beta=\gamma\left(\sigma\right)\right)\iff\left(\gamma\left(\sigma\right)=\beta\right)&\iff\left(D\left(\gamma\left(\sigma\right)\right)=D\left(\beta\right)\right)\\&\qquad\left(\text{by Lemma 13.194.1(a) (applied to }\alpha=\gamma\left(\sigma\right)\text{)}\right)\\&\iff\left(\operatorname{Des}\sigma=D\left(\beta\right)\right)\qquad\left(\text{by (13.194.1)}\right).\end{aligned}$$

This proves Lemma 13.194.1(b).

(c) For any two compositions $\alpha,\beta\in\operatorname{Comp}_n$, we have the logical equivalence

$$\left(\alpha\text{ refines }\beta\right)\iff\left(D\left(\alpha\right)\supset D\left(\beta\right)\right) \tag{13.194.2}$$

(by the definition of the relation "refines" in Definition 5.1.10).

Let $\sigma\in\mathfrak{S}_n$ be any permutation. Let $\beta\in\operatorname{Comp}_n$ be a composition. Note that $\gamma\left(\sigma\right)\in\operatorname{Comp}_n$ (since $\sigma\in\mathfrak{S}_n$). Now, we have the following chain of logical equivalences:

$$\begin{aligned}\left(\beta\text{ refines }\gamma\left(\sigma\right)\right)&\iff\left(D\left(\beta\right)\supset D\left(\gamma\left(\sigma\right)\right)\right)\\&\qquad\left(\begin{array}{c}\text{by (13.194.2) (applied to }\beta\text{ and }\gamma\left(\sigma\right)\text{ instead of }\alpha\text{ and }\beta\text{),}\\\text{since }\beta\in\operatorname{Comp}_n\text{ and }\gamma\left(\sigma\right)\in\operatorname{Comp}_n\end{array}\right)\\&\iff\left(D\left(\gamma\left(\sigma\right)\right)\subset D\left(\beta\right)\right)\iff\left(\operatorname{Des}\sigma\subset D\left(\beta\right)\right)\qquad\left(\text{by (13.194.1)}\right).\end{aligned}$$

This proves Lemma 13.194.1(c). □

Lemma 13.194.2. *Let $n \in \mathbb{N}$. Let $\beta = (\beta_1, \beta_2, \ldots, \beta_k) \in \operatorname{Comp}_n$. Let μ be the partition obtained by sorting the entries of β into decreasing order. Then,*

$$h_\mu = \sum_{\substack{\alpha \in \operatorname{Comp}_n; \\ \beta \text{ refines } \alpha}} s_{\operatorname{Rib}(\alpha)}.$$

Proof of Lemma 13.194.2. Consider the algebra homomorphism $\operatorname{NSym} \xrightarrow{\pi} \Lambda$ defined in Corollary 5.4.3.

It is easy to see that

$$\pi(H_\beta) = h_\mu. \tag{13.194.3}$$

[*Proof of (13.194.3):* The definition of μ shows that the nonzero entries of μ are precisely the entries of β (up to order). Since $\beta = (\beta_1, \beta_2, \ldots, \beta_k)$, we can restate this as follows: The nonzero entries of μ are precisely $\beta_1, \beta_2, \ldots, \beta_k$ (up to order). Hence, the definition of h_μ yields

$$h_\mu = h_{\beta_1} h_{\beta_2} \cdots h_{\beta_k}. \tag{13.194.4}$$

But $\beta_1, \beta_2, \ldots, \beta_k$ are positive integers (since $(\beta_1, \beta_2, \ldots, \beta_k) \in \operatorname{Comp}_n \subset \operatorname{Comp}$). Hence, the definition of π yields $\pi(H_{\beta_i}) = h_{\beta_i}$ for all $i \in \{1, 2, \ldots, k\}$. Multiplying these k equalities, we obtain $\pi(H_{\beta_1}) \pi(H_{\beta_2}) \cdots \pi(H_{\beta_k}) = h_{\beta_1} h_{\beta_2} \cdots h_{\beta_k}$.

Recall again that $\beta = (\beta_1, \beta_2, \ldots, \beta_k)$. Hence, (5.4.3) (applied to β and k instead of α and ℓ) yields $H_\beta = H_{\beta_1} H_{\beta_2} \cdots H_{\beta_k}$. Applying the map π to both sides of this equality, we find

$$\begin{aligned}\pi(H_\beta) &= \pi(H_{\beta_1} H_{\beta_2} \cdots H_{\beta_k}) = \pi(H_{\beta_1}) \pi(H_{\beta_2}) \cdots \pi(H_{\beta_k}) \qquad (\text{since } \pi \text{ is a } \mathbf{k}\text{-algebra homomorphism}) \\ &= h_{\beta_1} h_{\beta_2} \cdots h_{\beta_k}.\end{aligned}$$

Comparing this with (13.194.4), we find $\pi(H_\beta) = h_\mu$. Thus, (13.194.3) is proved.]

Theorem 5.4.10(b) shows that

$$\pi(R_\alpha) = s_{\operatorname{Rib}(\alpha)} \tag{13.194.5}$$

for each composition α.

The equality (5.4.9) (with the variables α and β renamed as β and α) says that

$$H_\beta = \underbrace{\sum_{\alpha \text{ coarsens } \beta}}_{\substack{= \sum_{\substack{\alpha \in \operatorname{Comp}_n; \\ \alpha \text{ coarsens } \beta}} = \sum_{\substack{\alpha \in \operatorname{Comp}_n; \\ \beta \text{ refines } \alpha}} \\ (\text{since the condition ``}\alpha \text{ coarsens } \beta\text{''} \\ \text{is equivalent to ``}\beta \text{ refines } \alpha\text{''})}} R_\alpha = \sum_{\substack{\alpha \in \operatorname{Comp}_n; \\ \beta \text{ refines } \alpha}} R_\alpha.$$

Applying the map π to both sides of this equality, we obtain

$$\begin{aligned}\pi(H_\beta) &= \pi\left(\sum_{\substack{\alpha \in \operatorname{Comp}_n; \\ \beta \text{ refines } \alpha}} R_\alpha\right) = \sum_{\substack{\alpha \in \operatorname{Comp}_n; \\ \beta \text{ refines } \alpha}} \underbrace{\pi(R_\alpha)}_{\substack{= s_{\operatorname{Rib}(\alpha)} \\ (\text{by } (13.194.5))}} \qquad (\text{since the map } \pi \text{ is } \mathbf{k}\text{-linear}) \\ &= \sum_{\substack{\alpha \in \operatorname{Comp}_n; \\ \beta \text{ refines } \alpha}} s_{\operatorname{Rib}(\alpha)}.\end{aligned}$$

Comparing this with (13.194.3), we obtain

$$h_\mu = \sum_{\substack{\alpha \in \operatorname{Comp}_n; \\ \beta \text{ refines } \alpha}} s_{\operatorname{Rib}(\alpha)}.$$

This proves Lemma 13.194.2. □

Proof of Proposition 6.6.54. Proposition 6.6.37 shows that the power series $\mathbf{GR}_\lambda$ belongs to Λ. In other words, $\mathbf{GR}_\lambda \in \Lambda$.

(b) For each permutation $\sigma \in \mathfrak{S}_n$, we have the logical equivalence

$$(\beta = \gamma(\sigma)) \iff (\operatorname{Des} \sigma = D(\beta))$$

(by Lemma 13.194.1(b)). Hence,

$$\begin{aligned}&\left(\text{the number of permutations } \sigma\in\mathfrak{S}_n \text{ satisfying } \operatorname{type}\sigma=\lambda \text{ and } \beta=\gamma\left(\sigma\right)\right)\\ &=\left(\text{the number of permutations } \sigma\in\mathfrak{S}_n \text{ satisfying } \operatorname{type}\sigma=\lambda \text{ and } \operatorname{Des}\sigma=D\left(\beta\right)\right).\end{aligned} \tag{13.194.6}$$

The definition of the noncommutative ribbon functions shows that the basis $\left(R_\alpha\right)_{\alpha\in\operatorname{Comp}}$ of NSym is dual to the basis $\left(L_\alpha\right)_{\alpha\in\operatorname{Comp}}$ of QSym (with respect to the dual pairing $\operatorname{NSym}\otimes\operatorname{QSym}\xrightarrow{(\cdot,\cdot)}\mathbf{k}$). In other words,

$$\left(R_\alpha,L_\zeta\right)=\delta_{\alpha,\zeta} \qquad \text{for all } \alpha,\zeta\in\operatorname{Comp}. \tag{13.194.7}$$

Now, consider the inclusion $\Lambda\overset{i}{\hookrightarrow}\operatorname{QSym}$ and the algebra homomorphism $\operatorname{NSym}\xrightarrow{\pi}\Lambda$ defined in Corollary 5.4.3. We have seen in Corollary 5.4.3 that these two maps i and π are adjoint (with respect to the dual pairing $\operatorname{NSym}\otimes\operatorname{QSym}\xrightarrow{(\cdot,\cdot)}\mathbf{k}$). In other words,

$$\left(f,i\left(g\right)\right)=\left(\pi\left(f\right),g\right) \qquad \text{for any } f\in\operatorname{NSym} \text{ and } g\in\Lambda. \tag{13.194.8}$$

Theorem 5.4.10(b) shows that $\pi\left(R_\alpha\right)=s_{\operatorname{Rib}(\alpha)}$ for every composition α. Applying this to $\alpha=\beta$, we find $\pi\left(R_\beta\right)=s_{\operatorname{Rib}(\beta)}$.

Recall that $\mathbf{GR}_\lambda\in\Lambda$. Thus, (13.194.8) (applied to $f=R_\beta$ and $g=\mathbf{GR}_\lambda$) yields

$$\left(R_\beta,i\left(\mathbf{GR}_\lambda\right)\right)=\left(\underbrace{\pi\left(R_\beta\right)}_{=s_{\operatorname{Rib}(\beta)}},\mathbf{GR}_\lambda\right)=\left(s_{\operatorname{Rib}(\beta)},\mathbf{GR}_\lambda\right)=\left(\mathbf{GR}_\lambda,s_{\operatorname{Rib}(\beta)}\right) \tag{13.194.9}$$

(since the Hall inner product is symmetric).

But i is just an inclusion map. Hence,

$$i\left(\mathbf{GR}_\lambda\right)=\mathbf{GR}_\lambda=\sum_{\substack{\sigma\in\mathfrak{S}_n;\\ \sigma \text{ has cycle type } \lambda}} L_{\gamma(\sigma)} \qquad \text{(by Proposition 6.6.40)}.$$

Now, (13.194.9) shows that

$$\left(\mathbf{GR}_\lambda, s_{\mathrm{Rib}(\beta)}\right) = \left(R_\beta, \underbrace{i\left(\mathbf{GR}_\lambda\right)}_{=\sum\limits_{\substack{\sigma\in\mathfrak{S}_n;\\ \sigma \text{ has cycle type } \lambda}} L_{\gamma(\sigma)}}\right) = \left(R_\beta, \sum_{\substack{\sigma\in\mathfrak{S}_n;\\ \sigma \text{ has cycle type } \lambda}} L_{\gamma(\sigma)}\right)$$

$$= \underbrace{\sum_{\substack{\sigma\in\mathfrak{S}_n;\\ \sigma \text{ has cycle type } \lambda}}}_{\substack{=\sum\limits_{\substack{\sigma\in\mathfrak{S}_n;\\ \operatorname{type}\sigma=\lambda}}\\ \text{(since } \operatorname{type}\sigma \text{ denotes the cycle type of } \sigma)}} \underbrace{\left(R_\beta, L_{\gamma(\sigma)}\right)}_{\substack{=\delta_{\beta,\gamma(\sigma)}\\ \text{(by (13.194.7),}\\ \text{applied to } \alpha=\beta\\ \text{and } \zeta=\gamma(\sigma))}}$$

$$\left(\text{since the dual pairing } \mathrm{NSym}\otimes\mathrm{QSym} \xrightarrow{(\cdot,\cdot)} \mathbf{k} \text{ is } \mathbf{k}\text{-bilinear}\right)$$

$$= \sum_{\substack{\sigma\in\mathfrak{S}_n;\\ \operatorname{type}\sigma=\lambda}} \delta_{\beta,\gamma(\sigma)} = \sum_{\substack{\sigma\in\mathfrak{S}_n;\\ \operatorname{type}\sigma=\lambda;\\ \beta=\gamma(\sigma)}} \underbrace{\delta_{\beta,\gamma(\sigma)}}_{\substack{=1\\ \text{(since } \beta=\gamma(\sigma))}} + \sum_{\substack{\sigma\in\mathfrak{S}_n;\\ \operatorname{type}\sigma=\lambda;\\ \beta\neq\gamma(\sigma)}} \underbrace{\delta_{\beta,\gamma(\sigma)}}_{\substack{=0\\ \text{(since } \beta\neq\gamma(\sigma))}}$$

$$\left(\begin{array}{c}\text{since every } \sigma\in\mathfrak{S}_n \text{ satisfies either } \beta=\gamma(\sigma) \text{ or } \beta\neq\gamma(\sigma)\\ \text{(but not both at the same time)}\end{array}\right)$$

$$= \sum_{\substack{\sigma\in\mathfrak{S}_n;\\ \operatorname{type}\sigma=\lambda;\\ \beta=\gamma(\sigma)}} 1 + \underbrace{\sum_{\substack{\sigma\in\mathfrak{S}_n;\\ \operatorname{type}\sigma=\lambda;\\ \beta\neq\gamma(\sigma)}} 0}_{=0} = \sum_{\substack{\sigma\in\mathfrak{S}_n;\\ \operatorname{type}\sigma=\lambda;\\ \beta=\gamma(\sigma)}} 1$$

$$= \left(\text{the number of permutations } \sigma\in\mathfrak{S}_n \text{ satisfying } \operatorname{type}\sigma=\lambda \text{ and } \beta=\gamma(\sigma)\right)\cdot 1$$

$$= \left(\text{the number of permutations } \sigma\in\mathfrak{S}_n \text{ satisfying } \operatorname{type}\sigma=\lambda \text{ and } \beta=\gamma(\sigma)\right).$$

Combining this with (13.194.6), we obtain

$$\begin{aligned}&\left(\text{the number of permutations } \sigma\in\mathfrak{S}_n \text{ satisfying } \operatorname{type}\sigma=\lambda \text{ and } \beta=\gamma(\sigma)\right)\\ &= \left(\text{the number of permutations } \sigma\in\mathfrak{S}_n \text{ satisfying } \operatorname{type}\sigma=\lambda \text{ and } \operatorname{Des}\sigma=D(\beta)\right)\\ &= \left(\mathbf{GR}_\lambda, s_{\mathrm{Rib}(\beta)}\right).\end{aligned}$$

This proves Proposition 6.6.54(b).

(a) For each permutation $\sigma\in\mathfrak{S}_n$, we have the logical equivalence

$$\left(\beta \text{ refines } \gamma(\sigma)\right) \iff \left(\operatorname{Des}\sigma\subset D(\beta)\right)$$

(by Lemma 13.194.1(c)). Hence,

$$\begin{aligned}&\left(\text{the number of permutations } \sigma\in\mathfrak{S}_n \text{ satisfying } \operatorname{type}\sigma=\lambda \text{ such that } \beta \text{ refines } \gamma(\sigma)\right)\\ &= \left(\text{the number of permutations } \sigma\in\mathfrak{S}_n \text{ satisfying } \operatorname{type}\sigma=\lambda \text{ and } \operatorname{Des}\sigma\subset D(\beta)\right).\end{aligned}\tag{13.194.10}$$

On the other hand, if $\alpha\in\mathrm{Comp}_n$ is a composition, then

$$\begin{aligned}&\left(\text{the number of permutations } \sigma\in\mathfrak{S}_n \text{ satisfying } \operatorname{type}\sigma=\lambda \text{ and } \gamma(\sigma)=\alpha\right)\\ &= \left(\mathbf{GR}_\lambda, s_{\mathrm{Rib}(\alpha)}\right).\end{aligned}\tag{13.194.11}$$

[*Proof of (13.194.11):* Let $\alpha\in\mathrm{Comp}_n$ be a composition. Write α in the form $\alpha=(\alpha_1,\alpha_2,\ldots,\alpha_\ell)$. Thus, $\alpha=(\alpha_1,\alpha_2,\ldots,\alpha_\ell)\in\mathrm{Comp}_n$. Hence, Proposition 6.6.54(b) (applied to α and ℓ instead of β and k) yields

$$\begin{aligned}&\left(\text{the number of permutations } \sigma\in\mathfrak{S}_n \text{ satisfying } \operatorname{type}\sigma=\lambda \text{ and } \alpha=\gamma(\sigma)\right)\\ &= \left(\text{the number of permutations } \sigma\in\mathfrak{S}_n \text{ satisfying } \operatorname{type}\sigma=\lambda \text{ and } \operatorname{Des}\sigma=D(\alpha)\right)\\ &= \left(\mathbf{GR}_\lambda, s_{\mathrm{Rib}(\alpha)}\right) \qquad \left(\text{this is the Hall inner product of } \mathbf{GR}_\lambda\in\Lambda \text{ and } s_{\mathrm{Rib}(\alpha)}\in\Lambda\right).\end{aligned}$$

But this immediately yields (13.194.11) (since the condition "$\gamma(\sigma) = \alpha$" is equivalent to "$\alpha = \gamma(\sigma)$").]

If $\sigma \in \mathfrak{S}_n$ is any permutation, then $\gamma(\sigma) \in \operatorname{Comp}_n$ (by the definition of $\gamma(\sigma)$). Hence,

$$\begin{aligned} &(\text{the number of permutations } \sigma \in \mathfrak{S}_n \text{ satisfying } \operatorname{type}\sigma = \lambda \text{ such that } \beta \text{ refines } \gamma(\sigma)) \\ &= \sum_{\substack{\alpha \in \operatorname{Comp}_n; \\ \beta \text{ refines } \alpha}} \underbrace{(\text{the number of permutations } \sigma \in \mathfrak{S}_n \text{ satisfying } \operatorname{type}\sigma = \lambda \text{ such that } \gamma(\sigma) = \alpha)}_{\substack{= \left(\mathbf{GR}_\lambda, s_{\operatorname{Rib}(\alpha)}\right) \\ (\text{by (13.194.11)})}} \\ &= \sum_{\substack{\alpha \in \operatorname{Comp}_n; \\ \beta \text{ refines } \alpha}} \left(\mathbf{GR}_\lambda, s_{\operatorname{Rib}(\alpha)}\right). \end{aligned}$$

Comparing this with

$$\left(\mathbf{GR}_\lambda, \underbrace{h_\mu}_{\substack{= \sum_{\substack{\alpha \in \operatorname{Comp}_n; \\ \beta \text{ refines } \alpha}} s_{\operatorname{Rib}(\alpha)} \\ (\text{by Lemma 13.194.2})}}\right) = \left(\mathbf{GR}_\lambda, \sum_{\substack{\alpha \in \operatorname{Comp}_n; \\ \beta \text{ refines } \alpha}} s_{\operatorname{Rib}(\alpha)}\right) = \sum_{\substack{\alpha \in \operatorname{Comp}_n; \\ \beta \text{ refines } \alpha}} \left(\mathbf{GR}_\lambda, s_{\operatorname{Rib}(\alpha)}\right)$$
$$(\text{since the Hall inner product is } \mathbf{k}\text{-bilinear}),$$

we obtain

$$\begin{aligned} &(\text{the number of permutations } \sigma \in \mathfrak{S}_n \text{ satisfying } \operatorname{type}\sigma = \lambda \text{ such that } \beta \text{ refines } \gamma(\sigma)) \\ &= (\mathbf{GR}_\lambda, h_\mu). \end{aligned} \tag{13.194.12}$$

Recall that $\mathbf{GR}_\lambda \in \Lambda$; also, μ is a weak composition (since μ is a partition). Furthermore, μ is a partition; thus, all the nonzero entries of μ are concentrated at the beginning of μ, and followed by an infinite string of zeroes. Hence, μ is the partition consisting of the nonzero entries of μ (sorted in decreasing order)[1229]. Therefore, Exercise 2.5.25 (applied to $\mathbf{GR}_\lambda$ and μ instead of f and β) yields

$$(\mathbf{GR}_\lambda, h_\mu) = (h_\mu, \mathbf{GR}_\lambda) = (\text{the coefficient of } \mathbf{x}^\mu \text{ in } \mathbf{GR}_\lambda). \tag{13.194.13}$$

Finally, let δ be the weak composition $(\beta_1, \beta_2, \ldots, \beta_k, 0, 0, 0, \ldots)$. Then, $\mathbf{x}^\delta = x_1^{\beta_1} x_2^{\beta_2} \cdots x_k^{\beta_k}$. Furthermore, the nonzero entries of δ are precisely the numbers $\beta_1, \beta_2, \ldots, \beta_k$ (since $(\beta_1, \beta_2, \ldots, \beta_k) \in \operatorname{Comp}$ shows that $\beta_1, \beta_2, \ldots, \beta_k$ are positive integers and thus nonzero). In other words, the nonzero entries of δ are precisely the entries of β (since $\beta = (\beta_1, \beta_2, \ldots, \beta_k)$). Hence, μ is the partition consisting of the nonzero entries of δ (sorted in decreasing order)[1230]. Therefore, Exercise 2.5.25 (applied to $\mathbf{GR}_\lambda$ and δ instead of f and δ) yields

$$(\mathbf{GR}_\lambda, h_\mu) = (h_\mu, \mathbf{GR}_\lambda) = \left(\text{the coefficient of } \mathbf{x}^\delta \text{ in } \mathbf{GR}_\lambda\right).$$

In view of $\mathbf{x}^\delta = x_1^{\beta_1} x_2^{\beta_2} \cdots x_k^{\beta_k}$, this rewrites as

$$(\mathbf{GR}_\lambda, h_\mu) = (h_\mu, \mathbf{GR}_\lambda) = \left(\text{the coefficient of } x_1^{\beta_1} x_2^{\beta_2} \cdots x_k^{\beta_k} \text{ in } \mathbf{GR}_\lambda\right).$$

Combining this equality with the equalities (13.194.10), (13.194.12) and (13.194.13), we obtain

$$\begin{aligned} &(\text{the number of permutations } \sigma \in \mathfrak{S}_n \text{ satisfying } \operatorname{type}\sigma = \lambda \text{ such that } \beta \text{ refines } \gamma(\sigma)) \\ &= (\text{the number of permutations } \sigma \in \mathfrak{S}_n \text{ satisfying } \operatorname{type}\sigma = \lambda \text{ and } \operatorname{Des}\sigma \subset D(\beta)) \\ &= \left(\text{the coefficient of } x_1^{\beta_1} x_2^{\beta_2} \cdots x_k^{\beta_k} \text{ in } \mathbf{GR}_\lambda\right) = (\text{the coefficient of } \mathbf{x}^\mu \text{ in } \mathbf{GR}_\lambda) \\ &= (\mathbf{GR}_\lambda, h_\mu) \qquad (\text{this is the Hall inner product of } \mathbf{GR}_\lambda \in \Lambda \text{ and } h_\mu \in \Lambda). \end{aligned}$$

This proves Proposition 6.6.54(a). □

[1229] Indeed, no sorting is required: The partition μ simply consists of the nonzero entries of μ.

[1230] Indeed, the definition of μ shows that μ is the partition consisting of the entries of β (sorted in decreasing order). But since the nonzero entries of δ are precisely the entries of β, we can rewrite this as follows: μ is the partition consisting of the nonzero entries of δ (sorted in decreasing order).

Thus, Proposition 6.6.54(a) is proved, so that Exercise 6.6.55 is solved.

13.195. **Solution to Exercise 7.1.9.** *Solution to Exercise 7.1.9.* Fix $m \in \{0, 1, 2, \ldots\}$. In order to check that $\zeta_Q^{\star m}(f) = \operatorname{ps}^1(f)(m)$, it is clearly enough to show that $\zeta_Q^{\star m}(M_\alpha) = \operatorname{ps}^1(M_\alpha)(m)$ for all compositions α. So let α be a composition. Iterated application of Proposition 5.1.7 yields

$$\Delta^{(m-1)} M_\alpha = \sum_{\substack{(\beta_1,\beta_2,\ldots,\beta_m):\\ \beta_1\beta_2\cdots\beta_m=\alpha}} M_{\beta_1} \otimes M_{\beta_2} \otimes \cdots \otimes M_{\beta_m}.$$

An addend on the right hand side of this equality is annihilated by the map $\zeta_Q^{\otimes m} : \operatorname{QSym}^{\otimes m} \to \mathbf{k}^{\otimes m} \cong \mathbf{k}$ unless each of the compositions $\beta_1, \beta_2, \ldots, \beta_m$ has length ≤ 1; all remaining terms are mapped to $1 \cdot 1 \cdots 1 = 1$. Hence,

$$\zeta_Q^{\otimes m}\left(\Delta^{(m-1)} M_\alpha\right) = \sum_{\substack{(\beta_1,\beta_2,\ldots,\beta_m):\\ \beta_1\beta_2\cdots\beta_m=\alpha;\\ \text{each of the compositions } \beta_1,\beta_2,\ldots,\beta_m \text{ has length } \leq 1}} 1 = \binom{m}{\ell},$$

where ℓ is the length of α. Hence,

$$\zeta_Q^{\star m}(M_\alpha) = \zeta_Q^{\otimes m}\left(\Delta^{(m-1)} M_\alpha\right) = \binom{m}{\ell} = \operatorname{ps}^1(M_\alpha)(m)$$

(by Proposition 7.1.7(i)). This completes the proof.

13.196. **Solution to Exercise 7.3.14.** *Solution to Exercise 7.3.14.*

Proof of Proposition 7.3.9. We recall the following fundamental fact (a version of inclusion-exclusion or a very simple special case of Möbius inversion): If R is a finite set, then

$$\sum_{T \subset R} (-1)^{|T|} = \delta_{R,\varnothing}. \tag{13.196.1}$$

[1231] We can use this to show a slightly more complicated fact: If P and R are two finite sets, then

$$\sum_{\substack{F \subset R;\\ F \supset P}} (-1)^{|F \setminus P|} = \delta_{P,R}. \tag{13.196.2}$$

[1231] For the sake of completeness, here is a short *proof of (13.196.1):* Let R be a finite set. We have

$$\sum_{T\subset R}(-1)^{|T|} = \sum_{k\in\mathbb{N}}\sum_{\substack{T\subset R;\\ |T|=k}} \underbrace{(-1)^{|T|}}_{\substack{=(-1)^k\\ (\text{since } |T|=k)}} = \sum_{k\in\mathbb{N}} \underbrace{\sum_{\substack{T\subset R;\\ |T|=k}}(-1)^k}_{=|\{T\subset R \ \mid\ |T|=k\}|(-1)^k} = \sum_{k\in\mathbb{N}} \underbrace{|\{T\subset R \ \mid\ |T|=k\}|}_{\substack{=(\text{the number of all } k\text{-element subsets of } R)\\ =\binom{|R|}{k}\\ (\text{by the combinatorial interpretation of}\\ \text{binomial coefficients})}} (-1)^k$$

$$= \sum_{k\in\mathbb{N}} \binom{|R|}{k} (-1)^k = \Big(\underbrace{1+(-1)}_{=0}\Big)^{|R|}$$

$$\left(\text{since } (1+(-1))^{|R|} = \sum_{k\in\mathbb{N}} \binom{|R|}{k} (-1)^k \text{ (by the binomial formula)}\right)$$

$$= 0^{|R|} = \delta_{|R|,0} = \delta_{R,\varnothing} \qquad (\text{since } |R| = 0 \text{ holds if and only if } R = \varnothing).$$

This proves (13.196.1).

[1232]

(a) Let $G=(V,E)$ be a finite graph. We have

$$\sum_{\substack{H=(V,E');\\ E'\cap E=\varnothing}} [H]^\sharp = \underbrace{\sum_{\substack{K=(V,F);\\ F\cap E=\varnothing}}}_{\substack{=\sum_{\substack{K=(V,F);\\ F\subset E^c}} \\ \text{(since } F\cap E=\varnothing \text{ is} \\ \text{equivalent to } F\subset E^c\text{)}}} \underbrace{[K]^\sharp}_{\substack{=\sum_{\substack{H=(V,E');\\ E'\supset F^c}} (-1)^{\left|E'\setminus F^c\right|}[H] \\ \text{(by the definition of } [K]^\sharp\text{)}}}$$

(here, we renamed H and E' as K and F in the sum)

$$= \underbrace{\sum_{\substack{K=(V,F);\\ F\subset E^c}} \sum_{\substack{H=(V,E');\\ E'\supset F^c}}}_{=\sum_{H=(V,E')} \sum_{\substack{K=(V,F);\\ F\subset E^c;\\ E'\supset F^c}}} \underbrace{(-1)^{\left|E'\setminus F^c\right|}}_{\substack{=(-1)^{\left|F\setminus (E')^c\right|} \\ \text{(since } E'\setminus F^c=E'\cup F=F\setminus \left(E'\right)^c\text{)}}} [H]$$

$$= \sum_{H=(V,E')} \underbrace{\sum_{\substack{K=(V,F);\\ F\subset E^c;\\ E'\supset F^c}}}_{\substack{=\sum_{\substack{F\subset E^c;\\ E'\supset F^c}} = \sum_{\substack{F\subset E^c;\\ F\supset \left(E'\right)^c}} \\ \text{(since } E'\supset F^c \text{ is equivalent to } F\supset \left(E'\right)^c\text{)}}} (-1)^{\left|F\setminus \left(E'\right)^c\right|} [H]$$

$$= \sum_{H=(V,E')} \underbrace{\sum_{\substack{F\subset E^c;\\ F\supset \left(E'\right)^c}} (-1)^{\left|F\setminus \left(E'\right)^c\right|}}_{\substack{=\delta_{(E')^c,E^c} \\ \text{(by (13.196.2), applied to} \\ R=E^c \text{ and } P=\left(E'\right)^c\text{)}}} [H]$$

$$= \sum_{H=(V,E')} \underbrace{\delta_{(E')^c,E^c}}_{\substack{=\delta_{E',E} \\ \text{(since } \left(E'\right)^c=E^c \text{ holds if} \\ \text{and only if } E'=E\text{)}}} [H] = \sum_{H=(V,E')} \delta_{E',E}\,[H] = [(V,E)] = [G]\,.$$

[1232] *Proof of (13.196.2):* Let P and R be two finite sets.

Let us first assume that $P\not\subset R$. Then, there exists no $F\subset R$ such that $F\supset P$ (because if such a F would exist, then we would have $P\subset F\subset R$, which would contradict $P\not\subset R$). Hence, the sum $\sum_{\substack{F\subset R;\\ F\supset P}} (-1)^{|F\setminus P|}$ is empty, and thus it simplifies to $\sum_{\substack{F\subset R;\\ F\supset P}} (-1)^{|F\setminus P|}=0$. On the other hand, $P\not\subset R$, so that $P\neq R$ and thus $\delta_{P,R}=0$. Hence, $\sum_{\substack{F\subset R;\\ F\supset P}} (-1)^{|F\setminus P|}=0=\delta_{P,R}$. Thus, (13.196.2) is proven under the assumption that $P\not\subset R$.

Now, let us forget that we have assumed $P\not\subset R$. We thus have shown that (13.196.2) holds if $P\not\subset R$. Hence, for the rest of this proof, we can WLOG assume that we don't have $P\not\subset R$. Assume this.

We have $P\subset R$ (since we don't have $P\not\subset R$). Hence, there exists a bijection from the set of all $F\subset R$ satisfying $F\supset P$ to the set of all $T\subset R\setminus P$; this bijection sends every F to $F\setminus P$. Hence, we can substitute T for $F\setminus P$ in the sum $\sum_{\substack{F\subset R;\\ F\supset P}} (-1)^{|F\setminus P|}$. We thus obtain

$$\sum_{\substack{F\subset R;\\ F\supset P}} (-1)^{|F\setminus P|} = \sum_{T\subset R\setminus P} (-1)^{|T|} = \delta_{R\setminus P,\varnothing} \qquad \text{(by (13.196.1), applied to } R\setminus P \text{ instead of } R\text{)}$$
$$= \delta_{P,R} \qquad \text{(since } R\setminus P=\varnothing \text{ if and only if } P=R \text{ (because } P\subset R\text{))}\,.$$

This proves (13.196.2).

This proves Proposition 7.3.9(a).

(b) For every finite graph $G = (V, E)$, we define the **complement** of this graph G to be the graph (V, E^c) (where E^c is defined as in Definition 7.3.8). We shall denote this complement by $\mathbf{c}(G)$.

The complement of the complement of a graph G is G again. In other words,

$$\mathbf{c}(\mathbf{c}(G)) = G \qquad \text{for every finite graph } G. \tag{13.196.3}$$

Recall that if G is a finite graph, then we denote the isomorphism class of this graph G by $[G]$. Let GrIs be the set of all isomorphism classes of finite graphs. For every $n \in \mathbb{N}$, let GrIs_n be the set of all isomorphism classes of finite graphs with n vertices. Thus, $\operatorname{GrIs} = \bigsqcup_{n\in\mathbb{N}} \operatorname{GrIs}_n$.

Fix $n \in \mathbb{N}$. The family $([G])_{[G]\in\operatorname{GrIs}_n}$ is a basis of the $\mathbf{k}$-module $\mathcal{G}_n$ (by the definition of the grading on $\mathcal{G}$). We are going to prove that the family $\left([G]^\sharp\right)_{[G]\in\operatorname{GrIs}_n}$ is a basis of the $\mathbf{k}$-module $\mathcal{G}_n$ as well.

We define a map $\mathbf{c}_n : \operatorname{GrIs}_n \to \operatorname{GrIs}_n$ by setting

$$\left(\mathbf{c}_n [G] = [\mathbf{c}(G)] \qquad \text{for every } [G] \in \operatorname{GrIs}_n\right).$$

[1233]

We have $\mathbf{c}_n \circ \mathbf{c}_n = \mathrm{id}$ [1234]. Hence, the maps $\mathbf{c}_n$ and $\mathbf{c}_n$ are mutually inverse. Thus, the map $\mathbf{c}_n$ is invertible, i.e., a bijection. Therefore, the family $\left([G]^\sharp\right)_{[G]\in\operatorname{GrIs}_n}$ is a reindexing of the family $\left((\mathbf{c}_n [G])^\sharp\right)_{[G]\in\operatorname{GrIs}_n}$.

If $G = (V, E)$ is a finite graph with n vertices, then

$$(\mathbf{c}_n [G])^\sharp = \sum_{\substack{H=(V,E');\\ E'\supset E}} (-1)^{\left|E'\setminus E\right|} [H] \tag{13.196.4}$$

[1235].

We define a map $e : \operatorname{GrIs}_n \to \mathbb{N}$ by setting

$$e[(V, E)] = |E| \qquad \text{for every } [(V, E)] \in \operatorname{GrIs}_n .$$

[1236] Thus, the map e sends the isomorphism class of a graph to the number of edges of this graph.

[1233] This map $\mathbf{c}_n$ is well-defined because of the following two simple observations:

- For each $[G] \in \operatorname{GrIs}_n$, we have $[\mathbf{c}(G)] \in \operatorname{GrIs}_n$ (because if the graph G has n vertices, then so does the graph $\mathbf{c}(G)$).
- For each graph G with n vertices, the isomorphism class $[\mathbf{c}(G)]$ depends only on the isomorphism class $[G]$, not on the graph G itself (i.e., if G_1 and G_2 are two isomorphic graphs with n vertices, then the graphs $\mathbf{c}(G_1)$ and $\mathbf{c}(G_2)$ are also isomorphic).

[1234] *Proof.* Let $U \in \operatorname{GrIs}_n$. Thus, U is an isomorphism class of a finite graph with n vertices (since GrIs_n is the set of all isomorphism classes of finite graphs with n vertices). In other words, there exists a graph G with n vertices such that $U = [G]$. Consider this G. Now, $\mathbf{c}_n \underbrace{U}_{=[G]} = \mathbf{c}_n [G] = [\mathbf{c}(G)]$ (by the definition of $\mathbf{c}_n$). Now,

$$\begin{aligned}(\mathbf{c}_n \circ \mathbf{c}_n) U &= \mathbf{c}_n \left(\underbrace{\mathbf{c}_n U}_{=[\mathbf{c}(G)]}\right) = \mathbf{c}_n [\mathbf{c}(G)] = \left[\underbrace{\mathbf{c}(\mathbf{c}(G))}_{\substack{=G\\ (\text{by } (13.196.3))}}\right] \qquad (\text{by the definition of } \mathbf{c}_n)\\ &= [G] = U = \mathrm{id}\, U.\end{aligned}$$

Now, forget that we fixed U. We thus have proven that $(\mathbf{c}_n \circ \mathbf{c}_n) U = \mathrm{id}\, U$ for every $U \in \operatorname{GrIs}_n$. In other words, $\mathbf{c}_n \circ \mathbf{c}_n = \mathrm{id}$. Qed.

[1235] *Proof of (13.196.4):* Let $G = (V, E)$ be a finite graph with n vertices. Then, $\mathbf{c}(G) = (V, E^c)$ (by the definition of $\mathbf{c}(G)$). Hence, the definition of $[\mathbf{c}(G)]^\sharp$ yields

$$[\mathbf{c}(G)]^\sharp = \sum_{\substack{H=(V,E');\\ E'\supset(E^c)^c}} (-1)^{\left|E'\setminus(E^c)^c\right|} [H] = \sum_{\substack{H=(V,E');\\ E'\supset E}} (-1)^{\left|E'\setminus E\right|} [H]$$

(since $(E^c)^c = E$). Now, the definition of $\mathbf{c}_n$ yields $\mathbf{c}_n [G] = [\mathbf{c}(G)]$, and therefore $(\mathbf{c}_n [G])^\sharp = [\mathbf{c}(G)]^\sharp = \sum_{\substack{H=(V,E');\\ E'\supset E}} (-1)^{\left|E'\setminus E\right|} [H]$. This proves (13.196.4).

[1236] This map is well-defined, because for any finite graph (V, E), the number $|E|$ depends only on the isomorphism class $[(V, E)]$ of (V, E) (and not on (V, E) itself).

We define a binary relation $\prec$ on the set GrIs_n as follows: For two elements $[H]$ and $[G]$ of GrIs_n, we set $[H] \prec [G]$ if and only if $e[H] > e[G]$. It is clear that this binary relation $\prec$ is transitive, asymmetric and irreflexive. Thus, there is a partial order on the set GrIs_n whose smaller relation is $\prec$. Consider GrIs_n as a poset, equipped with this partial order.

Now, every finite graph G with n vertices satisfies

$$(\mathbf{c}_n [G])^\sharp = [G] + (\text{a } \mathbf{k}\text{-linear combination of the elements } [H] \text{ for } [H] \in \operatorname{GrIs}_n \text{ satisfying } [H] \prec [G]) \tag{13.196.5}$$

[1237]. In other words, the family $\left((\mathbf{c}_n [G])^\sharp\right)_{[G]\in\operatorname{GrIs}_n}$ expands unitriangularly in the family $([G])_{[G]\in\operatorname{GrIs}_n}$ (by Remark 11.1.17(c), applied to $\mathcal{G}_n$, GrIs_n, $\left((\mathbf{c}_n [G])^\sharp\right)_{[G]\in\operatorname{GrIs}_n}$ and $([G])_{[G]\in\operatorname{GrIs}_n}$ instead of M, S, $(e_s)_{s\in S}$ and $(f_s)_{s\in S}$). Hence, the family $\left((\mathbf{c}_n [G])^\sharp\right)_{[G]\in\operatorname{GrIs}_n}$ is a basis of the $\mathbf{k}$-module $\mathcal{G}_n$ if and only if the family $([G])_{[G]\in\operatorname{GrIs}_n}$ is a basis of the $\mathbf{k}$-module $\mathcal{G}_n$ (by Corollary 11.1.19(e), applied to $\mathcal{G}_n$, GrIs_n, $\left((\mathbf{c}_n [G])^\sharp\right)_{[G]\in\operatorname{GrIs}_n}$ and $([G])_{[G]\in\operatorname{GrIs}_n}$ instead of M, S, $(e_s)_{s\in S}$ and $(f_s)_{s\in S}$). Thus, the family $\left((\mathbf{c}_n [G])^\sharp\right)_{[G]\in\operatorname{GrIs}_n}$ is a basis of the $\mathbf{k}$-module $\mathcal{G}_n$ (since the family $([G])_{[G]\in\operatorname{GrIs}_n}$ is a basis of the $\mathbf{k}$-module $\mathcal{G}_n$). Therefore, the family $\left([G]^\sharp\right)_{[G]\in\operatorname{GrIs}_n}$ also is a basis of the $\mathbf{k}$-module $\mathcal{G}_n$ (since the family $\left([G]^\sharp\right)_{[G]\in\operatorname{GrIs}_n}$ is a reindexing of the family $\left((\mathbf{c}_n [G])^\sharp\right)_{[G]\in\operatorname{GrIs}_n}$).

Now, forget that we fixed n. We thus have shown that, for every $n \in \mathbb{N}$, the family $\left([G]^\sharp\right)_{[G]\in\operatorname{GrIs}_n}$ is a basis of the $\mathbf{k}$-module $\mathcal{G}_n$. Therefore, the disjoint union of these families (over all $n \in \mathbb{N}$) is a basis of the

[1237] *Proof of (13.196.5):* Let G be a finite graph with n vertices. Write G as $G = (V, E)$.

Let $H = (V, E')$ be any graph satisfying $E' \supset E$ and $E' \neq E$. Clearly, the graph H has n vertices (since it has the same vertex set as G, and since G has n vertices). Thus, $[H] \in \operatorname{GrIs}_n$. Moreover, E is a proper subset of E' (since $E' \supset E$ and $E' \neq E$); hence, $|E| < |E'|$.

The definition of e yields $e[(V,E)] = |E|$ and $e[(V,E')] = |E'|$. Now, $e\left[\underbrace{G}_{=(V,E)}\right] = e[(V,E)] = |E| < |E'|$, so that $|E'| > e[G]$. But $e\left[\underbrace{H}_{=(V,E')}\right] = e[(V,E')] = |E'| > e[G]$. In other words, $[H] \prec [G]$ (by the definition of the relation $\prec$).

Now, forget that we fixed H. We thus have shown that if $H = (V, E')$ is any graph satisfying $E' \supset E$ and $E' \neq E$, then $[H] \in \operatorname{GrIs}_n$ and $[H] \prec [G]$. Hence,

$$\begin{aligned} &\sum_{\substack{H=(V,E');\\ E'\supset E \text{ and } E'\neq E}} (-1)^{|E'\setminus E|} [H] \\ &= (\text{a } \mathbf{k}\text{-linear combination of the elements } [H] \text{ for } [H] \in \operatorname{GrIs}_n \text{ satisfying } [H] \prec [G]) . \end{aligned} \tag{13.196.6}$$

But (13.196.4) becomes

$$\begin{aligned} &(\mathbf{c}_n [G])^\sharp \\ &= \sum_{\substack{H=(V,E');\\ E'\supset E}} (-1)^{|E'\setminus E|} [H] \\ &= \underbrace{(-1)^{|E\setminus E|}}_{\substack{=1\\ (\text{since } |E\setminus E| = |\varnothing| = 0)}} \left[\underbrace{(V,E)}_{=G}\right] + \underbrace{\sum_{\substack{H=(V,E');\\ E'\supset E \text{ and } E'\neq E}} (-1)^{|E'\setminus E|} [H]}_{\substack{=(\text{a } \mathbf{k}\text{-linear combination of the elements } [H] \text{ for } [H]\in\operatorname{GrIs}_n \text{ satisfying } [H]\prec[G])\\ (\text{by } (13.196.6))}} \\ &\qquad \left(\text{here, we have split off the addend for } E' = E \text{ and } H = (V,E) \text{ from the sum}\right) \\ &= [G] + (\text{a } \mathbf{k}\text{-linear combination of the elements } [H] \text{ for } [H] \in \operatorname{GrIs}_n \text{ satisfying } [H] \prec [G]) . \end{aligned}$$

This proves (13.196.5).

direct sum $\bigoplus_{n\in\mathbb{N}}\mathcal{G}_n$. Since the former disjoint union is the family $\left([G]^\sharp\right)_{[G]\in\mathrm{GrIs}}$, whereas the latter direct sum is $\bigoplus_{n\in\mathbb{N}}\mathcal{G}_n=\mathcal{G}$, this rewrites as follows: The family $\left([G]^\sharp\right)_{[G]\in\mathrm{GrIs}}$ is a basis of the $\mathbf{k}$-module $\mathcal{G}$. In other words, the elements $[G]^\sharp$, where $[G]$ ranges over all isomorphism classes of finite graphs, form a basis of the $\mathbf{k}$-module $\mathcal{G}$ (because GrIs is the set of all isomorphism classes of finite graphs). Proposition 7.3.9(b) is thus proven.

(c) We define a $\mathbf{k}$-linear map $\Delta':\mathcal{G}\to\mathcal{G}\otimes\mathcal{G}$ by

$$\Delta'[H]^\sharp=\sum_{\substack{(V_1,V_2);\\ V=V_1\sqcup V_2;\\ H=H|_{V_1}\sqcup H|_{V_2}}}[H\mid_{V_1}]^\sharp\otimes[H\mid_{V_2}]^\sharp .$$

[1238] In order to prove Proposition 7.3.9(c), it clearly suffices to show that $\Delta'=\Delta$.

Let $G=(V,E)$ be a finite graph. If T is a set of two-element subsets of V, and if R is a subset of V, then $T\mid_R$ shall denote the subset $\{X\in T\mid X\subset R\}$ of T. (Thus, if R is a subset of V, then $E\mid_R$ is the set of edges of the graph $G\mid_R$.)

[1238] This is well-defined, since the $[G]^\sharp$ form a basis of the $\mathbf{k}$-module $\mathcal{G}$.

Now, Proposition 7.3.9(a) yields $[G] = \sum_{\substack{H=(V,E');\\ E'\cap E=\varnothing}} [H]^\sharp$. Hence,

$$\begin{aligned}
\Delta'[G] &= \Delta' \sum_{\substack{H=(V,E');\\ E'\cap E=\varnothing}} [H]^\sharp = \sum_{\substack{H=(V,E');\\ E'\cap E=\varnothing}} \underbrace{\Delta'[H]^\sharp}_{= \sum_{\substack{(V_1,V_2);\\ V=V_1\sqcup V_2;\\ H=H|_{V_1}\sqcup H|_{V_2}}} [H|_{V_1}]^\sharp\otimes[H|_{V_2}]^\sharp} \\
&= \underbrace{\sum_{\substack{H=(V,E');\\ E'\cap E=\varnothing}} \sum_{\substack{(V_1,V_2);\\ V=V_1\sqcup V_2;\\ H=H|_{V_1}\sqcup H|_{V_2}}}}_{= \sum_{\substack{(V_1,V_2);\\ V=V_1\sqcup V_2}} \sum_{\substack{H=(V,E');\\ E'\cap E=\varnothing;\\ H=H|_{V_1}\sqcup H|_{V_2}}}} [H\mid_{V_1}]^\sharp \otimes [H\mid_{V_2}]^\sharp \\
&= \sum_{\substack{(V_1,V_2);\\ V=V_1\sqcup V_2}} \underbrace{\sum_{\substack{H=(V,E');\\ E'\cap E=\varnothing;\\ H=H|_{V_1}\sqcup H|_{V_2}}} [H\mid_{V_1}]^\sharp \otimes [H\mid_{V_2}]^\sharp}_{\substack{= \sum_{\substack{H_1=(V_1,E_1');\\ E_1'\cap(E|_{V_1})=\varnothing}} \sum_{\substack{H_2=(V_2,E_2');\\ E_2'\cap(E|_{V_2})=\varnothing}} \sum_{\substack{H=(V,E');\\ E'\cap E=\varnothing;\\ H=H|_{V_1}\sqcup H|_{V_2};\\ H|_{V_1}=H_1;\ H|_{V_2}=H_2}} [H|_{V_1}]^\sharp\otimes[H|_{V_2}]^\sharp \\ \text{(because for every } H=(V,E') \text{ satisfying } E'\cap E=\varnothing, \\ \text{we can write the subgraph } H|_{V_1} \text{ in the form } H_1=(V_1,E_1') \\ \text{with some } E_1' \text{ satisfying } E_1'\cap(E|_{V_1})=\varnothing, \text{ and we can} \\ \text{write the subgraph } H|_{V_2} \text{ in the form } H_2=(V_2,E_2') \\ \text{with some } E_2' \text{ satisfying } E_2'\cap(E|_{V_2})=\varnothing)}} \\
&= \underbrace{\sum_{\substack{(V_1,V_2);\\ V=V_1\sqcup V_2}}}_{= \sum_{\substack{(V_1,V_2);\\ V_1\sqcup V_2=V}}} \sum_{\substack{H_1=(V_1,E_1');\\ E_1'\cap(E|_{V_1})=\varnothing}} \sum_{\substack{H_2=(V_2,E_2');\\ E_2'\cap(E|_{V_2})=\varnothing}} \sum_{\substack{H=(V,E');\\ E'\cap E=\varnothing;\\ H=H|_{V_1}\sqcup H|_{V_2};\\ H|_{V_1}=H_1;\ H|_{V_2}=H_2}} \Bigg[\underbrace{H\mid_{V_1}}_{=H_1}\Bigg]^\sharp \otimes \Bigg[\underbrace{H\mid_{V_2}}_{=H_2}\Bigg]^\sharp \\
&= \sum_{\substack{(V_1,V_2);\\ V_1\sqcup V_2=V}} \sum_{\substack{H_1=(V_1,E_1');\\ E_1'\cap(E|_{V_1})=\varnothing}} \sum_{\substack{H_2=(V_2,E_2');\\ E_2'\cap(E|_{V_2})=\varnothing}} \underbrace{\sum_{\substack{H=(V,E');\\ E'\cap E=\varnothing;\\ H=H|_{V_1}\sqcup H|_{V_2};\\ H|_{V_1}=H_1;\ H|_{V_2}=H_2}} [H_1]^\sharp \otimes [H_2]^\sharp}_{\substack{=[H_1]^\sharp\otimes[H_2]^\sharp \\ \text{(because this sum has only one addend} \\ \text{(indeed, there exists only one } H=(V,E') \text{ satisfying} \\ E'\cap E=\varnothing,\ H=H|_{V_1}\sqcup H|_{V_2},\ H|_{V_1}=H_1 \text{ and } H|_{V_2}=H_2))}} \\
&= \sum_{\substack{(V_1,V_2);\\ V_1\sqcup V_2=V}} \sum_{\substack{H_1=(V_1,E_1');\\ E_1'\cap(E|_{V_1})=\varnothing}} \sum_{\substack{H_2=(V_2,E_2');\\ E_2'\cap(E|_{V_2})=\varnothing}} [H_1]^\sharp \otimes [H_2]^\sharp .
\end{aligned}$$

Comparing this with

$$\Delta [G] = \sum_{\substack{(V_1,V_2);\\ V_1 \sqcup V_2 = V}} \underbrace{[G \mid_{V_1}]}_{\substack{= \sum_{\substack{H_1 = (V_1, E_1');\\ E_1' \cap (E|_{V_1}) = \varnothing}} [H_1]^\sharp \\ \text{(by Proposition 7.3.9(a))}}} \otimes \underbrace{[G \mid_{V_2}]}_{\substack{= \sum_{\substack{H_2 = (V_2, E_2');\\ E_2' \cap (E|_{V_2}) = \varnothing}} [H_2]^\sharp \\ \text{(by Proposition 7.3.9(a))}}}$$

$$= \sum_{\substack{(V_1,V_2);\\ V_1 \sqcup V_2 = V}} \sum_{\substack{H_1 = (V_1, E_1');\\ E_1' \cap (E|_{V_1}) = \varnothing}} \sum_{\substack{H_2 = (V_2, E_2');\\ E_2' \cap (E|_{V_2}) = \varnothing}} [H_1]^\sharp \otimes [H_2]^\sharp ,$$

we obtain $\Delta' [G] = \Delta [G]$. Since this holds for every graph G, we thus obtain $\Delta' = \Delta$. As we know, this proves Proposition 7.3.9(c).

(d) We define a $\mathbf{k}$-bilinear operation $\sharp : \mathcal{G} \times \mathcal{G} \to \mathcal{G}$ (written infix[1239]) by

$$[H_1]^\sharp \sharp [H_2]^\sharp = \sum_{\substack{H = (V_1 \sqcup V_2, E);\\ H|_{V_1} = H_1;\\ H|_{V_2} = H_2}} [H]^\sharp .$$

[1240] In order to prove Proposition 7.3.9(d), it clearly suffices to show that this operation $\sharp$ is identical with the usual multiplication on $\mathcal{G}$.

Let G_1 and G_2 be any two finite graphs. We shall show that $[G_1] \sharp [G_2] = [G_1] [G_2]$.

[1239] This means that we write $a \sharp b$ to denote the image of a pair $(a, b) \in \mathcal{G} \times \mathcal{G}$ under this operation $\sharp$.

[1240] This is well-defined, since the $[G]^\sharp$ form a basis of the $\mathbf{k}$-module $\mathcal{G}$.

Indeed, write G_1 and G_2 in the forms $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$. Then,

$$\underbrace{[G_1]}_{\substack{=\sum_{\substack{H_1=\left(V_1,E_1'\right);\\ E_1'\cap E_1=\varnothing}} [H_1]^\sharp \\ \text{(by Proposition 7.3.9(a))}}} \sharp \underbrace{[G_2]}_{\substack{=\sum_{\substack{H_2=\left(V_2,E_2'\right);\\ E_2'\cap E_2=\varnothing}} [H_2]^\sharp \\ \text{(by Proposition 7.3.9(a))}}}$$

$$= \left(\sum_{\substack{H_1=\left(V_1,E_1'\right);\\ E_1'\cap E_1=\varnothing}} [H_1]^\sharp \right) \sharp \left(\sum_{\substack{H_2=\left(V_2,E_2'\right);\\ E_2'\cap E_2=\varnothing}} [H_2]^\sharp \right) = \sum_{\substack{H_1=\left(V_1,E_1'\right);\\ E_1'\cap E_1=\varnothing}} \sum_{\substack{H_2=\left(V_2,E_2'\right);\\ E_2'\cap E_2=\varnothing}} \underbrace{[H_1]^\sharp \sharp [H_2]^\sharp}_{\substack{=\sum_{\substack{H=(V_1\sqcup V_2,E);\\ H\mid_{V_1}=H_1;\\ H\mid_{V_2}=H_2}} [H]^\sharp}}$$

$$= \underbrace{\sum_{\substack{H_1=\left(V_1,E_1'\right);\\ E_1'\cap E_1=\varnothing}} \sum_{\substack{H_2=\left(V_2,E_2'\right);\\ E_2'\cap E_2=\varnothing}} \sum_{\substack{H=(V_1\sqcup V_2,E);\\ H\mid_{V_1}=H_1;\\ H\mid_{V_2}=H_2}}}_{=\sum_{H_1=\left(V_1,E_1'\right)} \sum_{H_2=\left(V_2,E_2'\right)} \sum_{\substack{H=(V_1\sqcup V_2,E);\\ H\mid_{V_1}=H_1;\\ H\mid_{V_2}=H_2;\\ E_1'\cap E_1=\varnothing;\\ E_2'\cap E_2=\varnothing}}} [H]^\sharp$$

$$= \sum_{H_1=\left(V_1,E_1'\right)} \sum_{H_2=\left(V_2,E_2'\right)} \underbrace{\sum_{\substack{H=(V_1\sqcup V_2,E);\\ H\mid_{V_1}=H_1;\\ H\mid_{V_2}=H_2;\\ E_1'\cap E_1=\varnothing;\\ E_2'\cap E_2=\varnothing}}}_{\substack{=\sum_{\substack{H=(V_1\sqcup V_2,E);\\ H\mid_{V_1}=H_1;\\ H\mid_{V_2}=H_2;\\ \left(E\mid_{V_1}\right)\cap E_1=\varnothing;\\ \left(E\mid_{V_2}\right)\cap E_2=\varnothing}} \\ \text{(because for a graph } H=(V_1\sqcup V_2,E) \\ \text{satisfying } H\mid_{V_1}=H_1 \text{ and } H\mid_{V_2}=H_2, \\ \text{we have } E_1'=E\mid_{V_1} \text{ and } E_2'=E\mid_{V_2}\text{)}}} [H]^\sharp$$

$$= \underbrace{\sum_{H_1=\left(V_1,E_1'\right)} \sum_{H_2=\left(V_2,E_2'\right)} \sum_{\substack{H=(V_1\sqcup V_2,E);\\ H\mid_{V_1}=H_1;\\ H\mid_{V_2}=H_2;\\ \left(E\mid_{V_1}\right)\cap E_1=\varnothing;\\ \left(E\mid_{V_2}\right)\cap E_2=\varnothing}}}_{\substack{=\sum_{\substack{H=(V_1\sqcup V_2,E);\\ \left(E\mid_{V_1}\right)\cap E_1=\varnothing;\\ \left(E\mid_{V_2}\right)\cap E_2=\varnothing}} = \sum_{\substack{H=(V_1\sqcup V_2,E);\\ E\cap(E_1\sqcup E_2)=\varnothing}} \\ \text{(since the statement } \left(\left(E\mid_{V_1}\right)\cap E_1=\varnothing \text{ and } \left(E\mid_{V_2}\right)\cap E_2=\varnothing\right) \\ \text{is equivalent to } E\cap(E_1\sqcup E_2)=\varnothing\text{)}}} [H]^\sharp$$

$$= \sum_{\substack{H=(V_1\sqcup V_2,E);\\ E\cap(E_1\sqcup E_2)=\varnothing}} [H]^\sharp .$$

Compared with

$$\begin{aligned}[G_1][G_2] &= [G_1 \sqcup G_2] = [(V_1 \sqcup V_2, E_1 \sqcup E_2)] \\ &= \sum_{\substack{H=(V_1 \sqcup V_2, E);\\ E \cap (E_1 \sqcup E_2) = \varnothing}} [H]^\sharp \qquad \text{(by Proposition 7.3.9(a))},\end{aligned}$$

this yields $[G_1] \sharp [G_2] = [G_1][G_2]$.

We now forget that we fixed G_1 and G_2. We have thus shown that $[G_1] \sharp [G_2] = [G_1][G_2]$ for any two finite graphs G_1 and G_2. Thus, the operation $\sharp$ is identical with the usual multiplication on $\mathcal{G}$. This completes the proof of Proposition 7.3.9(d). □

Proof of Proposition 7.3.11. Proposition 7.3.9(a) shows that every finite graph $G = (V, E)$ satisfies

$$[G] = \sum_{\substack{H=(V,E');\\ E' \cap E = \varnothing}} [H]^\sharp = \sum_{\substack{K=(V,E');\\ E' \cap E = \varnothing}} [K]^\sharp \tag{13.196.7}$$

(here, we renamed the summation index H as K). Let us now show that

$$([G], [H]) = ([H], [G]) \tag{13.196.8}$$

for any two finite graphs G and H.

Proof of (13.196.8): We shall use the following notation: If V and W are two sets, and if $\varphi : V \to W$ is a map, then φ_* will denote the map from the powerset of V to the powerset of W which sends every $T \subset V$ to $\varphi(T) \subset W$. This map φ_* is a bijection if φ is a bijection.

Let G and H be two finite graphs. Let us write G and H in the forms $G = (V, E)$ and $H = (W, F)$. Now,

$$\begin{aligned}\left(\underbrace{[G]}_{\substack{= \sum\limits_{\substack{K=(V,E');\\ E' \cap E = \varnothing}} [K]^\sharp \\ \text{(by (13.196.7))}}}, [H] \right) &= \left(\sum_{\substack{K=(V,E');\\ E' \cap E = \varnothing}} [K]^\sharp, [H] \right) = \sum_{\substack{K=(V,E');\\ E' \cap E = \varnothing}} \underbrace{\left([K]^\sharp, [H]\right)}_{\substack{=|\operatorname{Iso}(K,H)| \\ \text{(by the definition of} \\ \text{the form } (\cdot,\cdot))}} \\ &\qquad \text{(since the form } (\cdot,\cdot) \text{ is } \mathbf{k}\text{-bilinear)} \\ &= \sum_{\substack{K=(V,E');\\ E' \cap E = \varnothing}} |\operatorname{Iso}(K, H)|. \end{aligned} \tag{13.196.9}$$

However, for every graph $K = (V, E')$, we have

$$\begin{aligned}\operatorname{Iso}(K, H) &= (\text{the set of all isomorphisms from } K \text{ to } H) \\ &= \{\varphi : V \to W \mid \varphi \text{ is a bijection and satisfies } \varphi_*(E') = F\}\end{aligned}$$

(because the isomorphisms from K to H are defined to be the bijections $\varphi : V \to W$ such that $\varphi_* (E') = F$). Hence,

$$
\begin{aligned}
&\sum_{\substack{K=(V,E');\\ E'\cap E=\varnothing}} \left| \underbrace{\operatorname{Iso}(K,H)}_{=\{\varphi : V\to W \ \mid\ \varphi \text{ is a bijection and satisfies } \varphi_*(E')=F\}} \right| \\
&= \underbrace{\sum_{\substack{K=(V,E');\\ E'\cap E=\varnothing}}}_{\substack{=\sum\limits_{\substack{E' \text{ is a set of}\\ \text{two-element subsets of } V;\\ E'\cap E=\varnothing}}}} \left| \left\{ \varphi : V \to W \ \mid\ \varphi \text{ is a bijection and satisfies } \underbrace{\varphi_*(E') = F}_{\substack{\text{this is equivalent to}\\ E'=(\varphi_*)^{-1}(F)\\ \text{(since } \varphi_* \text{ is a bijection}\\ \text{(since } \varphi \text{ is a bijection))}}} \right\} \right| \\
&= \sum_{\substack{E' \text{ is a set of}\\ \text{two-element subsets of } V;\\ E'\cap E=\varnothing}} \left| \left\{ \varphi : V \to W \ \mid\ \varphi \text{ is a bijection and satisfies } E' = (\varphi_*)^{-1}(F) \right\} \right| \\
&= \left| \left\{ \varphi : V \to W \ \mid\ \varphi \text{ is a bijection and satisfies } (\varphi_*)^{-1}(F) \cap E = \varnothing \right\} \right| .
\end{aligned}
$$

Now, (13.196.9) becomes

$$
\begin{aligned}
([G],[H]) &= \sum_{\substack{K=(V,E');\\ E'\cap E=\varnothing}} \left|\operatorname{Iso}(K,H)\right| \\
&= \left| \left\{ \varphi : V \to W \ \mid\ \varphi \text{ is a bijection and satisfies } (\varphi_*)^{-1}(F) \cap E = \varnothing \right\} \right| . \qquad (13.196.10)
\end{aligned}
$$

The same argument (applied to H, W, F, G, V and E instead of G, V, E, H, W and F) yields

$$
([H],[G]) = \left| \left\{ \varphi : W \to V \ \mid\ \varphi \text{ is a bijection and satisfies } (\varphi_*)^{-1}(E) \cap F = \varnothing \right\} \right| . \tag{13.196.11}
$$

Now, we define two sets $\mathfrak{P}$ and $\mathfrak{Q}$ by

$$
\mathfrak{P} = \left\{ \varphi : V \to W \ \mid\ \varphi \text{ is a bijection and satisfies } (\varphi_*)^{-1}(F) \cap E = \varnothing \right\}
$$

and

$$
\mathfrak{Q} = \left\{ \varphi : W \to V \ \mid\ \varphi \text{ is a bijection and satisfies } (\varphi_*)^{-1}(E) \cap F = \varnothing \right\} .
$$

Then, (13.196.10) becomes

$$
([G],[H]) = \left| \underbrace{\left\{ \varphi : V \to W \ \mid\ \varphi \text{ is a bijection and satisfies } (\varphi_*)^{-1}(F) \cap E = \varnothing \right\}}_{=\mathfrak{P}} \right| = |\mathfrak{P}| . \tag{13.196.12}
$$

Also, (13.196.11) becomes

$$
([H],[G]) = \left| \underbrace{\left\{ \varphi : W \to V \ \mid\ \varphi \text{ is a bijection and satisfies } (\varphi_*)^{-1}(E) \cap F = \varnothing \right\}}_{=\mathfrak{Q}} \right| = |\mathfrak{Q}| . \tag{13.196.13}
$$

But every $\psi \in \mathfrak{P}$ satisfies $\psi^{-1} \in \mathfrak{Q}$ [1241]. Hence, we can define a map

$$\begin{aligned} \mathbf{A} : \mathfrak{P} &\to \mathfrak{Q}, \\ \psi &\mapsto \psi^{-1}. \end{aligned}$$

Similarly, we can define a map

$$\begin{aligned} \mathbf{B} : \mathfrak{Q} &\to \mathfrak{P}, \\ \psi &\mapsto \psi^{-1}. \end{aligned}$$

Consider these two maps $\mathbf{A}$ and $\mathbf{B}$. Clearly, these maps $\mathbf{A}$ and $\mathbf{B}$ are mutually inverse[1242]. Hence, $\mathbf{A}$ is a bijection. Thus, there exists a bijection $\mathfrak{P} \to \mathfrak{Q}$ (namely, $\mathbf{A}$). Consequently, $|\mathfrak{P}| = |\mathfrak{Q}|$. Now, (13.196.12) becomes $([G], [H]) = |\mathfrak{P}| = |\mathfrak{Q}| = ([H], [G])$ (by (13.196.13)). This proves (13.196.8).

Now, let a and b be two elements of $\mathcal{G}$. We want to show that $(a, b) = (b, a)$. This equality is $\mathbf{k}$-linear in each of a and b. Therefore, we can WLOG assume that a and b belong to the family $([G])_{[G] \text{ is an isomorphism class of finite graphs}}$ (because this family is a basis of the $\mathbf{k}$-module $\mathcal{G}$). Assume this. Then, $a = [G]$ and $b = [H]$ for two finite graphs G and H. Consider these G and H. Now,

$$\begin{aligned} \Big(\underbrace{a}_{=[G]}, \underbrace{b}_{=[H]}\Big) = ([G], [H]) &= \Big(\underbrace{[H]}_{=b}, \underbrace{[G]}_{=a}\Big) \qquad \text{(by (13.196.8))} \\ &= (b, a). \end{aligned}$$

We thus have shown $(a, b) = (b, a)$.

Let us now forget that we fixed a and b. We thus have proven that $(a, b) = (b, a)$ for any $a \in \mathcal{G}$ and $b \in \mathcal{G}$. In other words, the form $(\cdot, \cdot) : \mathcal{G} \times \mathcal{G} \to \mathbf{k}$ is symmetric. This proves Proposition 7.3.11. □

[1241] *Proof.* Let $\psi \in \mathfrak{P}$. Then, $\psi \in \mathfrak{P} = \left\{ \varphi : V \to W \ \mid \ \varphi \text{ is a bijection and satisfies } (\varphi_*)^{-1}(F) \cap E = \varnothing \right\}$. Hence, ψ is a bijection $V \to W$ and satisfies $(\psi_*)^{-1}(F) \cap E = \varnothing$.

Let ρ be the map $\psi^{-1} : W \to V$. This map ρ is well-defined (since ψ is a bijection) and is a bijection itself (since $\rho = \psi^{-1}$).

But ψ is a bijection. Thus, ψ_* is a bijection as well, and satisfies $(\psi_*)^{-1} = \Big(\underbrace{\psi^{-1}}_{=\rho}\Big)_* = \rho_*$. Of course, ρ_* is also a bijection (since ρ is a bijection).

Now,

$$\begin{aligned} \rho_*\left((\rho_*)^{-1}(E) \cap F\right) &= \underbrace{\rho_*\left((\rho_*)^{-1}(E)\right)}_{\substack{=E \\ \text{(since } \rho_* \text{ is a bijection)}}} \cap \underbrace{\rho_*}_{=(\psi_*)^{-1}}(F) \qquad \text{(since } \rho_* \text{ is a bijection)} \\ &= E \cap (\psi_*)^{-1}(F) = (\psi_*)^{-1}(F) \cap E = \varnothing. \end{aligned}$$

Thus,

$$\begin{aligned} (\rho_*)^{-1}(E) \cap F &= \rho_*(\varnothing) \qquad \text{(since } \rho_* \text{ is a bijection)} \\ &= \varnothing. \end{aligned}$$

Hence, $\rho : W \to V$ is a bijection and satisfies $(\rho_*)^{-1}(E) \cap F = 0$. In other words,

$$\rho \in \left\{ \varphi : W \to V \ \mid \ \varphi \text{ is a bijection and satisfies } (\varphi_*)^{-1}(E) \cap F = \varnothing \right\} = \mathfrak{Q}.$$

Thus, $\psi^{-1} = \rho \in \mathfrak{Q}$, qed.

[1242] *Proof.* Every $\psi \in \mathfrak{P}$ satisfies

$$\begin{aligned} (\mathbf{B} \circ \mathbf{A})(\psi) = \mathbf{B}\left(\underbrace{\mathbf{A}(\psi)}_{\substack{=\psi^{-1} \\ \text{(by the definition} \\ \text{of } \mathbf{A})}}\right) &= \mathbf{B}\left(\psi^{-1}\right) = \left(\psi^{-1}\right)^{-1} \qquad \text{(by the definition of } \mathbf{B}) \\ &= \psi = \operatorname{id}(\psi). \end{aligned}$$

Hence, $\mathbf{B} \circ \mathbf{A} = \operatorname{id}$. Similarly, $\mathbf{A} \circ \mathbf{B} = \operatorname{id}$. Combining this with $\mathbf{B} \circ \mathbf{A} = \operatorname{id}$, we conclude that the maps $\mathbf{A}$ and $\mathbf{B}$ are mutually inverse. Qed.

Proof of Proposition 7.3.13. In this proof, we shall use all the notations that appeared in Definition 7.3.12. We shall also use the fact that the family $\left([G]^\sharp\right)_{[G] \text{ is an isomorphism class of finite graphs}}$ is a basis of the $\mathbf{k}$-module $\mathcal{G}$. (This fact is Proposition 7.3.9(b).)

(a) We first notice that any two finite graphs G and H satisfy

$$\left(\psi\left([G]^\sharp\right)\right)[H] = |\operatorname{Iso}(G,H)| \qquad (13.196.14)$$

[1243]. Now, let $a \in \mathcal{G}$ and $b \in \mathcal{G}$. We want to prove the equality $(\psi(a))(b) = (a,b)$. This equality is $\mathbf{k}$-linear in each of a and b. Thus, we can WLOG assume that a belongs to the family $\left([G]^\sharp\right)_{[G] \text{ is an isomorphism class of finite graphs}}$ and that b belongs to the family $([G])_{[G] \text{ is an isomorphism class of finite graphs}}$ (because both of these families are bases of the $\mathbf{k}$-module $\mathcal{G}$). Assume this. Then, $a = [G]^\sharp$ and $b = [H]$ for two finite graphs G and H. Consider these G and H. Now,

$$\begin{aligned}
\left(\psi\left(\underbrace{a}_{=[G]^\sharp}\right)\right)\left(\underbrace{b}_{=[H]}\right) &= \left(\psi\left([G]^\sharp\right)\right)[H] = |\operatorname{Iso}(G,H)| && (\text{by } (13.196.14)) \\
&= \left(\underbrace{[G]^\sharp}_{=a}, \underbrace{[H]}_{=b}\right) && \left(\text{since } \left([G]^\sharp, [H]\right) = |\operatorname{Iso}(G,H)|\right) \\
&= (a,b).
\end{aligned}$$

This proves Proposition 7.3.13(a).

[1243] *Proof of (13.196.14):* Let G and H be two finite graphs. Then,

$$\left(\underbrace{\psi\left([G]^\sharp\right)}_{=\operatorname{aut}(G)\cdot[G]^*}\right)[H] = \left(\operatorname{aut}(G)\cdot[G]^*\right)[H] = \operatorname{aut}(G)\cdot \underbrace{[G]^*[H]}_{=\delta_{[G],[H]}} = \operatorname{aut}(G)\cdot\delta_{[G],[H]}.$$

We are in one of the following two cases:

Case 1: The graphs G and H are isomorphic.

Case 2: The graphs G and H are not isomorphic.

Let us first consider Case 1. In this case, the graphs G and H are isomorphic. In other words, there exists a graph isomorphism α from G to H. Consider this α. Then,

$$\begin{aligned}
\operatorname{Iso}(G,G) &\to \operatorname{Iso}(G,H), \\
\psi &\mapsto \alpha\circ\psi
\end{aligned}$$

is a bijection (this is very easy to check). Thus, there exists a bijection $\operatorname{Iso}(G,G)\to\operatorname{Iso}(G,H)$. Hence, $|\operatorname{Iso}(G,G)| = |\operatorname{Iso}(G,H)|$. But the definition of $\operatorname{aut}(G)$ yields $\operatorname{aut}(G) = |\operatorname{Iso}(G,G)|$. Now,

$$\left(\psi\left([G]^\sharp\right)\right)[H] = \operatorname{aut}(G)\cdot \underbrace{\delta_{[G],[H]}}_{\substack{=1 \\ (\text{since } [G]=[H] \\ (\text{because the graphs } G \\ \text{and } H \text{ are isomorphic}))}} = \operatorname{aut}(G) = |\operatorname{Iso}(G,G)| = |\operatorname{Iso}(G,H)|.$$

Hence, (13.196.14) is proven in Case 1.

Let us now consider Case 2. In this case, the graphs G and H are not isomorphic. Thus, $[G] \neq [H]$, so that $\delta_{[G],[H]} = 0$.

But the graphs G and H are not isomorphic. Hence, the set of all isomorphisms from G to H is empty. In other words, $\operatorname{Iso}(G,H) = \varnothing$ (since $\operatorname{Iso}(G,H)$ is the set of all isomorphisms from G to H). Hence, $|\operatorname{Iso}(G,H)| = |\varnothing| = 0$. Now,

$$\left(\psi\left([G]^\sharp\right)\right)[H] = \operatorname{aut}(G)\cdot\underbrace{\delta_{[G],[H]}}_{=0} = 0 = |\operatorname{Iso}(G,H)|.$$

Hence, (13.196.14) is proven in Case 2.

Now, (13.196.14) is proven in both Cases 1 and 2. Hence, (13.196.14) always holds.

(b) We have $\psi(1_{\mathcal{G}}) = 1_{\mathcal{G}^o}$ [1244]. Also, every $a \in \mathcal{G}$ and $b \in \mathcal{G}$ satisfy $\psi(ab) = \psi(a) \cdot \psi(b)$ [1245]. Hence, $\psi : \mathcal{G} \to \mathcal{G}^o$ is a **k**-algebra morphism (since $\psi(1_{\mathcal{G}}) = 1_{\mathcal{G}^o}$).

Next, we notice that the **k**-linear map ψ is graded (because for any finite graph G, both $[G]^\sharp$ and $[G]^*$ are homogeneous elements having degree $|G|$). Thus, the **k**-linear map $\psi : \mathcal{G} \to \mathcal{G}^o$ gives rise to an adjoint **k**-linear map $\psi^* : (\mathcal{G}^o)^o \to \mathcal{G}^o$.

Now, the canonical **k**-module homomorphism $\mathcal{G} \to (\mathcal{G}^o)^o$ (which sends every $a \in \mathcal{G}$ to the map $\mathcal{G}^o \to \mathbf{k},\ f \mapsto f(a)$) is a **k**-module isomorphism (since $\mathcal{G}$ is of finite type). We thus identify $\mathcal{G}$ with $(\mathcal{G}^o)^o$ along this

[1244] *Proof.* Let $\mathbf{0}$ denote the empty graph.

We have $1_{\mathcal{G}^o} = \epsilon_{\mathcal{G}}$ (by the definition of the **k**-algebra $\mathcal{G}^o$). Thus, for every finite graph G, we have

$$\underbrace{1_{\mathcal{G}^o}}_{=\epsilon_{\mathcal{G}}}([G]) = \epsilon_{\mathcal{G}}([G]) = \delta_{[G],[\mathbf{0}]} \qquad \text{(by the definition of } \epsilon_{\mathcal{G}})$$
$$= \delta_{[\mathbf{0}],[G]} = [\mathbf{0}]^*[G] \qquad \left(\text{since } [\mathbf{0}]^*[G] = \delta_{[\mathbf{0}],[G]} \text{ (by the definition of } [\mathbf{0}]^*)\right).$$

Hence, the two maps $1_{\mathcal{G}^o}$ and $[\mathbf{0}]^*$ are equal to each other on every element of the basis $([G])_{[G] \text{ is an isomorphism class of finite graphs}}$ of $\mathcal{G}$. Since these two maps are **k**-linear, this yields that these two maps $1_{\mathcal{G}^o}$ and $[\mathbf{0}]^*$ are identical. In other words, $1_{\mathcal{G}^o} = [\mathbf{0}]^*$.

On the other hand, the definition of $[\mathbf{0}]^\sharp$ readily yields $[\mathbf{0}]^\sharp = [\mathbf{0}] = 1_{\mathcal{G}}$. But the definition of ψ yields $\psi\left([\mathbf{0}]^\sharp\right) = \underbrace{\operatorname{aut}(\mathbf{0})}_{=1} \cdot [\mathbf{0}]^* = [\mathbf{0}]^* = 1_{\mathcal{G}^o}$ (since $1_{\mathcal{G}^o} = [\mathbf{0}]^*$), so that $1_{\mathcal{G}^o} = \psi\left(\underbrace{[\mathbf{0}]^\sharp}_{=1_{\mathcal{G}}}\right) = \psi(1_{\mathcal{G}})$, qed.

[1245] *Proof.* Let $a \in \mathcal{G}$ and $b \in \mathcal{G}$. We need to prove the equality $\psi(ab) = \psi(a) \cdot \psi(b)$. Since this equality is **k**-linear in each of a and b, we can WLOG assume that a and b are elements of the family $\left([G]^\sharp\right)_{[G] \text{ is an isomorphism class of finite graphs}}$ (because this family is a basis of the **k**-module $\mathcal{G}$). Assume this.

There exist two finite graphs H_1 and H_2 such that $a = [H_1]^\sharp$ and $b = [H_2]^\sharp$ (since a and b are elements of the family $\left([G]^\sharp\right)_{[G] \text{ is an isomorphism class of finite graphs}}$). Consider these H_1 and H_2. Write the graphs H_1 and H_2 in the forms $H_1 = (V_1, E_1)$ and $H_2 = (V_2, E_2)$.

Multiplying the equalities $a = [H_1]^\sharp$ and $b = [H_2]^\sharp$, we obtain

$$ab = [H_1]^\sharp [H_2]^\sharp = \sum_{\substack{H=(V_1 \sqcup V_2, E);\\ H|_{V_1}=H_1;\\ H|_{V_2}=H_2}} [H]^\sharp \qquad \text{(by Proposition 7.3.9(d))}.$$

Now, fix a finite graph G. We shall show that $(\psi(ab))[G] = (\psi(a) \cdot \psi(b))[G]$.

Write the graph G in the form $G = (V, F)$. Then,

$$\left(\psi \underbrace{(ab)}_{= \sum_{\substack{H=(V_1 \sqcup V_2, E);\\ H|_{V_1}=H_1;\\ H|_{V_2}=H_2}} [H]^\sharp}\right)[G] = \left(\psi\left(\sum_{\substack{H=(V_1 \sqcup V_2, E);\\ H|_{V_1}=H_1;\\ H|_{V_2}=H_2}} [H]^\sharp\right)\right)[G] = \sum_{\substack{H=(V_1 \sqcup V_2, E);\\ H|_{V_1}=H_1;\\ H|_{V_2}=H_2}} \underbrace{\psi\left([H]^\sharp\right)[G]}_{\substack{=|\operatorname{Iso}(H,G)|\\ \text{(by (13.196.14), applied to}\\ H \text{ and } G \text{ instead of } G \text{ and } H)}}$$
$$= \sum_{\substack{H=(V_1 \sqcup V_2, E);\\ H|_{V_1}=H_1;\\ H|_{V_2}=H_2}} |\operatorname{Iso}(H,G)| = \left| \bigsqcup_{\substack{H=(V_1 \sqcup V_2, E);\\ H|_{V_1}=H_1;\\ H|_{V_2}=H_2}} \operatorname{Iso}(H,G) \right|.$$

On the other hand, Exercise 1.6.1(b) shows that the **k**-algebra structure defined on $\mathcal{G}^*$ is precisely the one defined on $\operatorname{Hom}(\mathcal{G}, \mathbf{k}) = \mathcal{G}^*$ according to Definition 1.4.1. Thus, the product of two elements of $\mathcal{G}^*$ is the convolution of these elements (viewed as maps from $\mathcal{G}$ to **k**). Applying this to the two elements $\psi(a)$ and $\psi(b)$, we obtain $\psi(a) \cdot \psi(b) = \psi(a) \star \psi(b) =$

$m_{\mathbf{k}} \circ (\psi(a) \otimes \psi(b)) \circ \Delta_{\mathcal{G}}$, so that

$$\begin{aligned}
(\psi(a) \cdot \psi(b))[G] &= (m_{\mathbf{k}} \circ (\psi(a) \otimes \psi(b)) \circ \Delta_{\mathcal{G}})([G]) \\
&= \left(m_{\mathbf{k}} \circ \left(\psi\left(\underbrace{a}_{=[H_1]^{\sharp}}\right) \otimes \psi\left(\underbrace{b}_{=[H_2]^{\sharp}}\right)\right)\right) \left(\underbrace{\Delta_{\mathcal{G}}([G])}_{\substack{=\sum_{\substack{(V_1,V_2);\\ V_1 \sqcup V_2 = V}} [G|_{V_1}] \otimes [G|_{V_2}] \\ \text{(by the definition of } \Delta_{\mathcal{G}})}}\right) \\
&= \left(m_{\mathbf{k}} \circ \left(\psi\left([H_1]^{\sharp}\right) \otimes \psi\left([H_2]^{\sharp}\right)\right)\right) \left(\sum_{\substack{(V_1,V_2);\\ V_1 \sqcup V_2 = V}} [G\mid_{V_1}] \otimes [G\mid_{V_2}]\right) \\
&= m_{\mathbf{k}} \left(\sum_{\substack{(V_1,V_2);\\ V_1 \sqcup V_2 = V}} \underbrace{\left(\psi\left([H_1]^{\sharp}\right) \otimes \psi\left([H_2]^{\sharp}\right)\right)([G\mid_{V_1}] \otimes [G\mid_{V_2}])}_{=(\psi([H_1]^{\sharp}))[G|_{V_1}] \otimes (\psi([H_2]^{\sharp}))[G|_{V_2}]}\right) \\
&= m_{\mathbf{k}} \left(\sum_{\substack{(V_1,V_2);\\ V_1 \sqcup V_2 = V}} \left(\psi\left([H_1]^{\sharp}\right)\right)[G\mid_{V_1}] \otimes \left(\psi\left([H_2]^{\sharp}\right)\right)[G\mid_{V_2}]\right) \\
&= \sum_{\substack{(V_1,V_2);\\ V_1 \sqcup V_2 = V}} \underbrace{\left(\psi\left([H_1]^{\sharp}\right)\right)[G\mid_{V_1}]}_{\substack{=\left|\operatorname{Iso}\left(H_1, G|_{V_1}\right)\right| \\ \text{(by (13.196.14), applied to} \\ H_1 \text{ and } G|_{V_1} \text{ instead of } G \text{ and } H)}} \cdot \underbrace{\left(\psi\left([H_2]^{\sharp}\right)\right)[G\mid_{V_2}]}_{\substack{=\left|\operatorname{Iso}\left(H_2, G|_{V_2}\right)\right| \\ \text{(by (13.196.14), applied to} \\ H_2 \text{ and } G|_{V_2} \text{ instead of } G \text{ and } H)}} \\
&= \sum_{\substack{(V_1,V_2);\\ V_1 \sqcup V_2 = V}} \left|\operatorname{Iso}\left(H_1, G\mid_{V_1}\right)\right| \cdot \left|\operatorname{Iso}\left(H_2, G\mid_{V_2}\right)\right| = \sum_{\substack{(W_1,W_2);\\ W_1 \sqcup W_2 = V}} \left|\operatorname{Iso}\left(H_1, G\mid_{W_1}\right)\right| \cdot \left|\operatorname{Iso}\left(H_2, G\mid_{W_2}\right)\right| \\
&= \left|\bigsqcup_{\substack{(W_1,W_2);\\ W_1 \sqcup W_2 = V}} \operatorname{Iso}\left(H_1, G\mid_{W_1}\right) \times \operatorname{Iso}\left(H_2, G\mid_{W_2}\right)\right|.
\end{aligned}$$

Now, we are going to construct a bijection between the sets

$$\bigsqcup_{\substack{H=(V_1 \sqcup V_2, E);\\ H|_{V_1} = H_1;\\ H|_{V_2} = H_2}} \operatorname{Iso}(H, G) \qquad \text{and} \qquad \bigsqcup_{\substack{(W_1,W_2);\\ W_1 \sqcup W_2 = V}} \operatorname{Iso}\left(H_1, G\mid_{W_1}\right) \times \operatorname{Iso}\left(H_2, G\mid_{W_2}\right).$$

First, let us agree to encode the elements of the set

$$\bigsqcup_{\substack{H=(V_1 \sqcup V_2, E);\\ H|_{V_1} = H_1;\\ H|_{V_2} = H_2}} \operatorname{Iso}(H, G)$$

as triples (H, E, φ) consisting of a graph $H = (V_1 \sqcup V_2, E)$, its set of edges E, and an isomorphism $\varphi \in \operatorname{Iso}(H, G)$. Let us also agree to encode the elements of the set

$$\bigsqcup_{\substack{(W_1,W_2);\\ W_1 \sqcup W_2 = V}} \operatorname{Iso}\left(H_1, G\mid_{W_1}\right) \times \operatorname{Iso}\left(H_2, G\mid_{W_2}\right)$$

as triples $((W_1, W_2), \varphi_1, \varphi_2)$ consisting of a pair (W_1, W_2) of subsets of V (which satisfies $W_1 \sqcup W_2 = V$), an isomorphism $\varphi_1 \in \operatorname{Iso}\left(H_1, G\mid_{W_1}\right)$, and an isomorphism $\varphi_2 \in \operatorname{Iso}\left(H_2, G\mid_{W_2}\right)$.

isomorphism. Then, $(\mathcal{G}^o)^o = \mathcal{G}$ as Hopf algebras[1246]. Therefore, $\psi : (\mathcal{G}^o)^o \to \mathcal{G}^o$ is a $\mathbf{k}$-algebra morphism (since $\psi : \mathcal{G} \to \mathcal{G}^o$ is a $\mathbf{k}$-algebra morphism). Also, $\mathcal{G}^o$ is of finite type (since $\mathcal{G}$ is of finite type). Thus, Exercise 1.6.1(f) (applied to $C = \mathcal{G}$, $D = \mathcal{G}^o$ and $f = \psi$) yields that $\psi : \mathcal{G} \to \mathcal{G}^o$ is a $\mathbf{k}$-coalgebra morphism if and only if $\psi^* : (\mathcal{G}^o)^o \to \mathcal{G}^o$ is a $\mathbf{k}$-algebra morphism.

Let us now prove that $\psi^* = \psi$ (as maps from $\mathcal{G}$ to $\mathcal{G}^o$). Indeed, let $a \in \mathcal{G}$. Then, every $b \in \mathcal{G}$ satisfies

$$\begin{aligned}
(\psi^*(a))(b) &= a(\psi(b)) && \left(\begin{array}{c}\text{by the definition of the adjoint map } \psi^* : (\mathcal{G}^o)^o \to \mathcal{G}^o, \\ \text{where } a \in \mathcal{G} \text{ is regarded as an element of } (\mathcal{G}^o)^o\end{array}\right) \\
&= (\psi(b))(a) && \left(\begin{array}{c}\text{because the } \mathbf{k}\text{-module homomorphism } \mathcal{G} \to (\mathcal{G}^o)^o \text{ which we use to} \\ \text{identify } \mathcal{G} \text{ with } (\mathcal{G}^o)^o \text{ sends } a \in \mathcal{G} \text{ to the map } \mathcal{G}^o \to \mathbf{k},\ f \mapsto f(a)\end{array}\right) \\
&= (b, a) && (\text{by Proposition 7.3.13(a), applied to } b \text{ and } a \text{ instead of } a \text{ and } b) \\
&= (a, b) && (\text{since the form } (\cdot,\cdot) : \mathcal{G} \times \mathcal{G} \to \mathbf{k} \text{ is symmetric}) \\
&= (\psi(a))(b) && (\text{by Proposition 7.3.13(a)}).
\end{aligned}$$

Hence, $\psi^*(a) = \psi(a)$. Now, let us forget that we fixed a. We thus have shown that $\psi^*(a) = \psi(a)$ for every $a \in \mathcal{G}$. In other words, $\psi^* = \psi$. Hence, $\psi^* : (\mathcal{G}^o)^o \to \mathcal{G}^o$ is a $\mathbf{k}$-algebra morphism (since $\psi : (\mathcal{G}^o)^o \to \mathcal{G}^o$ is a $\mathbf{k}$-algebra morphism). Thus, $\psi : \mathcal{G} \to \mathcal{G}^o$ is a $\mathbf{k}$-coalgebra morphism (because we know that $\psi : \mathcal{G} \to \mathcal{G}^o$ is a $\mathbf{k}$-coalgebra morphism if and only if $\psi^* : (\mathcal{G}^o)^o \to \mathcal{G}^o$ is a $\mathbf{k}$-algebra morphism). This yields that $\psi : \mathcal{G} \to \mathcal{G}^o$ is a $\mathbf{k}$-bialgebra morphism (since we already know that $\psi : \mathcal{G} \to \mathcal{G}^o$ is a $\mathbf{k}$-algebra morphism). Thus, $\psi : \mathcal{G} \to \mathcal{G}^o$ is a Hopf algebra homomorphism (by Corollary 1.4.27, applied to $H_1 = \mathcal{G}$, $H_2 = \mathcal{G}^o$, $S_1 = S_{\mathcal{G}}$, $S_2 = S_{\mathcal{G}^o}$ and $\beta = \psi$). This proves Proposition 7.3.13(b).

Now, we claim that there exists a bijection from the set

$$\bigsqcup_{\substack{H=(V_1 \sqcup V_2, E);\\ H|_{V_1} = H_1;\\ H|_{V_2} = H_2}} \operatorname{Iso}(H, G)$$

to the set

$$\bigsqcup_{\substack{(W_1, W_2);\\ W_1 \sqcup W_2 = V}} \operatorname{Iso}\left(H_1, G\mid_{W_1}\right) \times \operatorname{Iso}\left(H_2, G\mid_{W_2}\right).$$

Namely, this bijection sends every (H, E, φ) (where $H = (V_1 \sqcup V_2, E)$ and $\varphi \in \operatorname{Iso}(H, G)$) to $((\varphi(V_1), \varphi(V_2)), \varphi_1, \varphi_2)$, where $\varphi_1 : V_1 \to \varphi(V_1)$ is the isomorphism from H_1 to $G\mid_{\varphi(V_1)}$ which is obtained by restricting φ to V_1, and where φ_2 is the isomorphism from H_2 to $G_2\mid_{\varphi(V_2)}$ which is obtained by restricting φ to V_2. (The inverse map of this bijection sends every $((W_1, W_2), \varphi_1, \varphi_2)$ (with $W_1 \sqcup W_2 = V$ and $\varphi_1 \in \operatorname{Iso}\left(H_1, G\mid_{W_1}\right)$ and $\varphi_2 \in \operatorname{Iso}\left(H_2, G\mid_{W_2}\right)$) to $((V_1 \sqcup V_2, E), E, \varphi)$, where $\varphi : V_1 \sqcup V_2 \to V$ is the map glued together from the maps $\varphi_1 : V_1 \to W_1$ and $\varphi_2 : V_2 \to W_2$, and where $E = \left(\varphi^{-1}\right)_*(F)$.) The existence of this bijection yields

$$\left|\bigsqcup_{\substack{H=(V_1 \sqcup V_2, E);\\ H|_{V_1} = H_1;\\ H|_{V_2} = H_2}} \operatorname{Iso}(H, G)\right| = \left|\bigsqcup_{\substack{(W_1, W_2);\\ W_1 \sqcup W_2 = V}} \operatorname{Iso}\left(H_1, G\mid_{W_1}\right) \times \operatorname{Iso}\left(H_2, G\mid_{W_2}\right)\right|.$$

Hence,

$$(\psi(ab))[G] = \left|\bigsqcup_{\substack{H=(V_1 \sqcup V_2, E);\\ H|_{V_1} = H_1;\\ H|_{V_2} = H_2}} \operatorname{Iso}(H, G)\right| = \left|\bigsqcup_{\substack{(W_1, W_2);\\ W_1 \sqcup W_2 = V}} \operatorname{Iso}\left(H_1, G\mid_{W_1}\right) \times \operatorname{Iso}\left(H_2, G\mid_{W_2}\right)\right| = (\psi(a) \cdot \psi(b))[G].$$

Let us now forget that we fixed G. We thus have shown that $(\psi(ab))[G] = (\psi(a) \cdot \psi(b))[G]$ for any finite graph G. In other words, the two maps $\psi(ab)$ and $\psi(a) \cdot \psi(b)$ are equal to each other on every element of the basis $([G])_{[G] \text{ is an isomorphism class of finite graphs}}$ of $\mathcal{G}$. Since these two maps are $\mathbf{k}$-linear, this yields that these two maps $\psi(ab)$ and $\psi(a) \cdot \psi(b)$ are identical. In other words, $\psi(ab) = \psi(a) \cdot \psi(b)$. Qed.

[1246] This is because the Hopf algebra structure on $\mathcal{G}^o$ was defined by taking adjoints of the structure maps of the Hopf algebra structure on $\mathcal{G}$ (for example, the comultiplication $\Delta_{\mathcal{G}^o}$ on $\mathcal{G}^o$ is the adjoint of the multiplication $m_{\mathcal{G}}$ on $\mathcal{G}$, if $(\mathcal{G} \otimes \mathcal{G})^o$ is identified with $\mathcal{G}^o \otimes \mathcal{G}^o$), and the Hopf algebra structure on $(\mathcal{G}^o)^o$ was defined by taking adjoints of the structure maps of the Hopf algebra structure on $\mathcal{G}^o$; but the adjoint of the adjoint of a linear map F between two graded $\mathbf{k}$-modules of finite type is the map F again.

(c) Assume that $\mathbb{Q}$ is a subring of $\mathbf{k}$. Then, $\operatorname{aut}(G)$ is invertible for every finite graph G (because $\operatorname{aut}(G)$ is a positive integer). Hence, $\left(\operatorname{aut}(G)\cdot [G]^*\right)_{[G] \text{ is an isomorphism class of finite graphs}}$ is a basis of the $\mathbf{k}$-module $\mathcal{G}^o$ (because $\left([G]^*\right)_{[G] \text{ is an isomorphism class of finite graphs}}$ is a basis of this $\mathbf{k}$-module). The map ψ thus sends the basis $\left([G]^\sharp\right)_{[G] \text{ is an isomorphism class of finite graphs}}$ of the $\mathbf{k}$-module $\mathcal{G}$ to the basis $\left(\operatorname{aut}(G)\cdot [G]^*\right)_{[G] \text{ is an isomorphism class of finite graphs}}$ of the $\mathbf{k}$-module $\mathcal{G}^o$ (because it satisfies $\psi\left([G]^\sharp\right) = \operatorname{aut}(G)\cdot [G]^*$ for every finite graph G). Hence, the $\mathbf{k}$-linear map ψ sends a basis of its domain to a basis of its codomain. Therefore, the map ψ is a $\mathbf{k}$-module isomorphism. Combined with the fact that ψ is a Hopf algebra homomorphism, this yields that ψ is a Hopf algebra isomorphism. Proposition 7.3.13(c) is thus proven. □

Now, all of Proposition 7.3.9, Proposition 7.3.11 and Proposition 7.3.13 are proven. Therefore, Exercise 7.3.14 is solved.

13.197. **Solution to Exercise 7.3.25.** *Solution to Exercise 7.3.25.* Proposition 7.3.9(b) yields that the elements $[G]^\sharp$, where $[G]$ ranges over all isomorphism classes of finite graphs, form a basis of the $\mathbf{k}$-module $\mathcal{G}$. Hence, we can define a $\mathbf{k}$-linear map $\Psi' : \mathcal{G} \to \mathbf{k}$ by requiring that

$$\Psi'\left([G]^\sharp\right) = \sum_{\substack{f:V\to\{1,2,3,\ldots\};\\ \operatorname{eqs} f = E}} \mathbf{x}_f \qquad \text{for every finite graph } G=(V,E)$$

(because $\sum_{\substack{f:V\to\{1,2,3,\ldots\};\\ \operatorname{eqs} f = E}} \mathbf{x}_f$ depends only on the isomorphism class $[G]$, but not on the graph G itself). Consider this map Ψ'. We shall show that $\Psi' = \Psi$.

Indeed, let $G=(V,E)$ be any finite graph. Then, Proposition 7.3.9(a) yields $[G] = \sum_{\substack{H=(V,E');\\ E'\cap E=\varnothing}} [H]^\sharp$. Applying the map Ψ' to both sides of this equality, we obtain

$$\Psi'[G] = \Psi'\left(\sum_{\substack{H=(V,E');\\ E'\cap E=\varnothing}} [H]^\sharp\right) = \sum_{\substack{H=(V,E');\\ E'\cap E=\varnothing}} \underbrace{\Psi'\left([H]^\sharp\right)}_{\substack{=\sum_{\substack{f:V\to\{1,2,3,\ldots\};\\ \operatorname{eqs} f=E'}} \mathbf{x}_f \\ \text{(by the definition of } \Psi')}} \qquad \text{(since the map } \Psi' \text{ is } \mathbf{k}\text{-linear)}$$

$$= \underbrace{\sum_{\substack{H=(V,E');\\ E'\cap E=\varnothing}}}_{=\sum_{\substack{E' \text{ is a set of two-element}\\ \text{subsets of } V;\\ E'\cap E=\varnothing}}} \sum_{\substack{f:V\to\{1,2,3,\ldots\};\\ \operatorname{eqs} f=E'}} \mathbf{x}_f = \underbrace{\sum_{\substack{E' \text{ is a set of two-element}\\ \text{subsets of } V;\\ E'\cap E=\varnothing}} \sum_{\substack{f:V\to\{1,2,3,\ldots\};\\ \operatorname{eqs} f=E'}}}_{=\sum_{\substack{f:V\to\{1,2,3,\ldots\};\\ (\operatorname{eqs} f)\cap E=\varnothing}}} \mathbf{x}_f$$

$$= \sum_{\substack{f:V\to\{1,2,3,\ldots\};\\ (\operatorname{eqs} f)\cap E=\varnothing}} \mathbf{x}_f. \tag{13.197.1}$$

However, if $f : V \to \{1,2,3,\ldots\}$ is any map, then we have the following logical equivalence:

$$((\operatorname{eqs} f)\cap E = \varnothing) \iff (f \text{ is a proper coloring of } G). \tag{13.197.2}$$

[1247] Hence, (13.197.1) becomes

$$\Psi'[G] = \underbrace{\sum_{\substack{f:V\to\{1,2,3,\ldots\};\\ (\operatorname{eqs} f)\cap E=\varnothing}}}_{\substack{=\sum_{\substack{f:V\to\{1,2,3,\ldots\};\\ f \text{ is a proper}\\ \text{coloring of } G}}\\ \text{(because of the equivalence (13.197.2))}}} \mathbf{x}_f = \underbrace{\sum_{\substack{f:V\to\{1,2,3,\ldots\};\\ f \text{ is a proper}\\ \text{coloring of } G}}}_{=\sum_{\substack{\text{proper colorings}\\ f:V\to\{1,2,3,\ldots\}}}} \mathbf{x}_f = \sum_{\substack{\text{proper colorings}\\ f:V\to\{1,2,3,\ldots\}}} \mathbf{x}_f$$

$$= \Psi[G] \qquad \left(\text{since Proposition 7.3.17 yields } \Psi[G] = \sum_{\substack{\text{proper colorings}\\ f:V\to\{1,2,3,\ldots\}}} \mathbf{x}_f\right).$$

Let us now forget that we fixed $[G]$. Thus, we have shown that $\Psi'[G] = \Psi[G]$ for any finite graph G. In other words, the two maps Ψ' and Ψ are equal to each other on each element of the family $([G])_{G \text{ is an isomorphism class of finite graphs}}$. Thus, the two maps Ψ' and Ψ must be identical (because these two maps are $\mathbf{k}$-linear, and because the family $([G])_{G \text{ is an isomorphism class of finite graphs}}$ is a basis of the $\mathbf{k}$-module $\mathcal{G}$). In other words, $\Psi' = \Psi$. Now, every finite graph $G = (V, E)$ satisfies

$$\underbrace{\Psi}_{=\Psi'}\left([G]^{\sharp}\right) = \Psi'\left([G]^{\sharp}\right) = \sum_{\substack{f:V\to\{1,2,3,\ldots\};\\ \operatorname{eqs} f=E}} \mathbf{x}_f.$$

This solves Exercise 7.3.25.

13.198. **Solution to Exercise 8.1.10.** *Solution to Exercise 8.1.10.* We begin with some preparations.

We shall regard permutations as words (over the alphabet $\{1,2,3,\ldots\}$), by identifying every permutation $\pi \in \mathfrak{S}_n$ with the word $(\pi(1), \pi(2), \ldots, \pi(n))$. For every $n \in \mathbb{N}$, the lexicographic order on words thus

[1247] *Proof of (13.197.2):* Let $f : V \to \{1,2,3,\ldots\}$ be any map.

Let us first assume that $(\operatorname{eqs} f) \cap E = \varnothing$. We shall then show that f is a proper coloring of G.

Indeed, let us assume that there exists an edge $e = \{v, v'\}$ in E such that $f(v) = f(v')$. Consider this edge $e = \{v, v'\}$. We are going to derive a contradiction.

The definition of $\operatorname{eqs} f$ yields $\operatorname{eqs} f = \{\{u, u'\} \mid u \in V,\ u' \in V,\ u \neq u' \text{ and } f(u) = f(u')\}$. But $v \in V$ and $v' \in V$ and $v \neq v'$ (since $\{v, v'\}$ is an edge of E) and $f(v) = f(v')$. Hence,

$$\{v, v'\} \in \left\{\{u, u'\} \mid u \in V,\ u' \in V,\ u \neq u' \text{ and } f(u) = f(u')\right\} = \operatorname{eqs} f.$$

Combined with $\{v, v'\} \in E$, this yields $\{v, v'\} \in (\operatorname{eqs} f) \cap E = \varnothing$. But this is absurd, because the empty set $\varnothing$ has no elements. Hence, we have found a contradiction. Thus, our assumption (that there exists an edge $e = \{v, v'\}$ in E such that $f(v) = f(v')$) was wrong. Hence, no edge $e = \{v, v'\}$ in E has $f(v) = f(v')$. In other words, f is a proper coloring of G (according to the definition of a "proper coloring").

Let us now forget that we assumed that $(\operatorname{eqs} f) \cap E = \varnothing$. We thus have proven the implication

$$((\operatorname{eqs} f) \cap E = \varnothing) \implies (f \text{ is a proper coloring of } G). \tag{13.197.3}$$

Let us now assume that f is a proper coloring of G. In other words, no edge $e = \{v, v'\}$ in E has $f(v) = f(v')$ (according to the definition of a "proper coloring").

Now, let us prove that $(\operatorname{eqs} f) \cap E = \varnothing$. Indeed, assume the contrary. Then, $(\operatorname{eqs} f) \cap E \neq \varnothing$. Hence, there exists some $g \in (\operatorname{eqs} f) \cap E$. Consider this g. We have

$$g \in (\operatorname{eqs} f) \cap E \subset \operatorname{eqs} f = \left\{\{u, u'\} \mid u \in V,\ u' \in V,\ u \neq u' \text{ and } f(u) = f(u')\right\}.$$

Hence, $g = \{u, u'\}$ for some $u \in V$ and $u' \in V$ satisfying $u \neq u'$ and $f(u) = f(u')$. Consider these u and u'. We have $u \neq u'$ and $\{u, u'\} = g \in (\operatorname{eqs} f) \cap E \subset E$. Hence, $\{u, u'\}$ is an edge in E. Moreover, $f(u) = f(u')$. Hence, the edge $\{u, u'\}$ is an edge $e = \{v, v'\}$ in E such that $f(v) = f(v')$ (namely, for $v = u$ and $v' = u'$). This contradicts the fact that no edge $e = \{v, v'\}$ in E has $f(v) = f(v')$. This contradiction proves that our assumption was wrong. Hence, we have shown that $(\operatorname{eqs} f) \cap E = \varnothing$.

Let us now forget that we assumed that (f is a proper coloring of G). We thus have proven the implication

$$(f \text{ is a proper coloring of } G) \implies ((\operatorname{eqs} f) \cap E = \varnothing).$$

Combining this with (13.197.3), we obtain the equivalence

$$((\operatorname{eqs} f) \cap E = \varnothing) \iff (f \text{ is a proper coloring of } G).$$

This proves (13.197.2).

defines a total order on $\mathfrak{S}_n$; we will be using this order in the following when we make statements like "$\sigma < \tau$" for σ and τ being two permutations in $\mathfrak{S}_n$.

We denote the empty word by $\varnothing$. This empty word is identified with the trivial permutation in $\mathfrak{S}_0$. We shall refer to $\varnothing$ as the *empty permutation*. Every permutation other than $\varnothing$ will be called a *nonempty permutation*.

Let $\mathfrak{S}$ denote the disjoint union $\bigsqcup_{n\in\mathbb{N}} \mathfrak{S}_n$ of the posets $\mathfrak{S}_n$. While each poset $\mathfrak{S}_n$ is actually totally ordered (by the lexicographic order), the disjoint union $\mathfrak{S}$ is not, since elements of different $\mathfrak{S}_n$ are incomparable.[1248]

For every $k \in \mathbb{N}$ and $\ell \in \mathbb{N}$ and any permutations $u \in \mathfrak{S}_k$ and $v \in \mathfrak{S}_\ell$, we define a permutation $u \square v$ by $u \square v = u \cdot v\,[k]$ (where "$u \cdot v\,[k]$" has to be read as "$u \cdot (v\,[k])$" rather than as "$(u \cdot v)\,[k]$"). [1249] Thus, we have defined a binary operation $\square$ on the set $\mathfrak{S}$. It is easy to see that $\mathfrak{S}$ becomes a monoid with respect to this binary operation $\square$; the neutral element of this monoid is $\varnothing \in \mathfrak{S}_0$. Hence, products of the form $w_1 \square w_2 \square \cdots \square w_k$ for k-tuples $(w_1, w_2, \ldots, w_k) \in \mathfrak{S}^k$ of permutations are well-defined (without specifying a bracketing). The monoid $(\mathfrak{S}, \square)$ is left-cancellative[1250] and right-cancellative[1251].

It is easy to see that

$$\alpha \square \gamma > \beta \square \gamma \qquad (13.198.1)$$

for any $\alpha \in \mathfrak{S}$, $\beta \in \mathfrak{S}$ and $\gamma \in \mathfrak{S}$ satisfying $\alpha > \beta$ [1252].

Recall how we defined a connected permutation. Since we are regarding permutations as words, we can rewrite this definition as follows: A permutation $p \in \mathfrak{S}_n$ is connected if and only if it is nonempty and no nonempty proper prefix[1253] of p is itself a permutation. Hence, if a nonempty prefix of a connected permutation $p \in \mathfrak{S}_n$ is itself a permutation, then this prefix must be p.

It is easy to see that a permutation $w \in \mathfrak{S}_n$ is connected if and only if n is a positive integer and there exist no nonempty permutations u and v satisfying $w = u \square v$. It is furthermore easy to see that for every permutation $w \in \mathfrak{S}$, there is a unique way to write w in the form $w = w_1 \square w_2 \square \cdots \square w_k$ for some $k \in \mathbb{N}$ and

[1248] We could define an ordering on all of $\mathfrak{S}$ by restricting the total order on words defined in Definition 6.1.1; but we prefer not to.

[1249] This is indeed a permutation, because the word $u \cdot v\,[k]$ has each integer from 1 to $k + \ell$ appear exactly once in it.

[1250] A semigroup $(M, \cdot)$ is said to be *left-cancellative* if and only if it has the following property: If a, b and c are three elements of M satisfying $a \cdot b = a \cdot c$, then $b = c$.

[1251] A semigroup $(M, \cdot)$ is said to be *right-cancellative* if and only if it has the following property: If a, b and c are three elements of M satisfying $b \cdot a = c \cdot a$, then $b = c$.

[1252] *Proof of (13.198.1):* Let $\alpha \in \mathfrak{S}$, $\beta \in \mathfrak{S}$ and $\gamma \in \mathfrak{S}$ satisfy $\alpha > \beta$. Notice that the $n \in \mathbb{N}$ satisfying $\alpha \in \mathfrak{S}_n$ and the $n \in \mathbb{N}$ satisfying $\beta \in \mathfrak{S}_n$ must be identical (because otherwise, α and β would be incomparable in the poset $\mathfrak{S}$). In other words, the words α and β have the same length. Let k be this length. Thus, $\alpha \in \mathfrak{S}_k$ and $\beta \in \mathfrak{S}_k$.

Now, the words α and β have the same length and satisfy $\alpha > \beta$ in the lexicographic order. Hence, every word δ satisfies $\alpha \cdot \delta > \beta \cdot \delta$ (where $\cdot$ denotes concatenation). Applying this to $\delta = \gamma\,[k]$, we obtain $\alpha \cdot \gamma\,[k] > \beta \cdot \gamma\,[k]$. Now, the definition of $\alpha \square \gamma$ yields $\alpha \square \gamma = \alpha \cdot \gamma\,[k] > \beta \cdot \gamma\,[k] = \beta \square \gamma$ (by the definition of $\beta \square \gamma$), and thus (13.198.1) is proven.

[1253] A *proper prefix* of a word w is defined as a word u such that there exists a nonempty word v satisfying $w = uv$.

some k-tuple $(w_1, w_2, \ldots, w_k) \in \mathfrak{C}\mathfrak{S}^k$ of connected permutations[1254]. We refer to the tuple $(w_1, w_2, \ldots, w_k)$ as the *connected decomposition* of w.

For every $n \in \mathbb{N}$ and $w \in \mathfrak{S}_n$, define an element F_w^{conn} of FQSym by setting $F_w^{\mathrm{conn}} = F_{w_1} F_{w_2} \cdots F_{w_k}$, where $(w_1, w_2, \ldots, w_k)$ is the connected decomposition of w. It is clear that this F_w^{conn} is an element of FQSym_n. Hence, for every $n \in \mathbb{N}$, the family $(F_w^{\mathrm{conn}})_{w \in \mathfrak{S}_n}$ is a family of elements of FQSym_n. We shall prove that it is a basis of the $\mathbf{k}$-module FQSym_n. Once this is proven, the claim of the exercise will quickly follow (as we will see later).

We need the following easy fact:

Lemma 13.198.1. *Let $u \in \mathfrak{S}$ and $v \in \mathfrak{S}$. Then,*

$$F_u F_v = F_{u \square v} + (\text{a sum of terms } F_t \text{ for } t \in \mathfrak{S} \text{ satisfying } t > u \square v).$$

(The symbol $>$ refers to the partial order on the poset $\mathfrak{S}$.)

Proof of Lemma 13.198.1. Let k and ℓ be the nonnegative integers satisfying $u \in \mathfrak{S}_k$ and $v \in \mathfrak{S}_\ell$. Write the words u and $v[k]$ in the forms $u = (u_1, u_2, \ldots, u_k)$ and $v[k] = (p_1, p_2, \ldots, p_\ell)$, respectively. Let $(c_1, c_2, \ldots, c_{k+\ell})$ denote the concatenation $u \cdot v[k] = (u_1, u_2, \ldots, u_k, p_1, p_2, \ldots, p_\ell)$. By the definition of $u \sqcup\!\sqcup\, v[k]$, we therefore have

$$u \sqcup\!\sqcup\, v[k] = \left\{ \left(c_{w(1)}, c_{w(2)}, \ldots, c_{w(k+\ell)} \right) \ : \ w \in \mathrm{Sh}_{k,\ell} \right\}_{\text{multiset}}. \tag{13.198.2}$$

[1255] We also have $(c_1, c_2, \ldots, c_{k+\ell}) = u \cdot v[k] = u \square v$.

[1254] *Proof.* Fix $w \in \mathfrak{S}$. We need to show that a decomposition of w in the form $w = w_1 \square w_2 \square \cdots \square w_k$ with $k \in \mathbb{N}$ and $(w_1, w_2, \ldots, w_k) \in \mathfrak{C}\mathfrak{S}^k$ exists and is unique.

The existence of such a decomposition $w = w_1 \square w_2 \square \cdots \square w_k$ is easy to check (just take a decomposition $w = w_1 \square w_2 \square \cdots \square w_k$ with $k \in \mathbb{N}$ and $(w_1, w_2, \ldots, w_k) \in (\mathfrak{S} \setminus \{\varnothing\})^k$ such that k is maximum, and argue that the maximality of k forces each of w_1, w_2, $\ldots$, w_k to be connected). It remains to prove that such a decomposition is unique. In other words, we need to prove the following statement:

> *Statement $\mathcal{U}$:* If $w \in \mathfrak{S}$, $k \in \mathbb{N}$, $(w_1, w_2, \ldots, w_k) \in \mathfrak{C}\mathfrak{S}^k$, $\ell \in \mathbb{N}$ and $(v_1, v_2, \ldots, v_\ell) \in \mathfrak{C}\mathfrak{S}^\ell$ are such that $w = w_1 \square w_2 \square \cdots \square w_k$ and $w = v_1 \square v_2 \square \cdots \square v_\ell$, then $k = \ell$ and $(w_1, w_2, \ldots, w_k) = (v_1, v_2, \ldots, v_\ell)$.

Proof of Statement $\mathcal{U}$: Let us prove Statement $\mathcal{U}$ by strong induction by the size of w (that is, the number $n \in \mathbb{N}$ satisfying $w \in \mathfrak{S}_n$). So let $N \in \mathbb{N}$, and assume (as the induction hypothesis) that Statement $\mathcal{U}$ is proven in the case when $w \in \mathfrak{S}_M$ for $M < N$.

Let $w \in \mathfrak{S}_N$, $k \in \mathbb{N}$, $(w_1, w_2, \ldots, w_k) \in \mathfrak{C}\mathfrak{S}^k$, $\ell \in \mathbb{N}$ and $(v_1, v_2, \ldots, v_\ell) \in \mathfrak{C}\mathfrak{S}^\ell$ be such that $w = w_1 \square w_2 \square \cdots \square w_k$ and $w = v_1 \square v_2 \square \cdots \square v_\ell$. We need to show that $k = \ell$ and $(w_1, w_2, \ldots, w_k) = (v_1, v_2, \ldots, v_\ell)$.

All of the permutations w_1, w_2, $\ldots$, w_k and v_1, v_2, $\ldots$, v_ℓ are connected, and thus nonempty.

We assume WLOG that w is nonempty (else, Statement $\mathcal{U}$ is obvious). Hence, $k > 0$ and $\ell > 0$. Thus, w_1 and v_1 are well-defined. We shall now prove that $w_1 = v_1$.

By the definition of the operation $\square$, it is clear that the word v_1 is a prefix of the word $v_1 \square v_2 \square \cdots \square v_\ell$. In other words, the word v_1 is a prefix of the word w (since $w = v_1 \square v_2 \square \cdots \square v_\ell$). Similarly, the word w_1 is a prefix of the word w as well. But it is well-known that if α and β are two prefixes of a word γ, then either α is a prefix of β or β is a prefix of α. Applying this to $\alpha = v_1$, $\beta = w_1$ and $\gamma = w$, we conclude that either v_1 is a prefix of w_1 or w_1 is a prefix of v_1. We WLOG assume that v_1 is a prefix of w_1.

But recall that if a nonempty prefix of a connected permutation $p \in \mathfrak{S}_n$ is itself a permutation, then this prefix must be p. Applying this to $p = w_1$ and the nonempty prefix v_1 of w_1, we conclude that v_1 must be w_1. That is, $v_1 = w_1$.

Let now w' be the permutation $w_2 \square w_3 \square \cdots \square w_k$. Then,

$$\underbrace{v_1}_{=w_1} \square \underbrace{w'}_{=w_2 \square w_3 \square \cdots \square w_k} = w_1 \square (w_2 \square w_3 \square \cdots \square w_k) = w_1 \square w_2 \square \cdots \square w_k = w = v_1 \square v_2 \square \cdots \square v_\ell = v_1 \square (v_2 \square v_3 \square \cdots \square v_\ell).$$

Since the monoid $(\mathfrak{S}, \square)$ is left-cancellative, we can cancel v_1 from this equality, and obtain $w' = v_2 \square v_3 \square \cdots \square v_\ell$.

Now, let M be the length of the word w'. Since w' is shorter than the word $v_1 \square w' = w$ (because v_1 is nonempty), this length M is smaller than the length of w, which is N (since $w \in \mathfrak{S}_N$). Thus, $M < N$. Thus, we have $w' \in \mathfrak{S}_M$ with $M < N$. Hence, by the induction hypothesis, we can apply Statement $\mathcal{U}$ to w', $k-1$, $(w_2, w_3, \ldots, w_k)$, $\ell - 1$ and $(v_2, v_3, \ldots, v_\ell)$ instead of w, k, $(w_1, w_2, \ldots, w_k)$, ℓ and $(v_1, v_2, \ldots, v_\ell)$ (since $w' = w_2 \square w_3 \square \cdots \square w_k$ and $v' = v_2 \square v_3 \square \cdots \square v_\ell$). As a consequence, we obtain $k - 1 = \ell - 1$ and $(w_2, w_3, \ldots, w_k) = (v_2, v_3, \ldots, v_\ell)$. From $k - 1 = \ell - 1$, we obtain $k = \ell$. Combining $w_1 = v_1$ with $(w_2, w_3, \ldots, w_k) = (v_2, v_3, \ldots, v_\ell)$, we conclude $(w_1, w_2, \ldots, w_k) = (v_1, v_2, \ldots, v_\ell)$. Thus, the induction step is complete, and Statement $\mathcal{U}$ is proven. Hence, the uniqueness of the decomposition is proven, and we are done.

[1255] See Definition 1.6.2 for the definition of $\mathrm{Sh}_{k,\ell}$.

From (8.1.1), we have

$$\begin{aligned}
F_u F_v &= \sum_{w \in u \sqcup\!\sqcup v[k]} F_w \\
&= \sum_{w \in \mathrm{Sh}_{k,\ell}} F_{\left(c_{w(1)}, c_{w(2)}, \ldots, c_{w(k+\ell)}\right)} \qquad \text{(by (13.198.2))} \\
&= \underbrace{F_{\left(c_{\mathrm{id}(1)}, c_{\mathrm{id}(2)}, \ldots, c_{\mathrm{id}(k+\ell)}\right)}}_{\substack{=F_{u \square v} \\ \text{(since } \left(c_{\mathrm{id}(1)}, c_{\mathrm{id}(2)}, \ldots, c_{\mathrm{id}(k+\ell)}\right) \\ =(c_1, c_2, \ldots, c_{k+\ell}) = u \square v)}} + \sum_{\substack{w \in \mathrm{Sh}_{k,\ell}; \\ w \neq \mathrm{id}}} F_{\left(c_{w(1)}, c_{w(2)}, \ldots, c_{w(k+\ell)}\right)} \\
&\qquad \text{(here, we have split off the addend for } w = \mathrm{id} \text{ from the sum, since } \mathrm{id} \in \mathrm{Sh}_{k,\ell}) \\
&= F_{u \square v} + \sum_{\substack{w \in \mathrm{Sh}_{k,\ell}; \\ w \neq \mathrm{id}}} F_{\left(c_{w(1)}, c_{w(2)}, \ldots, c_{w(k+\ell)}\right)}. \qquad (13.198.3)
\end{aligned}$$

But it is easy to see that

$$\left(c_{w(1)}, c_{w(2)}, \ldots, c_{w(k+\ell)}\right) > u \square v \qquad \text{for every } w \in \mathrm{Sh}_{k,\ell} \text{ satisfying } w \neq \mathrm{id}. \qquad (13.198.4)$$

[*Proof of (13.198.4):* Let $w \in \mathrm{Sh}_{k,\ell}$ be such that $w \neq \mathrm{id}$. The map w is a permutation, hence bijective and thus injective. There exists some $i \in \{1, 2, \ldots, k+\ell\}$ such that $w(i) \neq i$ (since $w \neq \mathrm{id}$). Let j be the smallest such i. Thus, $w(j) \neq j$, but

$$\text{every } i \in \{1, 2, \ldots, k+\ell\} \text{ satisfying } i < j \text{ satisfies } w(i) = i. \qquad (13.198.5)$$

We have $w \in \mathrm{Sh}_{k,\ell}$, and therefore $w^{-1}(1) < w^{-1}(2) < \cdots < w^{-1}(k)$ and $w^{-1}(k+1) < w^{-1}(k+2) < \cdots < w^{-1}(k+\ell)$. Thus, the restriction of the map w^{-1} to the set $\{1, 2, \ldots, k\}$ and the restriction of the map w^{-1} to the set $\{k+1, k+2, \ldots, k+\ell\}$ are strictly increasing.

Let us first show that $j \leq k$. Indeed, assume the contrary. Then, $j > k$. Hence, $k < j$. Therefore, every $i \in \{1, 2, \ldots, k\}$ satisfies $i \leq k < j$ and thus $w(i) = i$ (by (13.198.5)). Thus, $w(\{1, 2, \ldots, k\}) = \{1, 2, \ldots, k\}$, so that $w^{-1}(\{1, 2, \ldots, k\}) = \{1, 2, \ldots, k\}$ (since w is bijective). Notice also that $j \in \{k+1, k+2, \ldots, k+\ell\}$ (since $j > k$). Now,

$$\begin{aligned}
w^{-1}\left(\underbrace{\{k+1, k+2, \ldots, k+\ell\}}_{=\{1,2,\ldots,k+\ell\} \setminus \{1,2,\ldots,k\}}\right) &= w^{-1}(\{1, 2, \ldots, k+\ell\} \setminus \{1, 2, \ldots, k\}) \\
&= \underbrace{w^{-1}(\{1, 2, \ldots, k+\ell\})}_{=\{1,2,\ldots,k+\ell\}} \setminus \underbrace{w^{-1}(\{1, 2, \ldots, k\})}_{=\{1,2,\ldots,k\}} \\
&\qquad \text{(since } w \text{ is bijective)} \\
&= \{1, 2, \ldots, k+\ell\} \setminus \{1, 2, \ldots, k\} = \{k+1, k+2, \ldots, k+\ell\}.
\end{aligned}$$

Therefore, w^{-1} restricts to a map from $\{k+1, k+2, \ldots, k+\ell\}$ to $\{k+1, k+2, \ldots, k+\ell\}$. This latter map must be strictly increasing (since the restriction of the map w^{-1} to the set $\{k+1, k+2, \ldots, k+\ell\}$ is strictly increasing), and therefore is the identity map (because the only strictly increasing map from $\{k+1, k+2, \ldots, k+\ell\}$ to $\{k+1, k+2, \ldots, k+\ell\}$ is the identity map). In other words, $w^{-1}(i) = \mathrm{id}(i) = i$ for every $i \in \{k+1, k+2, \ldots, k+\ell\}$. Applied to $i = j$, this yields $w^{-1}(j) = j$ (since we know that $j \in \{k+1, k+2, \ldots, k+\ell\}$), whence $w(j) = j$. But this contradicts $w(j) \neq j$. This contradiction proves that our assumption was wrong. Thus, $j \leq k$. Therefore,

$$\begin{aligned}
c_j &= u_j \qquad \text{(since } (c_1, c_2, \ldots, c_{k+\ell}) = (u_1, u_2, \ldots, u_k, p_1, p_2, \ldots, p_\ell)) \\
&= u(j) \leq k \qquad \text{(since } u \in \mathfrak{S}_k). \qquad (13.198.6)
\end{aligned}$$

If we had $w(j) < j$, then we would have $w(w(j)) = w(j)$ (by (13.198.5), applied to $i = w(j)$), which would lead to $w(j) = j$ (since w is injective), which would contradict $w(j) \neq j$. Hence, we cannot have $w(j) < j$. Thus, we have $w(j) \geq j$, so that $w(j) > j$ (since $w(j) \neq j$). In other words, $j < w(j)$.

Let us next prove that $w(j) > k$. Indeed, assume the contrary. Thus, $w(j) \leq k$. Thus, $w(j)$ and j are two elements of $\{1, 2, \ldots, k\}$ (because $w(j) \leq k$ and $j \leq k$) satisfying $j < w(j)$. Hence, $w^{-1}(j) < w^{-1}(w(j))$ (since $w^{-1}(1) < w^{-1}(2) < \cdots < w^{-1}(k)$). Hence, $w^{-1}(j) < w^{-1}(w(j)) = j$. Therefore, (13.198.5) (applied to $i = w^{-1}(j)$) yields $w\left(w^{-1}(j)\right) = w^{-1}(j)$, so that $w^{-1}(j) = w\left(w^{-1}(j)\right) = j$. This contradicts $w^{-1}(j) < j$. This contradiction proves that our assumption was wrong. Hence, $w(j) > k$ is proven.

Now, write v in the form $(v_1, v_2, \ldots, v_\ell)$. Then, $v[k] = (k + v_1, k + v_2, \ldots, k + v_\ell)$ (by the definition of $v[k]$). Since $w(j) > k$, we have

$$\begin{aligned} c_{w(j)} &= p_{w(j)-k} \qquad \left(\text{since } (c_1, c_2, \ldots, c_{k+\ell}) = (u_1, u_2, \ldots, u_k, p_1, p_2, \ldots, p_\ell)\right) \\ &= k + \underbrace{v_{w(j)-k}}_{>0} \qquad \left(\text{since } (p_1, p_2, \ldots, p_\ell) = v[k] = (k + v_1, k + v_2, \ldots, k + v_\ell)\right) \\ &> k \geq c_j \qquad \left(\text{by (13.198.6)}\right). \end{aligned}$$

So we have $c_{w(j)} > c_j$, but on the other hand, for every $i \in \{1, 2, \ldots, k + \ell\}$ satisfying $i < j$, we have $c_{w(i)} = c_i$ (because (13.198.5) yields $w(i) = i$). Thus,

$$\begin{aligned} \left(c_{w(1)}, c_{w(2)}, \ldots, c_{w(k+\ell)}\right) &> (c_1, c_2, \ldots, c_{k+\ell}) \qquad \text{(by the definition of lexicographic order)} \\ &= u \,\square\, v, \end{aligned}$$

and thus (13.198.4) is proven.]

Hence, (13.198.3) becomes

$$\begin{aligned} F_u F_v &= F_{u \square v} + \underbrace{\sum_{\substack{w \in \operatorname{Sh}_{k,\ell}; \\ w \neq \operatorname{id}}} F_{\left(c_{w(1)}, c_{w(2)}, \ldots, c_{w(k+\ell)}\right)}}_{\substack{=(\text{a sum of terms } F_t \text{ for } t \in \mathfrak{S} \text{ satisfying } t > u \square v) \\ (\text{by (13.198.4)})}} \\ &= F_{u \square v} + (\text{a sum of terms } F_t \text{ for } t \in \mathfrak{S} \text{ satisfying } t > u \,\square\, v). \end{aligned}$$

This proves Lemma 13.198.1. □

Corollary 13.198.2. *Every $n \in \mathbb{N}$ and $w \in \mathfrak{S}_n$ satisfy*

$$F_w^{\operatorname{conn}} = F_w + (\textit{a sum of terms } F_t \textit{ for } t \in \mathfrak{S}_n \textit{ satisfying } t > w).$$

Proof of Corollary 13.198.2. We shall first prove that every $k \in \mathbb{N}$ and every $(w_1, w_2, \ldots, w_k) \in \mathfrak{S}^k$ satisfy

(13.198.7)
$$F_{w_1} F_{w_2} \cdots F_{w_k} = F_{w_1 \square w_2 \square \cdots \square w_k} + (\text{a sum of terms } F_t \text{ for } t \in \mathfrak{S} \text{ satisfying } t > w_1 \,\square\, w_2 \,\square \cdots \square\, w_k).$$

Proof of (13.198.7): We will prove (13.198.7) by induction over k:

Induction base: For $k = 0$, the statement of (13.198.7) is obvious. Hence, the induction base is complete.

Induction step: Let $K \in \mathbb{N}$. Assume that (13.198.7) holds for $k = K$. We need to prove that (13.198.7) holds for $k = K + 1$.

Let $(w_1, w_2, \ldots, w_{K+1}) \in \mathfrak{S}^{K+1}$. By the induction hypothesis, we can apply (13.198.7) to $k = K$, and thus obtain

$$F_{w_1} F_{w_2} \cdots F_{w_K} = F_{w_1 \square w_2 \square \cdots \square w_K} + (\text{a sum of terms } F_t \text{ for } t \in \mathfrak{S} \text{ satisfying } t > w_1 \,\square\, w_2 \,\square \cdots \square\, w_K).$$

In other words, there exists a finite family $(t_i)_{i \in I}$ of elements t of $\mathfrak{S}$ satisfying $t > w_1 \,\square\, w_2 \,\square \cdots \square\, w_K$ such that

$$F_{w_1} F_{w_2} \cdots F_{w_K} = F_{w_1 \square w_2 \square \cdots \square w_K} + \sum_{i \in I} F_{t_i}.$$

Consider this family $(t_i)_{i \in I}$. We have

(13.198.8)
$$t_i \,\square\, w_{K+1} > w_1 \,\square\, w_2 \,\square \cdots \square\, w_{K+1} \qquad \text{for every } i \in I.$$

[1256] Now,

$$\begin{aligned}
&F_{w_1} F_{w_2} \cdots F_{w_{K+1}} \\
&= \underbrace{\left(F_{w_1} F_{w_2} \cdots F_{w_K}\right)}_{= F_{w_1 \square w_2 \square \cdots \square w_K} + \sum_{i \in I} F_{t_i}} F_{w_{K+1}} \\
&= \left(F_{w_1 \square w_2 \square \cdots \square w_K} + \sum_{i \in I} F_{t_i} \right) F_{w_{K+1}} \\
&= \underbrace{F_{w_1 \square w_2 \square \cdots \square w_K} F_{w_{K+1}}}_{\substack{= F_{(w_1 \square w_2 \square \cdots \square w_K) \square w_{K+1}} + (\text{a sum of terms } F_t \text{ for } t \in \mathfrak{S} \text{ satisfying } t > (w_1 \square w_2 \square \cdots \square w_K) \square w_{K+1}) \\ (\text{by Lemma 13.198.1, applied to } u = w_1 \square w_2 \square \cdots \square w_K \text{ and } v = w_{K+1})}} \\
&\qquad + \sum_{i \in I} \underbrace{F_{t_i} F_{w_{K+1}}}_{\substack{= F_{t_i \square w_{K+1}} + (\text{a sum of terms } F_t \text{ for } t \in \mathfrak{S} \text{ satisfying } t > t_i \square w_{K+1}) \\ (\text{by Lemma 13.198.1, applied to } u = t_i \text{ and } v = w_{K+1})}} \\
&= F_{(w_1 \square w_2 \square \cdots \square w_K) \square w_{K+1}} + \left(\text{a sum of terms } F_t \text{ for } t \in \mathfrak{S} \text{ satisfying } t > (w_1 \square w_2 \square \cdots \square w_K) \square w_{K+1}\right) \\
&\qquad + \sum_{i \in I} \left(F_{t_i \square w_{K+1}} + \left(\text{a sum of terms } F_t \text{ for } t \in \mathfrak{S} \text{ satisfying } t > t_i \square w_{K+1}\right) \right) \\
&= F_{w_1 \square w_2 \square \cdots \square w_{K+1}} + \left(\text{a sum of terms } F_t \text{ for } t \in \mathfrak{S} \text{ satisfying } t > w_1 \square w_2 \square \cdots \square w_{K+1}\right) \\
&\qquad + \sum_{i \in I} \underbrace{\left(F_{t_i \square w_{K+1}} + \left(\text{a sum of terms } F_t \text{ for } t \in \mathfrak{S} \text{ satisfying } t > t_i \square w_{K+1}\right) \right)}_{= (\text{a sum of terms } F_t \text{ for } t \in \mathfrak{S} \text{ satisfying } t \geq t_i \square w_{K+1})} \\
&\qquad \left(\text{since } \left(w_1 \square w_2 \square \cdots \square w_K\right) \square w_{K+1} = w_1 \square w_2 \square \cdots \square w_{K+1}\right) \\
&= F_{w_1 \square w_2 \square \cdots \square w_{K+1}} + \left(\text{a sum of terms } F_t \text{ for } t \in \mathfrak{S} \text{ satisfying } t > w_1 \square w_2 \square \cdots \square w_{K+1}\right) \\
&\qquad + \sum_{i \in I} \underbrace{\left(\text{a sum of terms } F_t \text{ for } t \in \mathfrak{S} \text{ satisfying } t \geq t_i \square w_{K+1}\right)}_{\substack{= (\text{a sum of terms } F_t \text{ for } t \in \mathfrak{S} \text{ satisfying } t > w_1 \square w_2 \square \cdots \square w_{K+1}) \\ (\text{since every } t \in \mathfrak{S} \text{ satisfying } t \geq t_i \square w_{K+1} \text{ also satisfies} \\ t \geq t_i \square w_{K+1} > w_1 \square w_2 \square \cdots \square w_{K+1} \text{ (by (13.198.8)))}}} \\
&= F_{w_1 \square w_2 \square \cdots \square w_{K+1}} + \left(\text{a sum of terms } F_t \text{ for } t \in \mathfrak{S} \text{ satisfying } t > w_1 \square w_2 \square \cdots \square w_{K+1}\right) \\
&\qquad + \sum_{i \in I} \left(\text{a sum of terms } F_t \text{ for } t \in \mathfrak{S} \text{ satisfying } t > w_1 \square w_2 \square \cdots \square w_{K+1}\right) \\
&= F_{w_1 \square w_2 \square \cdots \square w_{K+1}} + \left(\text{a sum of terms } F_t \text{ for } t \in \mathfrak{S} \text{ satisfying } t > w_1 \square w_2 \square \cdots \square w_{K+1}\right).
\end{aligned}$$

In other words, (13.198.7) holds for $k = K+1$. This completes the induction step. Hence, (13.198.7) is proven by induction.

Now, let $n \in \mathbb{N}$ and $w \in \mathfrak{S}_n$. Let $(w_1, w_2, \ldots, w_k)$ be the connected decomposition of w. Then, $F_w^{\text{conn}} = F_{w_1} F_{w_2} \cdots F_{w_k}$ (by the definition of F_w^{conn}) and $w = w_1 \square w_2 \square \cdots \square w_k$ (by the definition of a connected decomposition). Now,

$$\begin{aligned}
F_w^{\text{conn}} &= F_{w_1} F_{w_2} \cdots F_{w_k} \\
&= F_{w_1 \square w_2 \square \cdots \square w_k} + \left(\text{a sum of terms } F_t \text{ for } t \in \mathfrak{S} \text{ satisfying } t > w_1 \square w_2 \square \cdots \square w_k\right) \\
&\qquad \left(\text{by (13.198.7)}\right) \\
&= F_w + \left(\text{a sum of terms } F_t \text{ for } t \in \mathfrak{S} \text{ satisfying } t > w\right) \qquad \left(\text{since } w_1 \square w_2 \square \cdots \square w_k = w\right).
\end{aligned}$$

Since every $t \in \mathfrak{S}$ satisfying $t > w$ must belong to $\mathfrak{S}_n$ [1257], this rewrites as

$$F_w^{\text{conn}} = F_w + \left(\text{a sum of terms } F_t \text{ for } t \in \mathfrak{S}_n \text{ satisfying } t > w\right).$$

[1256] *Proof of (13.198.8):* Let $i \in I$. Then, t_i is an element t of $\mathfrak{S}$ satisfying $t > w_1 \square w_2 \square \cdots \square w_K$. Hence, $t_i > w_1 \square w_2 \square \cdots \square w_K$. Thus, (13.198.1) (applied to $\alpha = t_i$, $\beta = w_1 \square w_2 \square \cdots \square w_K$ and $\gamma = w_{K+1}$) yields $t_i \square w_{K+1} > (w_1 \square w_2 \square \cdots \square w_K) \square w_{K+1} = w_1 \square w_2 \square \cdots \square w_{K+1}$, qed.

[1257] *Proof.* Let $t \in \mathfrak{S}$ be such that $t > w$. Recall that $w \in \mathfrak{S}_n$. Therefore, if we had $t \notin \mathfrak{S}_n$, then t and w would be incomparable in the poset $\mathfrak{S}$ (because of the construction of the poset $\mathfrak{S}$), which would contradict the fact that $t > w$. Hence, we cannot have $t \notin \mathfrak{S}_n$. Thus, $t \in \mathfrak{S}_n$, qed.

This proves Corollary 13.198.2. □

Now, let us fix $n \in \mathbb{N}$. The family $(F_w)_{w \in \mathfrak{S}_n}$ is a basis of the $\mathbf{k}$-module FQSym_n. We regard the set $\mathfrak{S}_n$ as a poset whose smaller relation is the relation $>$ inherited from $\mathfrak{S}$ (yes, you are reading it right: the order on $\mathfrak{S}_n$ is opposite to that on $\mathfrak{S}$). Then, Corollary 13.198.2 shows that the family $(F_w^{\operatorname{conn}})_{w \in \mathfrak{S}_n}$ expands unitriangularly[1258] in the family $(F_w)_{w \in \mathfrak{S}_n}$ (by Remark 11.1.17(c)). Thus, the family $(F_w^{\operatorname{conn}})_{w \in \mathfrak{S}_n}$ expands invertibly triangularly in the family $(F_w)_{w \in \mathfrak{S}_n}$. Consequently, Corollary 11.1.19(e) (applied to FQSym_n, $\mathfrak{S}_n$, $(F_w^{\operatorname{conn}})_{w \in \mathfrak{S}_n}$ and $(F_w)_{w \in \mathfrak{S}_n}$ instead of M, S, $(e_s)_{s \in S}$ and $(f_s)_{s \in S}$) shows that the family $(F_w^{\operatorname{conn}})_{w \in \mathfrak{S}_n}$ is a basis of the $\mathbf{k}$-module FQSym_n if and only if the family $(F_w)_{w \in \mathfrak{S}_n}$ is a basis of the $\mathbf{k}$-module FQSym_n. Hence, the family $(F_w^{\operatorname{conn}})_{w \in \mathfrak{S}_n}$ is a basis of the $\mathbf{k}$-module FQSym_n (since the family $(F_w)_{w \in \mathfrak{S}_n}$ is a basis of the $\mathbf{k}$-module FQSym_n).

Now, let us forget that we fixed n. We thus have proven that the family $(F_w^{\operatorname{conn}})_{w \in \mathfrak{S}_n}$ is a basis of the $\mathbf{k}$-module FQSym_n for every $n \in \mathbb{N}$. Hence, the family $(F_w^{\operatorname{conn}})_{w \in \mathfrak{S}}$ (being the disjoint union of the families $(F_w^{\operatorname{conn}})_{w \in \mathfrak{S}_n}$ for all $n \in \mathbb{N}$) is a basis of the $\mathbf{k}$-module $\bigoplus_{n \in \mathbb{N}} \operatorname{FQSym}_n = \operatorname{FQSym}$.

Now, it is easy to see that the family $(F_{w_1} F_{w_2} \cdots F_{w_k})_{k \in \mathbb{N};\ (w_1, w_2, \ldots, w_k) \in \mathfrak{CS}^k}$ is a reindexing of the family $(F_w^{\operatorname{conn}})_{w \in \mathfrak{S}}$ [1259]. Thus, the family $(F_{w_1} F_{w_2} \cdots F_{w_k})_{k \in \mathbb{N};\ (w_1, w_2, \ldots, w_k) \in \mathfrak{CS}^k}$ is a basis of the $\mathbf{k}$-module FQSym (since the family $(F_w^{\operatorname{conn}})_{w \in \mathfrak{S}}$ is a basis of the $\mathbf{k}$-module FQSym). In other words, FQSym is a free (noncommutative) $\mathbf{k}$-algebra with generators $(F_w)_{w \in \mathfrak{CS}}$. This solves Exercise 8.1.10.

13.199. **Solution to Exercise 11.1.11.** *Solution to Exercise 11.1.11.* Let us first agree on a convention: Whenever S is a poset, we let $\leq$ denote the smaller-or-equal relation of the poset S.

Before we prove Proposition 11.1.10, let us make a definition:

Definition 13.199.1. Let S be a poset.

(a) Let T_S denote the set of all triangular $S \times S$-matrices. Clearly, $\operatorname{T}_S \subset \mathbf{k}^{S \times S}$.
(b) Let IT_S denote the set of all invertibly triangular $S \times S$-matrices.
(c) Let UT_S denote the set of all unitriangular $S \times S$-matrices.

Now, we shall state several lemmas (most of them completely trivial, and stated merely for the purpose of easier reference).

Lemma 13.199.2. *Let S and T be two finite sets. Let $A = (a_{s,t})_{(s,t) \in S \times T}$ be an $S \times T$-matrix. Let $B = (b_{s,t})_{(s,t) \in T \times S}$ be a $T \times S$-matrix. Then,*

$$AB = \left(\sum_{k \in T} a_{s,k} b_{k,t} \right)_{(s,t) \in S \times S} .$$

Proof of Lemma 13.199.2. Lemma 13.199.2 is just a restatement of the definition of AB. □

[1258] See Definition 11.1.16 for the meaning of these words.

[1259] *Proof.* For every $k \in \mathbb{N}$ and $(w_1, w_2, \ldots, w_k) \in \mathfrak{CS}^k$, the connected decomposition of the permutation $w_1 \square w_2 \square \cdots \square w_k$ is the k-tuple $(w_1, w_2, \ldots, w_k)$ (by the definition of a connected decomposition). Hence, for every $k \in \mathbb{N}$ and $(w_1, w_2, \ldots, w_k) \in \mathfrak{CS}^k$, we have

$$F^{\operatorname{conn}}_{w_1 \square w_2 \square \cdots \square w_k} = F_{w_1} F_{w_2} \cdots F_{w_k} \qquad (13.198.9)$$

(by the definition of $F^{\operatorname{conn}}_{w_1 \square w_2 \square \cdots \square w_k}$).

Now, recall that for every permutation $w \in \mathfrak{S}$, there is a unique way to write w in the form $w = w_1 \square w_2 \square \cdots \square w_k$ for some $k \in \mathbb{N}$ and some k-tuple $(w_1, w_2, \ldots, w_k) \in \mathfrak{CS}^k$ of connected permutations. In other words, the map

$$\bigsqcup_{k \in \mathbb{N}} \mathfrak{CS}^k \to \mathfrak{S},$$
$$(w_1, w_2, \ldots, w_k) \mapsto w_1 \square w_2 \square \cdots \square w_k$$

is a bijection. Hence, the family $\left(F^{\operatorname{conn}}_{w_1 \square w_2 \square \cdots \square w_k} \right)_{k \in \mathbb{N};\ (w_1, w_2, \ldots, w_k) \in \mathfrak{CS}^k}$ is a reindexing of the family $(F_w^{\operatorname{conn}})_{w \in \mathfrak{S}}$. Due to (13.198.9), this rewrites as follows: The family $(F_{w_1} F_{w_2} \cdots F_{w_k})_{k \in \mathbb{N};\ (w_1, w_2, \ldots, w_k) \in \mathfrak{CS}^k}$ is a reindexing of the family $(F_w^{\operatorname{conn}})_{w \in \mathfrak{S}}$. Qed.

Lemma 13.199.3. *Let S be a poset. Let $A \in \mathrm{T}_S$. Then, A is a triangular $S \times S$-matrix.*

Proof of Lemma 13.199.3. Lemma 13.199.3 follows from the definition of T_S. □

Lemma 13.199.4. *Let S be a poset. Let $A = (a_{s,t})_{(s,t)\in S\times S}$ be a triangular $S \times S$-matrix. Then, every $(s,t) \in S \times S$ which does not satisfy $t \leq s$ must satisfy $a_{s,t} = 0$.*

Proof of Lemma 13.199.4. Lemma 13.199.4 follows from the definition of "triangular". □

Lemma 13.199.5. *Let S be a poset. Let $A = (a_{s,t})_{(s,t)\in S\times S}$ be an $S \times S$-matrix. Assume that every $(s,t) \in S \times S$ which does not satisfy $t \leq s$ must satisfy $a_{s,t} = 0$. Then, $A \in \mathrm{T}_S$.*

Proof of Lemma 13.199.5. Lemma 13.199.5 follows from the definitions of T_S and of "triangular". □

Lemma 13.199.6. *Let S be a poset. Let $A \in \mathrm{IT}_S$. Then, A is an invertibly triangular $S \times S$-matrix.*

Proof of Lemma 13.199.6. Lemma 13.199.6 follows from the definition of IT_S. □

Lemma 13.199.7. *Let S be a poset. Let A be an invertibly triangular $S \times S$-matrix. Then, $A \in \mathrm{T}_S$.*

Proof of Lemma 13.199.7. Lemma 13.199.7 follows from the definition of "invertibly triangular". □

Lemma 13.199.8. *Let S be a poset. Let $A = (a_{s,t})_{(s,t)\in S\times S}$ be an invertibly triangular $S \times S$-matrix. Then, for every $s \in S$, the element $a_{s,s}$ of $\mathbf{k}$ is invertible.*

Proof of Lemma 13.199.8. Lemma 13.199.8 follows from the definition of "invertibly triangular". □

Lemma 13.199.9. *Let S be a poset. Let $A = (a_{s,t})_{(s,t)\in S\times S}$ be an $S \times S$-matrix. Assume that $A \in \mathrm{T}_S$. Assume that, for every $s \in S$, the element $a_{s,s}$ of $\mathbf{k}$ is invertible. Then, $A \in \mathrm{IT}_S$.*

Proof of Lemma 13.199.9. Lemma 13.199.9 follows from the definitions of T_S, of IT_S and of "invertibly triangular". □

Lemma 13.199.10. *Let S be a poset. We have $\mathrm{IT}_S \subset \mathrm{T}_S$.*

Proof of Lemma 13.199.10. Lemma 13.199.10 just says that every invertibly triangular $S \times S$-matrix is triangular; this follows from the definition of "invertibly triangular". □

Lemma 13.199.11. *Let S be a finite poset. Let $A = (a_{s,t})_{(s,t)\in S\times S} \in \mathrm{T}_S$ and $B = (b_{s,t})_{(s,t)\in S\times S} \in \mathrm{T}_S$. Then, for every $s \in S$, we have*

$$\sum_{k\in S} a_{s,k} b_{k,s} = a_{s,s} b_{s,s}.$$

Proof of Lemma 13.199.11. We have $A \in \mathrm{T}_S$. Hence, Lemma 13.199.3 shows that A is a triangular $S \times S$-matrix. Lemma 13.199.4 thus shows that

(13.199.1) every $(s,t) \in S \times S$ which does not satisfy $t \leq s$ must satisfy $a_{s,t} = 0$.

The same argument (applied to B and $b_{s,t}$ instead of A and $a_{s,t}$) shows that

(13.199.2) every $(s,t) \in S \times S$ which does not satisfy $t \leq s$ must satisfy $b_{s,t} = 0$.

Let $s \in S$. Every $k \in S$ satisfying $k \neq s$ must satisfy

(13.199.3) $$a_{s,k} b_{k,s} = 0$$

[1260]. Now, for every $s \in S$, we have

$$\sum_{k\in S} a_{s,k} b_{k,s} = a_{s,s} b_{s,s} + \sum_{\substack{k\in S;\\ k\neq s}} \underbrace{a_{s,k} b_{k,s}}_{\substack{=0\\ \text{(by (13.199.3))}}}$$

(here, we have split off the addend for $k = s$ from the sum)

$$= a_{s,s} b_{s,s} + \underbrace{\sum_{\substack{k\in S;\\ k\neq s}} 0}_{=0} = a_{s,s} b_{s,s}.$$

This proves Lemma 13.199.11. □

Lemma 13.199.12. *Let S be a poset. Let $A \in \mathrm{UT}_S$. Then, A is a unitriangular $S \times S$-matrix.*

Proof of Lemma 13.199.12. Lemma 13.199.12 follows from the definition of UT_S. □

Lemma 13.199.13. *Let S be a poset. Let A be a unitriangular $S \times S$-matrix. Then, $A \in \mathrm{T}_S$.*

Proof of Lemma 13.199.13. Lemma 13.199.13 follows from the definitions of T_S and of "unitriangular". □

Lemma 13.199.14. *Let S be a poset. Let $A = (a_{s,t})_{(s,t)\in S\times S}$ be a unitriangular $S \times S$-matrix. Then, for every $s \in S$, we have $a_{s,s} = 1$.*

Proof of Lemma 13.199.14. Lemma 13.199.14 follows from the definition of "unitriangular". □

Lemma 13.199.15. *Let S be a poset. Let $A = (a_{s,t})_{(s,t)\in S\times S}$ be an $S \times S$-matrix. Assume that $A \in \mathrm{T}_S$. Assume that, for every $s \in S$, we have $a_{s,s} = 1$. Then, $A \in \mathrm{UT}_S$.*

Proof of Lemma 13.199.15. Lemma 13.199.15 follows from the definitions of T_S, of UT_S and of "unitriangular". □

Lemma 13.199.16. *Let S be a poset. We have $\mathrm{UT}_S \subset \mathrm{T}_S$.*

Proof of Lemma 13.199.16. Lemma 13.199.16 just says that every unitriangular $S \times S$-matrix is triangular; this follows from the definition of "unitriangular". □

Lemma 13.199.17. *We have $\mathrm{UT}_S \subset \mathrm{IT}_S$.*

Proof of Lemma 13.199.17. Lemma 13.199.17 just says that every unitriangular $S \times S$-matrix is invertibly triangular. In order to prove it, we only need to compare the definitions of "unitriangular" and of "invertibly triangular", and observe that the former definition makes a stronger requirement than the latter (indeed, if $a_{s,s} = 1$, then $a_{s,s}$ is invertible). □

Proof of Proposition 11.1.10. We begin with some preparations.

For any $(s,t) \in S \times S$, we define a subset $[t,s]$ of S by

$$[t,s] = \{q \in S \mid t \leq q \leq s\}.$$

This subset $[t,s]$ is called the *interval of S bounded by t and s*. (Notice that it can be an empty set, since we have not required that $t \leq s$.) The following facts hold:

[1260] *Proof of (13.199.3):* Let $k \in S$ be such that $k \neq s$. We must prove (13.199.3).

We are in one of the following two cases:

Case 1: We have $s \leq k$.

Case 2: We do not have $s \leq k$.

Let us first consider Case 1. In this case, we have $s \leq k$. If we had $k \leq s$, then we would have $k = s$ (since $k \leq s \leq k$); but this would contradict $k \neq s$. Hence, we cannot have $k \leq s$. Thus, (13.199.1) (applied to (s,k) instead of (s,t)) shows that $a_{s,k} = 0$. Thus, $\underbrace{a_{s,k}}_{=0} b_{k,s} = 0$. Hence, (13.199.3) is proven in Case 1.

Let us now consider Case 2. In this case, we do not have $s \leq k$. Hence, (13.199.2) (applied to (k,s) instead of (s,t)) shows that $b_{k,s} = 0$. Thus, $a_{s,k} \underbrace{b_{k,s}}_{=0} = 0$. Hence, (13.199.3) is proven in Case 2.

Now, we have proven (13.199.3) in each of the two Cases 1 and 2. Hence, (13.199.3) always holds. Qed.

- For any $(s,t) \in S \times S$, the set $[t,s]$ is finite (since it is a subset of the finite set S). Hence, for any $(s,t) \in S \times S$, the cardinality $|[t,s]|$ is a well-defined nonnegative integer.
- If s, t and u are three elements of S such that $s \leq t < u$, then

$$[s,t] \text{ is a proper subset of } [s,u]. \tag{13.199.4}$$

(Indeed, it is straightforward to show that $[s,t]$ is a subset of $[s,u]$. To prove that this subset is proper, it suffices to observe that u belongs to $[s,u]$ but not to $[s,t]$.)
- If s, t and u are three elements of S such that $s < t \leq u$, then

$$[t,u] \text{ is a proper subset of } [s,u]. \tag{13.199.5}$$

(Indeed, it is straightforward to show that $[t,u]$ is a subset of $[s,u]$. To prove that this subset is proper, it suffices to observe that s belongs to $[s,u]$ but not to $[t,u]$.)

Let us now prove Proposition 11.1.10.

(a) Let $0_{S\times S} \in \mathbf{k}^{S\times S}$ be the matrix $(0)_{(s,t)\in S\times S}$. This matrix $0_{S\times S}$ is the zero of the $\mathbf{k}$-algebra $\mathbf{k}^{S\times S}$.

We shall now prove the following five claims:

Claim A1: We have $0_{S\times S} \in \mathrm{T}_S$.

Claim A2: For every $A \in \mathrm{T}_S$ and $B \in \mathrm{T}_S$, we have $A + B \in \mathrm{T}_S$.

Claim A3: For every $A \in \mathrm{T}_S$ and $u \in \mathbf{k}$, we have $uA \in \mathrm{T}_S$.

Claim A4: We have $I_S \in \mathrm{T}_S$.

Claim A5: For every $A \in \mathrm{T}_S$ and $B \in \mathrm{T}_S$, we have $AB \in \mathrm{T}_S$.

Proof of Claim A1: The definition of $0_{S\times S}$ yields $0_{S\times S} = (0)_{(s,t)\in S\times S}$. Clearly, every $(s,t) \in S \times S$ which does not satisfy $t \leq s$ must satisfy $0 = 0$. Thus, Lemma 13.199.5 (applied to $0_{S\times S}$ and 0 instead of A and $a_{s,t}$) shows that $0_{S\times S} \in \mathrm{T}_S$. This proves Claim A1.

Proof of Claim A2: Let $A \in \mathrm{T}_S$ and $B \in \mathrm{T}_S$. Write the $S \times S$-matrix A in the form $A = (a_{s,t})_{(s,t)\in S\times S}$. Write the $S \times S$-matrix B in the form $B = (b_{s,t})_{(s,t)\in S\times S}$.

Lemma 13.199.3 shows that A is a triangular $S \times S$-matrix (since $A \in \mathrm{T}_S$). Lemma 13.199.4 shows that

$$\text{every } (s,t) \in S \times S \text{ which does not satisfy } t \leq s \text{ must satisfy } a_{s,t} = 0. \tag{13.199.6}$$

The same argument (applied to B and $b_{s,t}$ instead of A and $a_{s,t}$) shows that

$$\text{every } (s,t) \in S \times S \text{ which does not satisfy } t \leq s \text{ must satisfy } b_{s,t} = 0. \tag{13.199.7}$$

Adding the equalities $A = (a_{s,t})_{(s,t)\in S\times S}$ and $B = (b_{s,t})_{(s,t)\in S\times S}$, we obtain

$$A + B = (a_{s,t})_{(s,t)\in S\times S} + (b_{s,t})_{(s,t)\in S\times S} = (a_{s,t} + b_{s,t})_{(s,t)\in S\times S}$$

(by the definition of the sum of two $S \times S$-matrices). But every $(s,t) \in S \times S$ which does not satisfy $t \leq s$ must satisfy $\underbrace{a_{s,t}}_{\substack{=0 \\ \text{(by (13.199.6))}}} + \underbrace{b_{s,t}}_{\substack{=0 \\ \text{(by (13.199.7))}}} = 0+0 = 0$. Hence, Lemma 13.199.5 (applied to $A+B$ and $a_{s,t}+b_{s,t}$ instead of A and $a_{s,t}$) shows that $A + B \in \mathrm{T}_S$ (since $A + B = (a_{s,t} + b_{s,t})_{(s,t)\in S\times S}$). This proves Claim A2.

Proof of Claim A3: The proof of Claim A3 is similar to our proof of Claim A2 above, and is left to the reader.

Proof of Claim A4: The definition of I_S yields $I_S = (\delta_{s,t})_{(s,t)\in S\times S}$. But every $(s,t) \in S \times S$ which does not satisfy $t \leq s$ must satisfy $\delta_{s,t} = 0$ [1261]. Hence, Lemma 13.199.5 (applied to I_S and $\delta_{s,t}$ instead of A and $a_{s,t}$) shows that $I_S \in \mathrm{T}_S$. This proves Claim A4.

Proof of Claim A5: Let $A \in \mathrm{T}_S$ and $B \in \mathrm{T}_S$.

As in the proof of Claim A2, we can see that the statements (13.199.6) and (13.199.7) hold.

Lemma 13.199.2 (applied to $T = S$) yields

$$AB = \left(\sum_{k\in S} a_{s,k} b_{k,t} \right)_{(s,t)\in S\times S}.$$

[1261] *Proof.* Let $(s,t) \in S \times S$ be such that we do not have $t \leq s$. We must prove that $\delta_{s,t} = 0$.

If we had $s = t$, then we would have $t = s \leq s$, which would contradict the fact that we do not have $t \leq s$. Hence, we cannot have $s = t$. Thus, $\delta_{s,t} = 0$, qed.

But every $(s,t) \in S \times S$ and every $k \in S$ which do not satisfy $t \leq s$ must satisfy

$$a_{s,k} b_{k,t} = 0 \tag{13.199.8}$$

[1262]. Thus, every $(s,t) \in S \times S$ which does not satisfy $t \leq s$ must satisfy

$$\sum_{k \in S} \underbrace{a_{s,k} b_{k,t}}_{\substack{=0 \\ \text{(by (13.199.8))}}} = \sum_{k \in S} 0 = 0.$$

Hence, Lemma 13.199.5 (applied to AB and $\sum_{k \in S} a_{s,k} b_{k,t}$ instead of A and $a_{s,t}$) shows that $AB \in \mathrm{T}_S$ (since $AB = \left(\sum_{k \in S} a_{s,k} b_{k,t}\right)_{(s,t) \in S \times S}$). This proves Claim A5.

Recall that T_S is a subset of $\mathbf{k}^{S \times S}$. This subset T_S is a $\mathbf{k}$-submodule of $\mathbf{k}^{S \times S}$ (by Claim A1, Claim A2 and Claim A3), and therefore is a $\mathbf{k}$-subalgebra of $\mathbf{k}^{S \times S}$ (by Claim A4 and Claim A5). In other words, the set of all triangular $S \times S$-matrices is a $\mathbf{k}$-subalgebra of $\mathbf{k}^{S \times S}$ (since T_S is the set of all triangular $S \times S$-matrices). This proves Proposition 11.1.10(a).

(b) We shall first prove the following claims:

Claim B1: We have $I_S \in \mathrm{IT}_S$.

Claim B2: For every $A \in \mathrm{IT}_S$ and $B \in \mathrm{IT}_S$, we have $AB \in \mathrm{IT}_S$.

Claim B3: Let $A \in \mathrm{IT}_S$. Then, A is invertible (as an element of the ring $\mathbf{k}^{S \times S}$), and its inverse A^{-1} belongs to IT_S.

Proof of Claim B1: The definition of I_S yields $I_S = (\delta_{s,t})_{(s,t) \in S \times S}$. Claim A4 in our proof of Proposition 11.1.10(a) yields $I_S \in \mathrm{T}_S$. For every $s \in S$, we have $\delta_{s,s} = 1$. Hence, for every $s \in S$, the element $\delta_{s,s}$ of $\mathbf{k}$ is invertible. Thus, Lemma 13.199.9 (applied to I_S and $\delta_{s,t}$ instead of A and $a_{s,t}$) yields $I_S \in \mathrm{IT}_S$. This proves Claim B1.

Proof of Claim B2: Let $A \in \mathrm{IT}_S$ and $B \in \mathrm{IT}_S$.

We know that A is an invertibly triangular $S \times S$-matrix (by Lemma 13.199.6). Hence, $A \in \mathrm{T}_S$ (by Lemma 13.199.7). Hence, Lemma 13.199.3 shows that A is a triangular $S \times S$-matrix. Write the $S \times S$-matrix A in the form $A = (a_{s,t})_{(s,t) \in S \times S}$. Lemma 13.199.8 yields that

$$\text{for every } s \in S, \text{ the element } a_{s,s} \text{ of } \mathbf{k} \text{ is invertible.} \tag{13.199.9}$$

We know that B is an invertibly triangular $S \times S$-matrix (by Lemma 13.199.6, applied to B instead of A). Hence, $B \in \mathrm{T}_S$ (by Lemma 13.199.7, applied to B instead of A). Hence, Lemma 13.199.3 (applied to B instead of A) shows that B is a triangular $S \times S$-matrix. Write the $S \times S$-matrix B in the form $B = (b_{s,t})_{(s,t) \in S \times S}$. Lemma 13.199.8 (applied to B and $b_{s,t}$ instead of A and $a_{s,t}$) yields that

$$\text{for every } s \in S, \text{ the element } b_{s,s} \text{ of } \mathbf{k} \text{ is invertible.} \tag{13.199.10}$$

From $A \in \mathrm{T}_S$ and $B \in \mathrm{T}_S$, we obtain $AB \in \mathrm{T}_S$ (by Claim A5 in our proof of Proposition 11.1.10(a)).

Lemma 13.199.2 (applied to $T = S$) yields

$$AB = \left(\sum_{k \in S} a_{s,k} b_{k,t}\right)_{(s,t) \in S \times S}.$$

[1262] *Proof of (13.199.8):* Let $(s,t) \in S \times S$ and $k \in S$ be such that we do not have $t \leq s$. We must prove the equality (13.199.8).

We are in one of the following two cases:

Case 1: We have $t \leq k$.

Case 2: We do not have $t \leq k$.

Let us first consider Case 1. In this case, we have $t \leq k$. If we had $k \leq s$, then we would have $t \leq k \leq s$; but this would contradict the fact that we do not have $t \leq s$. Hence, we cannot have $k \leq s$. Thus, (13.199.6) (applied to (s,k) instead of (s,t)) shows that $a_{s,k} = 0$. Thus, $\underbrace{a_{s,k}}_{=0} b_{k,t} = 0$. Hence, (13.199.8) is proven in Case 1.

Let us now consider Case 2. In this case, we do not have $t \leq k$. Hence, (13.199.7) (applied to (k,t) instead of (s,t)) shows that $b_{k,t} = 0$. Thus, $a_{s,k} \underbrace{b_{k,t}}_{=0} = 0$. Hence, (13.199.8) is proven in Case 2.

Now, we have proven (13.199.8) in each of the two Cases 1 and 2. This shows that (13.199.8) always holds. Qed.

Lemma 13.199.11 shows that, for every $s \in S$, we have $\sum_{k\in S} a_{s,k}b_{k,s} = a_{s,s}b_{s,s}$. Hence, for every $s \in S$, the element $\sum_{k\in S} a_{s,k}b_{k,s}$ of $\mathbf{k}$ is invertible[1263]. Thus, Lemma 13.199.9 (applied to AB and $\sum_{k\in S} a_{s,k}b_{k,t}$ instead of A and $a_{s,t}$) shows that $AB \in \mathrm{IT}_S$ (since $AB = \left(\sum_{k\in S} a_{s,k}b_{k,t}\right)_{(s,t)\in S\times S}$ and $AB \in \mathrm{T}_S$). This proves Claim B2.

Proof of Claim B3: We know that A is an invertibly triangular $S\times S$-matrix (by Lemma 13.199.6). Hence, $A \in \mathrm{T}_S$ (by Lemma 13.199.7). Hence, Lemma 13.199.3 shows that A is a triangular $S \times S$-matrix. Write the $S \times S$-matrix A in the form $A = (a_{s,t})_{(s,t)\in S\times S}$. Lemma 13.199.4 shows that

$$\text{every } (s,t) \in S \times S \text{ which does not satisfy } t \leq s \text{ must satisfy } a_{s,t} = 0. \tag{13.199.11}$$

Lemma 13.199.8 yields that

$$\text{for every } s \in S, \text{ the element } a_{s,s} \text{ of } \mathbf{k} \text{ is invertible.} \tag{13.199.12}$$

We shall now define an element $b_{s,t}$ for each $(s,t) \in S\times S$. In fact, we will define these elements recursively, by strong induction on $|[t,s]|$ [1264]: Let $N \in \mathbb{N}$. Assume that

$$\text{an element } b_{s,t} \in \mathbf{k} \text{ is already defined for each } (s,t) \in S \times S \text{ satisfying } |[t,s]| < N. \tag{13.199.13}$$

We shall now define an element $b_{s,t} \in \mathbf{k}$ for each $(s,t) \in S \times S$ satisfying $|[t,s]| = N$.

Indeed, let $(s,t) \in S \times S$ be such that $|[t,s]| = N$. We must define an element $b_{s,t} \in \mathbf{k}$.

For every $u \in S$ satisfying $t < u \leq s$, the element $b_{s,u}$ is already defined[1265]. Thus, the sum $\sum\limits_{\substack{u\in S;\\ t<u\leq s}} b_{s,u}a_{u,t} \in \mathbf{k}$ is well-defined. Furthermore, the element $a_{t,t}$ of $\mathbf{k}$ is invertible (by (13.199.12), applied to t instead of s). Hence, the element $(a_{t,t})^{-1}$ of $\mathbf{k}$ is well-defined. Now, we set

$$b_{s,t} = (a_{t,t})^{-1}\left(\delta_{s,t} - \sum_{\substack{u\in S;\\ t<u\leq s}} b_{s,u}a_{u,t}\right) \tag{13.199.14}$$

(this makes sense since both $(a_{t,t})^{-1}$ and $\sum\limits_{\substack{u\in S;\\ t<u\leq s}} b_{s,u}a_{u,t}$ are well-defined). Thus, we have defined an element $b_{s,t} \in \mathbf{k}$. This completes the recursive definition of $b_{s,t}$.

We have now defined an element $b_{s,t} \in \mathbf{k}$ for each $(s,t) \in S \times S$. In other words, we have defined a family $(b_{s,t})_{(s,t)\in S\times S} \in \mathbf{k}^{S\times S}$. This family is clearly an $S \times S$-matrix. Denote this $S \times S$-matrix by B. Thus, $B = (b_{s,t})_{(s,t)\in S\times S} \in \mathbf{k}^{S\times S}$. We shall now show (in several steps) that $BA = I_S$ and that $B \in \mathrm{IT}_S$.

First, we notice that

$$\text{every } (s,t) \in S \times S \text{ which does not satisfy } t \leq s \text{ must satisfy } b_{s,t} = 0 \tag{13.199.15}$$

[1266]. Hence, Lemma 13.199.5 (applied to B and $b_{s,t}$ instead of A and $a_{s,t}$) yields $B \in \mathrm{T}_S$.

[1263] *Proof.* Let $s \in S$. Then, the element $a_{s,s}$ of $\mathbf{k}$ is invertible (by (13.199.9)). The element $b_{s,s}$ of $\mathbf{k}$ is invertible as well (by (13.199.10)). Thus, the product $a_{s,s}b_{s,s}$ of these two elements is also invertible (since the product of two invertible elements is always invertible). In other words, $\sum_{k\in S} a_{s,k}b_{k,s}$ is invertible (since $\sum_{k\in S} a_{s,k}b_{k,s} = a_{s,s}b_{s,s}$). Qed.

[1264] Here, we are using the fact that $|[t,s]|$ is a nonnegative integer for every $(s,t) \in S \times S$.

[1265] *Proof.* Let $u \in S$ be such that $t < u \leq s$. We have to show that the element $b_{s,u}$ is already defined.

From (13.199.5) (applied to t, u and s instead of s, t and u), we conclude that $[u,s]$ is a proper subset of $[t,s]$. Hence, $|[u,s]| < |[t,s]|$ (since $[t,s]$ is a finite set). Thus, $|[u,s]| < |[t,s]| = N$. Therefore, (13.199.13) (applied to (s,u) instead of (s,t)) shows that an element $b_{s,u}$ is already defined.

[1266] *Proof of (13.199.15):* Let $(s,t) \in S \times S$ be such that we do not have $t \leq s$. We must prove that $b_{s,t} = 0$.

If we had $s = t$, then we would have $t = s \leq s$, which would contradict the fact that we do not have $t \leq s$. Hence, we cannot have $s = t$. Thus, $\delta_{s,t} = 0$.

If there was an $u \in S$ satisfying $t < u \leq s$, then we would have $t \leq s$, which would contradict the fact that we do not have $t \leq s$. Hence, there is no $u \in S$ satisfying $t < u \leq s$. Thus, the sum $\sum\limits_{\substack{u\in S;\\ t<u\leq s}} b_{s,u}a_{u,t}$ is empty. Hence, $\sum\limits_{\substack{u\in S;\\ t<u\leq s}} b_{s,u}a_{u,t} = (\text{empty sum}) = 0$.

For every $(s,t) \in S \times S$, we have

$$\sum_{\substack{k\in S;\\ t\leq k\leq s}} b_{s,k} a_{k,t} = \delta_{s,t} \tag{13.199.16}$$

[1267]. Thus, for every $(s,t) \in S \times S$, we have

$$\sum_{k\in S} b_{s,k} a_{k,t} = \delta_{s,t} \tag{13.199.17}$$

Now, the recursive definition of $b_{s,t}$ yields $b_{s,t} = (a_{t,t})^{-1} \underbrace{\left(\underbrace{\delta_{s,t}}_{=0} - \sum_{\substack{u\in S;\\ t<u\leq s}} b_{s,u} a_{u,t} \right)}_{=0} = (a_{t,t})^{-1} 0 = 0$. This proves (13.199.15).

[1267] *Proof of (13.199.16):* Let $(s,t) \in S \times S$. We must prove the equality (13.199.16).

We are in one of the following two cases:

Case 1: We have $t \leq s$.

Case 2: We do not have $t \leq s$.

Let us first consider Case 1. In this case, we have $t \leq s$. Thus, t is an element of S satisfying $t \leq t \leq s$. Hence, the sum $\sum_{\substack{u\in S;\\ t\leq u\leq s}} b_{s,u} a_{u,t}$ has an addend for $u = t$. Splitting off this addend, we obtain

$$\sum_{\substack{u\in S;\\ t\leq u\leq s}} b_{s,u} a_{u,t} = \sum_{\substack{u\in S;\\ t\leq u\leq s \text{ and } u\neq t}} b_{s,u} a_{u,t} + b_{s,t} a_{t,t}.$$

Now,

$$\begin{aligned}
\sum_{\substack{k\in S;\\ t\leq k\leq s}} b_{s,k} a_{k,t} &= \sum_{\substack{u\in S;\\ t\leq u\leq s}} b_{s,u} a_{u,t} \qquad \text{(here, we have renamed the summation index } k \text{ as } u\text{)} \\
&= \underbrace{\sum_{\substack{u\in S;\\ t\leq u\leq s \text{ and } u\neq t}}}_{\substack{= \sum_{\substack{u\in S;\\ t<u\leq s}} \\ \text{(because for any } u\in S\text{, the} \\ \text{condition } (t\leq u\leq s \text{ and } u\neq t) \\ \text{is equivalent to } (t<u\leq s))}} b_{s,u} a_{u,t} + \underbrace{b_{s,t}}_{\substack{=(a_{t,t})^{-1}\left(\delta_{s,t} - \sum_{\substack{u\in S;\\ t<u\leq s}} b_{s,u}a_{u,t}\right) \\ \text{(by the recursive definition of } b_{s,t}\text{)}}} a_{t,t} \\
&= \sum_{\substack{u\in S;\\ t<u\leq s}} b_{s,u} a_{u,t} + \underbrace{(a_{t,t})^{-1} \left(\delta_{s,t} - \sum_{\substack{u\in S;\\ t<u\leq s}} b_{s,u} a_{u,t} \right) a_{t,t}}_{=\delta_{s,t} - \sum_{\substack{u\in S;\\ t<u\leq s}} b_{s,u}a_{u,t}} \\
&= \sum_{\substack{u\in S;\\ t<u\leq s}} b_{s,u} a_{u,t} + \delta_{s,t} - \sum_{\substack{u\in S;\\ t<u\leq s}} b_{s,u} a_{u,t} = \delta_{s,t}.
\end{aligned}$$

Hence, (13.199.16) is proven in Case 1.

Let us now consider Case 2. In this case, we do not have $t \leq s$.

If we had $s = t$, then we would have $t = s \leq s$, which would contradict the fact that we do not have $t \leq s$. Hence, we do not have $s = t$. Therefore, $\delta_{s,t} = 0$.

If there was an $k \in S$ satisfying $t \leq k \leq s$, then we would have $t \leq s$, which would contradict the fact that we do not have $t \leq s$. Hence, there is no $k \in S$ satisfying $t \leq k \leq s$. Thus, the sum $\sum_{\substack{k\in S;\\ t\leq k\leq s}} b_{s,k} a_{k,t}$ is empty. Hence, $\sum_{\substack{k\in S;\\ t\leq k\leq s}} b_{s,k} a_{k,t} = (\text{empty sum}) = 0 = \delta_{s,t}$ (since $\delta_{s,t} = 0$). Hence, (13.199.16) is proven in Case 2.

We have now proven (13.199.16) in each of the two Cases 1 and 2. Thus, (13.199.16) always holds.

[1268]. Hence, $BA = I_S$ [1269]. Furthermore, for every $s \in S$, we have

$$b_{s,s} = (a_{s,s})^{-1} \tag{13.199.18}$$

[1270]. Hence,

$$\text{for every } s \in S, \text{ the element } b_{s,s} \text{ of } \mathbf{k} \text{ is invertible}$$

(because $b_{s,s}$ is an inverse (namely, $b_{s,s} = (a_{s,s})^{-1}$)). Thus, Lemma 13.199.9 (applied to B and $b_{s,t}$ instead of A and $a_{s,t}$) shows that $B \in \mathrm{IT}_S$ (since $B \in \mathrm{T}_S$). In other words, B belongs to IT_S.

So far, we do not know that B is the inverse of A; we merely know that B is a left inverse of A (that is, we know that $BA = I_S$). We shall now construct yet another $S \times S$-matrix C and subsequently show that $AC = I_S$; this will easily yield that A is invertible (because an element of a ring that has a left inverse and a right inverse must be invertible) and its inverse is $C = B$.

The construction of C will be rather similar to that of B, so that we will be briefer than before.

We shall define an element $c_{s,t}$ for each $(s,t) \in S \times S$. In fact, we will define these elements recursively, by strong induction on $|[t,s]|$: Let $N \in \mathbb{N}$. Assume that

$$\text{an element } c_{s,t} \in \mathbf{k} \text{ is already defined for each } (s,t) \in S \times S \text{ satisfying } |[t,s]| < N. \tag{13.199.20}$$

We shall now define an element $c_{s,t} \in \mathbf{k}$ for each $(s,t) \in S \times S$ satisfying $|[t,s]| = N$.

Indeed, let $(s,t) \in S \times S$ be such that $|[t,s]| = N$. We must define an element $c_{s,t} \in \mathbf{k}$.

For every $u \in S$ satisfying $t \le u < s$, the element $c_{u,t}$ is already defined[1271]. Thus, the sum $\sum\limits_{\substack{u\in S;\\ t\le u<s}} a_{s,u}c_{u,t} \in \mathbf{k}$ is well-defined. Furthermore, the element $a_{s,s}$ of $\mathbf{k}$ is invertible (by (13.199.12)). Hence, the element $(a_{s,s})^{-1}$

[1268] *Proof of (13.199.17):* Let $(s,t) \in S \times S$. Then,

$$\begin{aligned}
\sum_{k\in S} b_{s,k}a_{k,t} &= \sum_{\substack{k\in S;\\ t\le k}} b_{s,k}a_{k,t} + \sum_{\substack{k\in S;\\ \text{not } t\le k}} b_{s,k} \underbrace{a_{k,t}}_{\substack{=0\\ \text{(by (13.199.11), applied to}\\ (k,t) \text{ instead of } (s,t))}} = \sum_{\substack{k\in S;\\ t\le k}} b_{s,k}a_{k,t} + \underbrace{\sum_{\substack{k\in S;\\ \text{not } t\le k}} b_{s,k}0}_{=0} \\
&= \sum_{\substack{k\in S;\\ t\le k}} b_{s,k}a_{k,t} = \underbrace{\sum_{\substack{k\in S;\\ t\le k \text{ and } k\le s}}}_{=\sum_{\substack{k\in S;\\ t\le k\le s}}} b_{s,k}a_{k,t} + \sum_{\substack{k\in S;\\ t\le k \text{ and not } k\le s}} \underbrace{b_{s,k}}_{\substack{=0\\ \text{(by (13.199.15), applied to}\\ (s,k) \text{ instead of } (s,t))}} a_{k,t} \\
&= \sum_{\substack{k\in S;\\ t\le k\le s}} b_{s,k}a_{k,t} + \underbrace{\sum_{\substack{k\in S;\\ t\le k \text{ and not } k\le s}} 0a_{k,t}}_{=0} = \sum_{\substack{k\in S;\\ t\le k\le s}} b_{s,k}a_{k,t} = \delta_{s,t}
\end{aligned}$$

(by (13.199.16)). This proves (13.199.17).

[1269] *Proof.* Lemma 13.199.2 (applied to S, B, $b_{s,t}$, A and $a_{s,t}$ instead of T, A, $a_{s,t}$, B and $b_{s,t}$) shows that

$$BA = \left(\underbrace{\sum_{k\in S} b_{s,k}a_{k,t}}_{\substack{=\delta_{s,t}\\ \text{(by (13.199.17))}}} \right)_{(s,t)\in S\times S} = (\delta_{s,t})_{(s,t)\in S\times S} = I_S$$

(since $I_S = (\delta_{s,t})_{(s,t)\in S\times S}$ (by the definition of I_S)). Qed.

[1270] *Proof of (13.199.18):* Lemma 13.199.11 (applied to B, $b_{s,t}$, A and $a_{s,t}$ instead of A, $a_{s,t}$, B and $b_{s,t}$) yields that for every $s \in S$, we have

$$\sum_{k\in S} b_{s,k}a_{k,s} = b_{s,s}a_{s,s}. \tag{13.199.19}$$

Let $s \in S$. Then, (13.199.17) (applied to (s,s) instead of (s,t)) yields $\sum_{k\in S} b_{s,k}a_{k,s} = \delta_{s,s} = 1$ (since $s = s$). Comparing this with (13.199.19), we obtain $b_{s,s}a_{s,s} = 1$. Hence, $b_{s,s} = (a_{s,s})^{-1}$. This proves (13.199.18).

[1271] *Proof.* Let $u \in S$ be such that $t \le u < s$. We have to show that the element $c_{u,t}$ is already defined.

From (13.199.4) (applied to t, u and s instead of s, t and u), we conclude that $[t,u]$ is a proper subset of $[t,s]$. Hence, $|[t,u]| < |[t,s]|$ (since $[t,s]$ is a finite set). Thus, $|[t,u]| < |[t,s]| = N$. Therefore, (13.199.20) (applied to (u,t) instead of (s,t)) shows that an element $c_{u,t}$ is already defined.

of $\mathbf{k}$ is well-defined. Now, we set

$$c_{s,t} = (a_{s,s})^{-1} \left(\delta_{s,t} - \sum_{\substack{u \in S; \\ t \leq u < s}} a_{s,u} c_{u,t} \right) \tag{13.199.21}$$

(this makes sense since both $(a_{s,s})^{-1}$ and $\sum_{\substack{u \in S; \\ t \leq u < s}} a_{s,u} c_{u,t}$ are well-defined). Thus, we have defined an element $c_{s,t} \in \mathbf{k}$. This completes the recursive definition of $c_{s,t}$.

We have now defined an element $c_{s,t} \in \mathbf{k}$ for each $(s,t) \in S \times S$. In other words, we have defined a family $(c_{s,t})_{(s,t) \in S \times S} \in \mathbf{k}^{S \times S}$. This family is clearly an $S \times S$-matrix. Denote this $S \times S$-matrix by C. Thus, $C = (c_{s,t})_{(s,t) \in S \times S} \in \mathbf{k}^{S \times S}$. We shall now show (in several steps) that $AC = I_S$ and that $C \in \mathrm{IT}_S$.

First, we notice that

$$\text{every } (s,t) \in S \times S \text{ which does not satisfy } t \leq s \text{ must satisfy } c_{s,t} = 0 \tag{13.199.22}$$

[1272]. Hence, Lemma 13.199.5 (applied to C and $c_{s,t}$ instead of A and $a_{s,t}$) yields $C \in \mathrm{T}_S$.

For every $(s,t) \in S \times S$, we have

$$\sum_{\substack{k \in S; \\ t \leq k \leq s}} a_{s,k} c_{k,t} = \delta_{s,t} \tag{13.199.23}$$

[1273]. Thus, for every $(s,t) \in S \times S$, we have

$$\sum_{k \in S} a_{s,k} c_{k,t} = \delta_{s,t} \tag{13.199.24}$$

[1272] *Proof of (13.199.22):* The statement (13.199.22) is an analogue of (13.199.15), and has an analogous proof.

[1273] *Proof of (13.199.23):* The statement (13.199.23) is analogous to (13.199.16) and has an analogous proof. (The main difference is that we now need to split off the addend for $u = s$ from the sum $\sum_{\substack{u \in S; \\ t \leq u \leq s}} a_{s,u} c_{u,t}$, instead of splitting off the addend for $u = t$ from the sum $\sum_{\substack{u \in S; \\ t \leq u \leq s}} b_{s,u} a_{u,t}$.)

[1274]. Hence, $AC = I_S$ [1275]. Now, using the associativity of the $\mathbf{k}$-algebra $\mathbf{k}^{S\times S}$, we can make the following computation:

$$B = B \underbrace{I_S}_{=AC} = \underbrace{BA}_{=I_S} C = I_S C = C.$$

Hence, $AC = I_S$ rewrites as $AB = I_S$. Combining this with $BA = I_S$, we see that B is an inverse of A in the ring $\mathbf{k}^{S\times S}$. In particular, the element A of $\mathbf{k}^{S\times S}$ is invertible. Its inverse A^{-1} is B (as we have just proven), and therefore belongs to IT_S (since we know that B belongs to IT_S). This proves Claim B3.

Recall that IT_S is a subset of $\mathbf{k}^{S\times S}$. Claim B1 and Claim B2 (combined) show that this set IT_S is a submonoid of the multiplicative monoid of $\mathbf{k}^{S\times S}$. Claim B3 furthermore proves that this submonoid is a group. Thus, IT_S is a group with respect to multiplication. In other words, the set of all invertibly triangular $S \times S$-matrices is a group with respect to multiplication (since IT_S is the set of all invertibly triangular $S \times S$-matrices). This proves Proposition 11.1.10(b).

(c) We shall first prove the following claims:

Claim C1: We have $I_S \in \mathrm{UT}_S$.

Claim C2: For every $A \in \mathrm{UT}_S$ and $B \in \mathrm{UT}_S$, we have $AB \in \mathrm{UT}_S$.

Claim C3: Let $A \in \mathrm{UT}_S$. Then, A is invertible (as an element of the ring $\mathbf{k}^{S\times S}$), and its inverse A^{-1} belongs to UT_S.

Proof of Claim C1: The definition of I_S yields $I_S = (\delta_{s,t})_{(s,t)\in S\times S}$. Claim A4 in our proof of Proposition 11.1.10(a) yields $I_S \in \mathrm{T}_S$. For every $s \in S$, we have $\delta_{s,s} = 1$. Thus, Lemma 13.199.15 (applied to I_S and $\delta_{s,t}$ instead of A and $a_{s,t}$) yields $I_S \in \mathrm{UT}_S$. This proves Claim C1.

Proof of Claim C2: Let $A \in \mathrm{UT}_S$ and $B \in \mathrm{UT}_S$.

Write the $S \times S$-matrix A in the form $A = (a_{s,t})_{(s,t)\in S\times S}$. Write the $S \times S$-matrix B in the form $B = (b_{s,t})_{(s,t)\in S\times S}$.

We know that A is unitriangular $S \times S$-matrix (by Lemma 13.199.12). Hence, $A \in \mathrm{T}_S$ (by Lemma 13.199.13). Similarly, $B \in \mathrm{T}_S$.

Lemma 13.199.14 yields that

$$\text{for every } s \in S, \text{ we have } a_{s,s} = 1. \tag{13.199.25}$$

[1274] *Proof of (13.199.24):* Let $(s,t) \in S \times S$. Then,

$$\begin{aligned}\sum_{k\in S} a_{s,k}c_{k,t} &= \sum_{\substack{k\in S;\\ t\le k}} a_{s,k}c_{k,t} + \sum_{\substack{k\in S;\\ \text{not } t\le k}} a_{s,k} \underbrace{c_{k,t}}_{\substack{=0\\ \text{(by (13.199.22), applied to}\\ (k,t) \text{ instead of } (s,t))}} = \sum_{\substack{k\in S;\\ t\le k}} a_{s,k}c_{k,t} + \underbrace{\sum_{\substack{k\in S;\\ \text{not } t\le k}} a_{s,k}0}_{=0}\\
&= \sum_{\substack{k\in S;\\ t\le k}} a_{s,k}c_{k,t} = \underbrace{\sum_{\substack{k\in S;\\ t\le k \text{ and } k\le s}}}_{=\sum_{\substack{k\in S;\\ t\le k\le s}}} a_{s,k}c_{k,t} + \sum_{\substack{k\in S;\\ t\le k \text{ and not } k\le s}} \underbrace{a_{s,k}}_{\substack{=0\\ \text{(by (13.199.11), applied to}\\ (s,k) \text{ instead of } (s,t))}} c_{k,t}\\
&= \sum_{\substack{k\in S;\\ t\le k\le s}} a_{s,k}c_{k,t} + \underbrace{\sum_{\substack{k\in S;\\ t\le k \text{ and not } k\le s}} 0c_{k,t}}_{=0} = \sum_{\substack{k\in S;\\ t\le k\le s}} a_{s,k}c_{k,t} = \delta_{s,t}\end{aligned}$$

(by (13.199.23)). This proves (13.199.24).

[1275] *Proof.* Lemma 13.199.2 (applied to S, C and $c_{s,t}$ instead of T, B and $b_{s,t}$) shows that

$$AC = \left(\underbrace{\sum_{k\in S} a_{s,k}c_{k,t}}_{\substack{=\delta_{s,t}\\ \text{(by (13.199.24))}}} \right)_{(s,t)\in S\times S} = (\delta_{s,t})_{(s,t)\in S\times S} = I_S$$

(since $I_S = (\delta_{s,t})_{(s,t)\in S\times S}$ (by the definition of I_S)). Qed.

The same argument (applied to B and $b_{s,t}$ instead of A and $a_{s,t}$) yields that

$$\text{for every } s \in S\text{, we have } b_{s,s} = 1. \tag{13.199.26}$$

Claim A5 in our proof of Proposition 11.1.10(a) yields $AB \in \mathrm{T}_S$. Lemma 13.199.2 (applied to $T = S$) yields

$$AB = \left(\sum_{k \in S} a_{s,k} b_{k,t}\right)_{(s,t) \in S \times S}.$$

Lemma 13.199.11 yields that, for every $s \in S$, we have

$$\sum_{k \in S} a_{s,k} b_{k,s} = \underbrace{a_{s,s}}_{\substack{=1 \\ \text{(by (13.199.25))}}} \underbrace{b_{s,s}}_{\substack{=1 \\ \text{(by (13.199.26))}}} = 1.$$

Hence, Lemma 13.199.15 (applied to AB and $\sum_{k \in S} a_{s,k} b_{k,t}$ instead of A and $a_{s,t}$) shows that $AB \in \mathrm{UT}_S$ (since $AB = \left(\sum_{k \in S} a_{s,k} b_{k,t}\right)_{(s,t) \in S \times S}$ and $AB \in \mathrm{T}_S$). This proves Claim C2.

Proof of Claim C3: We know that A is unitriangular $S \times S$-matrix (by Lemma 13.199.12). Hence, $A \in \mathrm{T}_S$ (by Lemma 13.199.13).

Write the $S \times S$-matrix A in the form $A = (a_{s,t})_{(s,t) \in S \times S}$. Lemma 13.199.14 yields that

$$\text{for every } s \in S\text{, we have } a_{s,s} = 1. \tag{13.199.27}$$

We have $A \in \mathrm{UT}_S \subset \mathrm{IT}_S$ (by Lemma 13.199.17). Thus, the conditions of Claim B3 (in our proof of Proposition 11.1.10(b)) are satisfied. Define an $S \times S$-matrix $B = (b_{s,t})_{(s,t) \in S \times S} \in \mathbf{k}^{S \times S}$ as in the proof of Claim B3 (in our proof of Proposition 11.1.10(b)). Then, we have the following facts (which were shown in the proof of Claim B3):

- For every $s \in S$, we have
$$b_{s,s} = (a_{s,s})^{-1}. \tag{13.199.28}$$
- We have $B \in \mathrm{T}_S$.
- The matrix B is an inverse of A in the ring $\mathbf{k}^{S \times S}$.

Now, for every $s \in S$, we have

$$b_{s,s} = \left(\underbrace{a_{s,s}}_{\substack{=1 \\ \text{(by (13.199.27))}}} \right)^{-1} = 1^{-1} = 1.$$

Hence, Lemma 13.199.15 (applied to B and $b_{s,t}$ instead of A and $a_{s,t}$) shows that $B \in \mathrm{UT}_S$ (since $B = (b_{s,t})_{(s,t) \in S \times S}$ and $B \in \mathrm{T}_S$). In other words, B belongs to UT_S.

Now, recall that B is an inverse of A in the ring $\mathbf{k}^{S \times S}$. In particular, the element A of $\mathbf{k}^{S \times S}$ is invertible. Its inverse A^{-1} is B (as we have just proven), and therefore belongs to UT_S (since we know that B belongs to UT_S). This proves Claim C3.

Recall that UT_S is a subset of $\mathbf{k}^{S \times S}$. Claim C1 and Claim C2 (combined) show that this set UT_S is a submonoid of the multiplicative monoid of $\mathbf{k}^{S \times S}$. Claim C3 furthermore proves that this submonoid is a group. Thus, UT_S is a group with respect to multiplication. In other words, the set of all unitriangular $S \times S$-matrices is a group with respect to multiplication (since UT_S is the set of all unitriangular $S \times S$-matrices). This proves Proposition 11.1.10(c).

(d) Let A be an invertibly triangular $S \times S$-matrix. In other words, $A \in \mathrm{IT}_S$ (since IT_S is the set of all invertibly triangular $S \times S$-matrices). Hence, Claim B3 (in our proof of Proposition 11.1.10(b)) shows that A is invertible, and that its inverse A^{-1} belongs to IT_S. The matrix A^{-1} is therefore an invertibly triangular $S \times S$-matrix (by Lemma 13.199.6 (applied to A^{-1} instead of A)).

Now, forget that we fixed A. We thus have shown that if A is an invertibly triangular $S \times S$-matrix, then A is invertible, and its inverse A^{-1} is again invertibly triangular. This proves Proposition 11.1.10(d).

(e) Let A be a unitriangular $S \times S$-matrix. In other words, $A \in \mathrm{UT}_S$ (since UT_S is the set of all unitriangular $S \times S$-matrices). Hence, Claim C3 (in our proof of Proposition 11.1.10(c)) shows that A is invertible, and that its inverse A^{-1} belongs to UT_S. The matrix A^{-1} is therefore a unitriangular $S \times S$-matrix (by Lemma 13.199.12 (applied to A^{-1} instead of A)).

Now, forget that we fixed A. We thus have shown that if A is a unitriangular $S \times S$-matrix, then A is invertible, and its inverse A^{-1} is again unitriangular. This proves Proposition 11.1.10(e). □

Thus, Exercise 11.1.11 is solved.

13.200. **Solution to Exercise 11.1.15.** *Solution to Exercise 11.1.15.* Before we prove Theorem 11.1.14, let us introduce a notation:

Definition 13.200.1. Let M be a $\mathbf{k}$-module. Let $(h_p)_{p \in P}$ be a family of elements of M. Then, we let $\langle h_p \mid p \in P \rangle$ denote the $\mathbf{k}$-submodule of M spanned by this family $(h_p)_{p \in P}$.

Proof of Theorem 11.1.14. Write the matrix A in the form $A = (a_{s,t})_{(s,t) \in S \times T}$.

We have assumed that the family $(e_s)_{s \in S}$ expands in the family $(f_t)_{t \in T}$ through the matrix A. In other words,

$$\text{every } s \in S \text{ satisfies } e_s = \sum_{t \in T} a_{s,t} f_t \tag{13.200.1}$$

(by the definition of "the family $(e_s)_{s \in S}$ expands in the family $(f_t)_{t \in T}$ through the matrix A").

Let B be the $T \times S$-matrix A^{-1}. Write the $T \times S$-matrix B as $B = (b_{s,t})_{(s,t) \in T \times S}$. From $B = A^{-1}$, we obtain $BA = I_T$ and $AB = I_S$. These two equalities show that the matrix B has an inverse (namely, A); thus, the matrix B is invertible. In other words, the matrix A^{-1} is invertible (since $B = A^{-1}$).

Lemma 13.199.2 (applied to T, S, B, $b_{s,t}$, A and $a_{s,t}$ instead of S, T, A, $a_{s,t}$, B and $b_{s,t}$) yields

$$BA = \left(\sum_{k \in S} b_{s,k} a_{k,t} \right)_{(s,t) \in T \times T}.$$

Hence,

$$\left(\sum_{k \in S} b_{s,k} a_{k,t} \right)_{(s,t) \in T \times T} = BA = I_T = (\delta_{s,t})_{(s,t) \in T \times T}$$

(by the definition of I_T). In other words,

$$\sum_{k \in S} b_{s,k} a_{k,t} = \delta_{s,t} \qquad \text{for every } (s,t) \in T \times T. \tag{13.200.2}$$

Now,

$$\text{every } u \in T \text{ satisfies } f_u = \sum_{k \in S} b_{u,k} e_k \tag{13.200.3}$$

[1276]. Renaming the indices u and k as s and t in this statement, we obtain the following:

$$\text{every } s \in T \text{ satisfies } f_s = \sum_{t\in S} b_{s,t} e_t. \tag{13.200.4}$$

We have $(f_t)_{t\in T} = (f_s)_{s\in T}$ (here, we renamed the index t as s) and $(e_s)_{s\in S} = (e_t)_{t\in S}$ (here, we renamed the index s as t).

Recall that $B = (b_{s,t})_{(s,t)\in T\times S}$. Hence, the family $(f_s)_{s\in T}$ expands in the family $(e_t)_{t\in S}$ through the matrix B if and only if

$$\text{every } s \in T \text{ satisfies } f_s = \sum_{t\in S} b_{s,t} e_t$$

(by the definition of "the family $(f_s)_{s\in T}$ expands in the family $(e_t)_{t\in S}$ through the matrix B"). Thus, we conclude that the family $(f_s)_{s\in T}$ expands in the family $(e_t)_{t\in S}$ through the matrix B (since every $s \in T$ satisfies $f_s = \sum_{t\in S} b_{s,t} e_t$). In other words, the family $(f_t)_{t\in T}$ expands in the family $(e_s)_{s\in S}$ through the matrix A^{-1} (since $(f_t)_{t\in T} = (f_s)_{s\in T}$ and $(e_s)_{s\in S} = (e_t)_{t\in S}$ and $A^{-1} = B$). This proves Theorem 11.1.14(a).

(b) We shall prove the following two claims:

Claim B1: We have $\langle e_s \mid s \in S\rangle \subset \langle f_t \mid t \in T\rangle$.

Claim B2: We have $\langle f_t \mid t \in T\rangle \subset \langle e_s \mid s \in S\rangle$.

Proof of Claim B1: We need to prove that $\langle e_s \mid s \in S\rangle \subset \langle f_t \mid t \in T\rangle$. Since $\langle f_t \mid t \in T\rangle$ is a $\mathbf{k}$-module, we only need to show that $e_s \in \langle f_t \mid t \in T\rangle$ for each $s \in S$. But the latter fact follows from (13.200.1). Thus, Claim B1 is proven.

Proof of Claim B2: Theorem 11.1.14(a) shows that the family $(f_t)_{t\in T}$ expands in the family $(e_s)_{s\in S}$ through the matrix A^{-1}. Moreover, we know that the matrix A^{-1} is invertible. Hence, we can apply Claim B1 to T, S, $(f_t)_{t\in T}$, $(e_s)_{s\in S}$ and A^{-1} instead of S, T, $(e_s)_{s\in S}$, $(f_t)_{t\in T}$ and A. As a result, we conclude that $\langle f_t \mid t \in T\rangle \subset \langle e_s \mid s \in S\rangle$. This proves Claim B2.

Combining Claim B1 with Claim B2, we obtain $\langle e_s \mid s \in S\rangle = \langle f_t \mid t \in T\rangle$. Now,

$$\begin{aligned}
&\left(\text{the } \mathbf{k}\text{-submodule of } M \text{ spanned by the family } (e_s)_{s\in S}\right)\\
&= \langle e_s \mid s \in S\rangle = \langle f_t \mid t \in T\rangle\\
&= \left(\text{the } \mathbf{k}\text{-submodule of } M \text{ spanned by the family } (f_t)_{t\in T}\right). \qquad (13.200.5)
\end{aligned}$$

This proves Theorem 11.1.14(b).

[1276] *Proof of (13.200.3):* Let $u \in T$. Then,

$$\begin{aligned}
\sum_{k\in S} b_{u,k} \underbrace{e_k}_{\substack{=\sum_{t\in T} a_{k,t} f_t \\ \text{(by (13.200.1), applied to } s=k)}}
&= \sum_{k\in S} b_{u,k} \sum_{t\in T} a_{k,t} f_t = \underbrace{\sum_{k\in S}\sum_{t\in T}}_{=\sum_{t\in T}\sum_{k\in S}} b_{u,k} a_{k,t} f_t\\
&= \sum_{t\in T}\sum_{k\in S} b_{u,k} a_{k,t} f_t = \sum_{t\in T} \underbrace{\left(\sum_{k\in S} b_{u,k} a_{k,t}\right)}_{\substack{=\delta_{u,t} \\ \text{(by (13.200.2), applied to} \\ (u,t) \text{ instead of } (s,t))}} f_t\\
&= \sum_{t\in T} \delta_{u,t} f_t = \underbrace{\delta_{u,u}}_{=1} f_u + \sum_{\substack{t\in T;\\ t\neq u}} \underbrace{\delta_{u,t}}_{\substack{=0\\ \text{(since } t\neq u)}} f_t\\
&\qquad \text{(here, we have split off the addend for } t = u \text{ from the sum)}\\
&= f_u + \underbrace{\sum_{\substack{t\in T;\\ t\neq u}} 0 f_t}_{=0} = f_u.
\end{aligned}$$

This proves (13.200.3).

(c) We have the following chain of logical equivalences:

$$\begin{aligned}
&\left(\text{the family } \left(e_s\right)_{s\in S} \text{ spans the } \mathbf{k}\text{-module } M\right)\\
\iff &\left(\underbrace{\left(\text{the } \mathbf{k}\text{-submodule of } M \text{ spanned by the family } \left(e_s\right)_{s\in S}\right)}_{\substack{=\left(\text{the } \mathbf{k}\text{-submodule of } M \text{ spanned by the family } \left(f_t\right)_{t\in T}\right)\\ \text{(by (13.200.5))}}} = M\right)\\
\iff &\left(\left(\text{the } \mathbf{k}\text{-submodule of } M \text{ spanned by the family } \left(f_t\right)_{t\in T}\right) = M\right)\\
\iff &\left(\text{the family } \left(f_t\right)_{t\in T} \text{ spans the } \mathbf{k}\text{-module } M\right).
\end{aligned}$$

In other words, the family $\left(e_s\right)_{s\in S}$ spans the **k**-module M if and only if the family $\left(f_t\right)_{t\in T}$ spans the **k**-module M. This proves Theorem 11.1.14(c).

(d) We shall show the following two claims:

> *Claim D1:* If the family $\left(f_t\right)_{t\in T}$ is **k**-linearly independent, then the family $\left(e_s\right)_{s\in S}$ is **k**-linearly independent.
>
> *Claim D2:* If the family $\left(e_s\right)_{s\in S}$ is **k**-linearly independent, then the family $\left(f_t\right)_{t\in T}$ is **k**-linearly independent.

Proof of Claim D1: Assume that the family $\left(f_t\right)_{t\in T}$ is **k**-linearly independent. In other words, if $\left(\mu_t\right)_{t\in T} \in \mathbf{k}^T$ is any family of elements of **k** satisfying $\sum_{t\in T} \mu_t f_t = 0$, then

$$\left(\mu_t\right)_{t\in T} = \left(0\right)_{t\in T}. \tag{13.200.6}$$

Let $\left(\lambda_s\right)_{s\in S} \in \mathbf{k}^S$ be a family of elements of S such that $\sum_{s\in S} \lambda_s e_s = 0$. We shall show that $\left(\lambda_s\right)_{s\in S} = \left(0\right)_{s\in S}$.

We have $\sum_{s\in S} \lambda_s e_s = 0$. Comparing this with

$$\begin{aligned}
\sum_{s\in S} \lambda_s \underbrace{e_s}_{\substack{=\sum_{t\in T} a_{s,t} f_t\\ \text{(by (13.200.1))}}} &= \sum_{s\in S} \lambda_s \left(\sum_{t\in T} a_{s,t} f_t\right) = \underbrace{\sum_{s\in S}\sum_{t\in T}}_{=\sum_{t\in T}\sum_{s\in S}} \lambda_s a_{s,t} f_t\\
&= \sum_{t\in T}\sum_{s\in S} \lambda_s a_{s,t} f_t = \sum_{t\in T} \left(\sum_{s\in S} \lambda_s a_{s,t}\right) f_t,
\end{aligned}$$

we obtain $\sum_{t\in T} \left(\sum_{s\in S} \lambda_s a_{s,t}\right) f_t = 0$. Hence, (13.200.6) (applied to $\mu_t = \sum_{s\in S} \lambda_s a_{s,t}$) yields $\left(\sum_{s\in S} \lambda_s a_{s,t}\right)_{t\in T} = \left(0\right)_{t\in T}$. In other words,

$$\sum_{s\in S} \lambda_s a_{s,t} = 0 \qquad \text{for every } t \in T. \tag{13.200.7}$$

But Lemma 13.199.2 yields

$$AB = \left(\sum_{k\in T} a_{s,k} b_{k,t}\right)_{(s,t)\in S\times S}.$$

Hence,

$$\left(\sum_{k\in T} a_{s,k} b_{k,t}\right)_{(s,t)\in S\times S} = AB = I_S = \left(\delta_{s,t}\right)_{(s,t)\in S\times S}$$

(by the definition of I_S). In other words,

$$\sum_{k\in T} a_{s,k} b_{k,t} = \delta_{s,t} \qquad \text{for every } (s,t) \in S \times S. \tag{13.200.8}$$

Now, fix $u \in S$. Then,

$$\begin{aligned}\sum_{s\in S}\sum_{k\in T}\lambda_s a_{s,k}b_{k,u} &= \sum_{s\in S}\lambda_s \underbrace{\sum_{k\in T} a_{s,k}b_{k,u}}_{\substack{=\delta_{s,u} \\ \text{(by (13.200.8)} \\ \text{(applied to } (s,u) \text{ instead of } (s,t)))}} = \sum_{s\in S}\lambda_s\delta_{s,u} \\ &= \lambda_u \underbrace{\delta_{u,u}}_{=1} + \sum_{\substack{s\in S;\\ s\neq u}} \lambda_s \underbrace{\delta_{s,u}}_{\substack{=0 \\ \text{(since } s\neq u)}} \\ &\qquad \text{(here, we have split off the addend for } s=u \text{ from the sum)} \\ &= \lambda_u + \underbrace{\sum_{\substack{s\in S;\\ s\neq u}} \lambda_s 0}_{=0} = \lambda_u.\end{aligned}$$

Hence,

$$\begin{aligned}\lambda_u &= \underbrace{\sum_{s\in S}\sum_{k\in T}}_{=\sum_{k\in T}\sum_{s\in S}} \lambda_s a_{s,k}b_{k,u} = \sum_{k\in T}\sum_{s\in S}\lambda_s a_{s,k}b_{k,u} \\ &= \sum_{k\in T} \underbrace{\left(\sum_{s\in S}\lambda_s a_{s,k}\right)}_{\substack{=0 \\ \text{(by (13.200.7) (applied to } t=k))}} b_{k,u} = \sum_{k\in T} 0 b_{k,u} = 0.\end{aligned}$$

Now, forget that we fixed u. We thus have proven that $\lambda_u = 0$ for every $u \in S$. Renaming the index u as s in this statement, we obtain the following: We have $\lambda_s = 0$ for every $s \in S$. In other words, $(\lambda_s)_{s\in S} = (0)_{s\in S}$.

Now, forget that we fixed $(\lambda_s)_{s\in S}$. We thus have shown that if $(\lambda_s)_{s\in S} \in \mathbf{k}^S$ is a family of elements of S such that $\sum_{s\in S}\lambda_s e_s = 0$, then $(\lambda_s)_{s\in S} = (0)_{s\in S}$. In other words, the family $(e_s)_{s\in S}$ is $\mathbf{k}$-linearly independent. This proves Claim D1.

Proof of Claim D2: Theorem 11.1.14(a) shows that the family $(f_t)_{t\in T}$ expands in the family $(e_s)_{s\in S}$ through the matrix A^{-1}. Moreover, we know that the matrix A^{-1} is invertible. Hence, we can apply Claim D1 to T, S, $(f_t)_{t\in T}$, $(e_s)_{s\in S}$ and A^{-1} instead of S, T, $(e_s)_{s\in S}$, $(f_t)_{t\in T}$ and A. As a result, we conclude that if the family $(e_s)_{s\in S}$ is $\mathbf{k}$-linearly independent, then the family $(f_t)_{t\in T}$ is $\mathbf{k}$-linearly independent. Claim D2 is thus proven.

Now, Claim D1 and Claim D2 are two mutually converse implications. Combining these two implications, we obtain an equivalence; this equivalence is precisely Theorem 11.1.14(d).

(e) The family $(e_s)_{s\in S}$ is a basis of the $\mathbf{k}$-module M if and only if it spans the $\mathbf{k}$-module M and is $\mathbf{k}$-linearly independent (by the definition of a basis). Thus, we have the following chain of logical equivalences:

$$\begin{aligned}&\left(\text{the family } (e_s)_{s\in S} \text{ is a basis of the } \mathbf{k}\text{-module } M\right) \\ \iff &\left(\text{the family } (e_s)_{s\in S} \text{ spans the } \mathbf{k}\text{-module } M \text{ and is } \mathbf{k}\text{-linearly independent}\right) \\ \iff &\underbrace{\left(\text{the family } (e_s)_{s\in S} \text{ spans the } \mathbf{k}\text{-module } M\right)}_{\substack{\iff \left(\text{the family } (f_t)_{t\in T} \text{ spans the } \mathbf{k}\text{-module } M\right) \\ \text{(by Theorem 11.1.14(c))}}} \\ &\quad \wedge \underbrace{\left(\text{the family } (e_s)_{s\in S} \text{ is } \mathbf{k}\text{-linearly independent}\right)}_{\substack{\iff \left(\text{the family } (f_t)_{t\in T} \text{ is } \mathbf{k}\text{-linearly independent}\right) \\ \text{(by Theorem 11.1.14(d))}}} \\ \iff &\left(\text{the family } (f_t)_{t\in T} \text{ spans the } \mathbf{k}\text{-module } M\right) \\ &\quad \wedge \left(\text{the family } (f_t)_{t\in T} \text{ is } \mathbf{k}\text{-linearly independent}\right). \end{aligned} \tag{13.200.9}$$

On the other hand, the family $(f_t)_{t\in T}$ is a basis of the $\mathbf{k}$-module M if and only if it spans the $\mathbf{k}$-module M and is $\mathbf{k}$-linearly independent (by the definition of a basis). Thus, we have the following chain of logical equivalences:

$$\begin{aligned}
&\left(\text{the family } (f_t)_{t\in T} \text{ is a basis of the } \mathbf{k}\text{-module } M\right)\\
\iff &\left(\text{the family } (f_t)_{t\in T} \text{ spans the } \mathbf{k}\text{-module } M \text{ and is } \mathbf{k}\text{-linearly independent}\right)\\
\iff &\left(\text{the family } (f_t)_{t\in T} \text{ spans the } \mathbf{k}\text{-module } M\right)\\
&\quad \wedge \left(\text{the family } (f_t)_{t\in T} \text{ is } \mathbf{k}\text{-linearly independent}\right)\\
\iff &\left(\text{the family } (e_s)_{s\in S} \text{ is a basis of the } \mathbf{k}\text{-module } M\right)\\
&\quad (\text{by } (13.200.9)).
\end{aligned}$$

This proves Theorem 11.1.14(e). □

Thus, Exercise 11.1.15 is solved.

13.201. **Solution to Exercise 11.1.20.** *Solution to Exercise 11.1.20.*

Proof of Remark 11.1.17. (a) We shall prove the following two claims:

Claim A1: If the family $(e_s)_{s\in S}$ expands triangularly in the family $(f_s)_{s\in S}$, then every $s\in S$ satisfies

$$e_s = (\text{a } \mathbf{k}\text{-linear combination of the elements } f_t \text{ for } t\in S \text{ satisfying } t\leq s).$$

Claim A2: If every $s\in S$ satisfies

$$e_s = (\text{a } \mathbf{k}\text{-linear combination of the elements } f_t \text{ for } t\in S \text{ satisfying } t\leq s),$$

then the family $(e_s)_{s\in S}$ expands triangularly in the family $(f_s)_{s\in S}$.

Proof of Claim A1: Assume that the family $(e_s)_{s\in S}$ expands triangularly in the family $(f_s)_{s\in S}$. In other words, there exists a triangular $S\times S$-matrix A such that the family $(e_s)_{s\in S}$ expands in the family $(f_s)_{s\in S}$ through the matrix A. Consider this A.

Write the matrix A in the form $A=(a_{s,t})_{(s,t)\in S\times S}$.

The family $(e_s)_{s\in S}$ expands in the family $(f_s)_{s\in S}$ through the matrix A. In other words,

$$\text{every } s\in S \text{ satisfies } e_s=\sum_{t\in S} a_{s,t} f_t. \tag{13.201.1}$$

Lemma 13.199.4 shows that

$$\text{every } (s,t)\in S\times S \text{ which does not satisfy } t\leq s \text{ must satisfy } a_{s,t}=0 \tag{13.201.2}$$

(since $A=(a_{s,t})_{(s,t)\in S\times S}$ is a triangular $S\times S$-matrix). Hence, every $s\in S$ satisfies

$$\begin{aligned}
e_s &= \sum_{t\in S} a_{s,t} f_t \qquad (\text{by } (13.201.1))\\
&= \sum_{\substack{t\in S;\\ t\leq s}} a_{s,t} f_t + \sum_{\substack{t\in S;\\ \text{not } t\leq s}} \underbrace{a_{s,t}}_{\substack{=0\\ (\text{by } (13.201.2))}} f_t\\
&= \sum_{\substack{t\in S;\\ t\leq s}} a_{s,t} f_t + \underbrace{\sum_{\substack{t\in S;\\ \text{not } t\leq s}} 0 f_t}_{=0} = \sum_{\substack{t\in S;\\ t\leq s}} a_{s,t} f_t\\
&= (\text{a } \mathbf{k}\text{-linear combination of the elements } f_t \text{ for } t\in S \text{ satisfying } t\leq s).
\end{aligned}$$

This proves Claim A1.

Proof of Claim A2: Assume that every $s\in S$ satisfies

$$e_s = (\text{a } \mathbf{k}\text{-linear combination of the elements } f_t \text{ for } t\in S \text{ satisfying } t\leq s). \tag{13.201.3}$$

In other words, for every $s \in S$, there exists a family $(c_{s,t})_{t\in S;\ t\leq s} \in \mathbf{k}^{\{t\in S \mid t\leq s\}}$ such that

$$e_s = \sum_{\substack{t\in S;\\ t\leq s}} c_{s,t} f_t. \tag{13.201.4}$$

Fix such a family for each $s \in S$.

Now, define an $S \times S$-matrix $B = (b_{s,t})_{(s,t)\in S\times S} \in \mathbf{k}^{S\times S}$ by

$$\left(b_{s,t} = \begin{cases} c_{s,t}, & \text{if } t \leq s; \\ 0, & \text{otherwise} \end{cases} \qquad \text{for every } (s,t) \in S \times S \right).$$

Every $(s,t) \in S \times S$ which does not satisfy $t \leq s$ must satisfy

$$b_{s,t} = 0 \qquad \text{(by the definition of } b_{s,t}\text{)}. \tag{13.201.5}$$

In other words, the $S \times S$-matrix B is triangular (by the definition of "triangular").

On the other hand, every $(s,t) \in S \times S$ which satisfies $t \leq s$ must satisfy

$$b_{s,t} = c_{s,t} \qquad \text{(by the definition of } b_{s,t}\text{)}. \tag{13.201.6}$$

Now,

$$\text{every } s \in S \text{ satisfies } e_s = \sum_{t\in S} b_{s,t} f_t \tag{13.201.7}$$

[1277]. In other words, the family $(e_s)_{s\in S}$ expands in the family $(f_s)_{s\in S}$ through the matrix B (since $B = (b_{s,t})_{(s,t)\in S\times S}$). Hence, the family $(e_s)_{s\in S}$ expands triangularly in the family $(f_s)_{s\in S}$ (since the matrix B is triangular). This proves Claim A2.

We have now proven Claim A1 and Claim A2. These two claims are mutually converse implications. Combining these two implications, we obtain an equivalence, which is precisely the statement of Remark 11.1.17(a). Thus, Remark 11.1.17(a) is proven.

(b) We shall prove the following two claims:

Claim B1: If the family $(e_s)_{s\in S}$ expands invertibly triangularly in the family $(f_s)_{s\in S}$, then every $s \in S$ satisfies

$$e_s = \alpha_s f_s + (\text{a } \mathbf{k}\text{-linear combination of the elements } f_t \text{ for } t \in S \text{ satisfying } t < s)$$

for some invertible $\alpha_s \in \mathbf{k}$.

Claim B2: If every $s \in S$ satisfies

$$e_s = \alpha_s f_s + (\text{a } \mathbf{k}\text{-linear combination of the elements } f_t \text{ for } t \in S \text{ satisfying } t < s)$$

for some invertible $\alpha_s \in \mathbf{k}$, then the family $(e_s)_{s\in S}$ expands invertibly triangularly in the family $(f_s)_{s\in S}$.

Proof of Claim B1: Assume that the family $(e_s)_{s\in S}$ expands invertibly triangularly in the family $(f_s)_{s\in S}$. In other words, there exists an invertibly triangular $S \times S$-matrix A such that the family $(e_s)_{s\in S}$ expands in the family $(f_s)_{s\in S}$ through the matrix A. Consider this A.

Write the matrix A in the form $A = (a_{s,t})_{(s,t)\in S\times S}$.

The family $(e_s)_{s\in S}$ expands in the family $(f_s)_{s\in S}$ through the matrix A. In other words,

$$\text{every } s \in S \text{ satisfies } e_s = \sum_{t\in S} a_{s,t} f_t. \tag{13.201.8}$$

[1277] *Proof of (13.201.7):* Let $s \in S$. Then, every $t \in S$ satisfies either $t \leq s$ or (not $t \leq s$). Hence,

$$\sum_{t\in S} b_{s,t} f_t = \sum_{\substack{t\in S;\\ t\leq s}} \underbrace{b_{s,t}}_{\substack{=c_{s,t}\\ \text{(by (13.201.6))}}} f_t + \sum_{\substack{t\in S;\\ \text{not } t\leq s}} \underbrace{b_{s,t}}_{\substack{=0\\ \text{(by (13.201.5))}}} f_t = \sum_{\substack{t\in S;\\ t\leq s}} c_{s,t} f_t + \underbrace{\sum_{\substack{t\in S;\\ \text{not } t\leq s}} 0 f_t}_{=0}$$

$$= \sum_{\substack{t\in S;\\ t\leq s}} c_{s,t} f_t = e_s \qquad \text{(by (13.201.4))}.$$

This proves (13.201.7).

Lemma 13.199.7 shows that $A \in \mathrm{T}_S$ (since A is an invertibly triangular $S \times S$-matrix). Thus, Lemma 13.199.3 shows that A is a triangular $S \times S$-matrix. Hence, Lemma 13.199.4 shows that

$$\text{every } (s,t) \in S \times S \text{ which does not satisfy } t \leq s \text{ must satisfy } a_{s,t} = 0 \tag{13.201.9}$$

(since $A = (a_{s,t})_{(s,t) \in S \times S}$ is a triangular $S \times S$-matrix). Now, it is easy to show that every $s \in S$ satisfies

$$e_s = \alpha_s f_s + (\text{a } \mathbf{k}\text{-linear combination of the elements } f_t \text{ for } t \in S \text{ satisfying } t < s)$$

for some invertible $\alpha_s \in \mathbf{k}$ [1278]. Thus, Claim B1 is proven.

Proof of Claim B2: Assume that every $s \in S$ satisfies

$$e_s = \alpha_s f_s + (\text{a } \mathbf{k}\text{-linear combination of the elements } f_t \text{ for } t \in S \text{ satisfying } t < s) \tag{13.201.10}$$

for some invertible $\alpha_s \in \mathbf{k}$. Consider this α_s.

Now, for every $s \in S$, there exists a family $(c_{s,t})_{t \in S;\ t < s} \in \mathbf{k}^{\{t \in S \mid t < s\}}$ such that

$$e_s - \alpha_s f_s = \sum_{\substack{t \in S; \\ t < s}} c_{s,t} f_t \tag{13.201.11}$$

[1279]. Fix such a family for each $s \in S$.

Now, define an $S \times S$-matrix $B = (b_{s,t})_{(s,t) \in S \times S} \in \mathbf{k}^{S \times S}$ by

$$\left(b_{s,t} = \begin{cases} c_{s,t}, & \text{if } t < s; \\ \delta_{s,t} \alpha_s, & \text{otherwise} \end{cases} \qquad \text{for every } (s,t) \in S \times S \right).$$

Every $(s,t) \in S \times S$ which does not satisfy $t \leq s$ must satisfy

$$b_{s,t} = 0 \tag{13.201.12}$$

[1280]. Hence, Lemma 13.199.5 (applied to B and $b_{s,t}$ instead of A and $a_{s,t}$) shows that $B \in \mathrm{T}_S$.

[1278] *Proof.* Let $s \in S$. Then, Lemma 13.199.8 shows that the element $a_{s,s}$ of $\mathbf{k}$ is invertible.

But (13.201.8) yields

$$\begin{aligned} e_s &= \sum_{t \in S} a_{s,t} f_t = \sum_{\substack{t \in S; \\ t \leq s}} a_{s,t} f_t + \sum_{\substack{t \in S; \\ \text{not } t \leq s}} \underbrace{a_{s,t}}_{\substack{=0 \\ \text{(by (13.201.9))}}} f_t \\ &= \sum_{\substack{t \in S; \\ t \leq s}} a_{s,t} f_t + \underbrace{\sum_{\substack{t \in S; \\ \text{not } t \leq s}} 0 f_t}_{=0} = \sum_{\substack{t \in S; \\ t \leq s}} a_{s,t} f_t = a_{s,s} f_s + \underbrace{\sum_{\substack{t \in S; \\ t \leq s \text{ and } t \neq s}}}_{= \sum_{\substack{t \in S; \\ t < s}}} a_{s,t} f_t \\ &\qquad \text{(here, we have split off the addend for } t = s \text{ from the sum)} \\ &= a_{s,s} f_s + \underbrace{\sum_{\substack{t \in S; \\ t < s}} a_{s,t} f_t}_{=(\text{a } \mathbf{k}\text{-linear combination of the elements } f_t \text{ for } t \in S \text{ satisfying } t < s)} \\ &= a_{s,s} f_s + (\text{a } \mathbf{k}\text{-linear combination of the elements } f_t \text{ for } t \in S \text{ satisfying } t < s). \end{aligned}$$

Hence,

$$e_s = \alpha_s f_s + (\text{a } \mathbf{k}\text{-linear combination of the elements } f_t \text{ for } t \in S \text{ satisfying } t < s)$$

for some invertible $\alpha_s \in \mathbf{k}$ (namely, for $\alpha_s = a_{s,s}$). Qed.

[1279] *Proof.* Let $s \in S$. Subtracting $\alpha_s f_s$ from both sides of (13.201.10), we obtain

$$e_s - \alpha_s f_s = (\text{a } \mathbf{k}\text{-linear combination of the elements } f_t \text{ for } t \in S \text{ satisfying } t < s).$$

In other words, there exists a family $(c_{s,t})_{t \in S;\ t < s} \in \mathbf{k}^{\{t \in S \mid t < s\}}$ such that $e_s - \alpha_s f_s = \sum_{\substack{t \in S; \\ t < s}} c_{s,t} f_t$.

[1280] *Proof of (13.201.12):* Let $(s,t) \in S \times S$ be such that we do not have $t \leq s$.

We do not have $t \leq s$. Thus, we do not have $s = t$ (since $s = t$ would imply $t \leq s$). Hence, $\delta_{s,t} = 0$.

On the other hand, every $(s,t) \in S \times S$ which satisfies $t < s$ must satisfy

$$\begin{aligned} b_{s,t} &= \begin{cases} c_{s,t}, & \text{if } t < s; \\ \delta_{s,t}\alpha_s, & \text{otherwise} \end{cases} \qquad \text{(by the definition of } b_{s,t}\text{)} \\ &= c_{s,t} \qquad \text{(since } t < s\text{)}. \end{aligned} \tag{13.201.13}$$

Moreover, every $s \in S$ satisfies

$$\begin{aligned} b_{s,s} &= \begin{cases} c_{s,s}, & \text{if } s < s; \\ \delta_{s,s}\alpha_s, & \text{otherwise} \end{cases} \qquad \text{(by the definition of } b_{s,s}\text{)} \\ &= \underbrace{\delta_{s,s}}_{=1} \alpha_s \qquad \text{(since we do not have } s < s\text{)} \\ &= \alpha_s. \end{aligned} \tag{13.201.14}$$

Now, recall that, for every $s \in S$, the element α_s of $\mathbf{k}$ is invertible. In other words, for every $s \in S$, the element $b_{s,s}$ of $\mathbf{k}$ is invertible (since $b_{s,s} = \alpha_s$ for every $s \in S$). Hence, Lemma 13.199.9 (applied to B and $b_{s,t}$ instead of A and $a_{s,t}$) shows that $B \in \mathrm{IT}_S$. Thus, Lemma 13.199.6 (applied to B instead of A) shows that B is an invertibly triangular $S \times S$-matrix.

Now,

$$\text{every } s \in S \text{ satisfies } e_s = \sum_{t \in S} b_{s,t} f_t \tag{13.201.15}$$

[1281]. In other words, the family $(e_s)_{s \in S}$ expands in the family $(f_s)_{s \in S}$ through the matrix B (since $B = (b_{s,t})_{(s,t) \in S \times S}$). Hence, the family $(e_s)_{s \in S}$ expands invertibly triangularly in the family $(f_s)_{s \in S}$ (since the matrix B is invertibly triangular). This proves Claim B2.

We have now proven Claim B1 and Claim B2. These two claims are mutually converse implications. Combining these two implications, we obtain an equivalence, which is precisely the statement of Remark 11.1.17(b). Thus, Remark 11.1.17(b) is proven.

If we had $t < s$, then we would have $t \leq s$, which would contradict the fact that we do not have $t \leq s$. Hence, we do not have $t < s$. Now, the definition of $b_{s,t}$ yields

$$\begin{aligned} b_{s,t} &= \begin{cases} c_{s,t}, & \text{if } t < s; \\ \delta_{s,t}\alpha_s, & \text{otherwise} \end{cases} = \underbrace{\delta_{s,t}}_{=0} \alpha_s \qquad \text{(since we do not have } t < s\text{)} \\ &= 0. \end{aligned}$$

This proves (13.201.12).

[1281] *Proof of (13.201.15):* Let $s \in S$. Then,

$$\begin{aligned} \sum_{t \in S} b_{s,t} f_t &= \sum_{\substack{t \in S; \\ t \leq s}} b_{s,t} f_t + \sum_{\substack{t \in S; \\ \text{not } t \leq s}} \underbrace{b_{s,t}}_{\substack{=0 \\ \text{(by (13.201.9))}}} f_t = \sum_{\substack{t \in S; \\ t \leq s}} b_{s,t} f_t + \underbrace{\sum_{\substack{t \in S; \\ \text{not } t \leq s}} 0 f_t}_{=0} \\ &= \sum_{\substack{t \in S; \\ t \leq s}} b_{s,t} f_t = \underbrace{b_{s,s}}_{\substack{=\alpha_s \\ \text{(by (13.201.14))}}} f_s + \underbrace{\sum_{\substack{t \in S; \\ t \leq s \text{ and } t \neq s}}}_{= \sum_{\substack{t \in S; \\ t < s}}} b_{s,t} f_t \\ &\qquad \text{(here, we have split off the addend for } t = s \text{ from the sum)} \\ &= \alpha_s f_s + \sum_{\substack{t \in S; \\ t < s}} \underbrace{b_{s,t}}_{\substack{=c_{s,t} \\ \text{(by (13.201.13))}}} f_t \\ &= \alpha_s f_s + \underbrace{\sum_{\substack{t \in S; \\ t < s}} c_{s,t} f_t}_{\substack{=e_s - \alpha_s f_s \\ \text{(by (13.201.11))}}} = \alpha_s f_s + (e_s - \alpha_s f_s) = e_s. \end{aligned}$$

This proves (13.201.15).

(c) The proof of Remark 11.1.17(c) is analogous to our above proof of Remark 11.1.17(b); the main differences are that "invertibly triangular" has to be replaced by "unitriangular", and correspondingly conditions of the form "$a_{s,s}$ is invertible" are replaced by "$a_{s,s} = 1$", and finally every occurrence of "α_s" has to be replaced by "1". We leave the details to the reader. □

Proof of Corollary 11.1.19. The family $(e_s)_{s\in S}$ expands invertibly triangularly in the family $(f_s)_{s\in S}$. In other words, there exists an invertibly triangular $S \times S$-matrix A such that the family $(e_s)_{s\in S}$ expands in the family $(f_s)_{s\in S}$ through the matrix A (by the definition of what it means to "expand invertibly triangularly"). Denote this A by B. Thus, B is an invertibly triangular $S \times S$-matrix such that the family $(e_s)_{s\in S}$ expands in the family $(f_s)_{s\in S}$ through the matrix B.

Proposition 11.1.10(d) says that any invertibly triangular $S \times S$-matrix is invertible, and that its inverse is again invertibly triangular. In other words: If A is any invertibly triangular $S \times S$-matrix, then A is invertible, and its inverse A^{-1} is again invertibly triangular. We can apply this fact to $A = B$ (since B is an invertibly triangular $S \times S$-matrix). Thus, we conclude that B is invertible, and that its inverse B^{-1} is again invertibly triangular.

Since the matrix B is invertible, we can apply Theorem 11.1.14 to S, $(f_s)_{s\in S}$ and B instead of T, $(f_t)_{t\in T}$ and A. Hence, parts (b), (c), (d) and (e) of Corollary 11.1.19 follow immediately from parts (b), (c), (d) and (e) of Theorem 11.1.14 (applied to S, $(f_s)_{s\in S}$ and B instead of T, $(f_t)_{t\in T}$ and A). It remains to prove Corollary 11.1.19(a).

(a) Theorem 11.1.14(a) (applied to S, $(f_s)_{s\in S}$ and B instead of T, $(f_t)_{t\in T}$ and A) shows that the family $(f_s)_{s\in S}$ expands in the family $(e_s)_{s\in S}$ through the matrix B^{-1}. Hence, there exists an invertibly triangular $S \times S$-matrix A such that the family $(f_s)_{s\in S}$ expands in the family $(e_s)_{s\in S}$ through the matrix A (namely, $A = B^{-1}$). In other words, the family $(f_s)_{s\in S}$ expands invertibly triangularly in the family $(e_s)_{s\in S}$ (by the definition of "expands invertibly triangularly"). This proves Corollary 11.1.19(a). As we said, this completes the proof of Corollary 11.1.19. □

Thus, Exercise 11.1.20 is solved.

Email address: darijgrinberg@gmail.com

Drexel University, Korman Center, Room 263, 15 S 33rd Street, Philadelphia PA, 19104, USA // (temporary) Mathematisches Forschungsinstitut Oberwolfach, Schwarzwaldstrasse 9–11, 77709 Oberwolfach, Germany

Email address: reiner@math.umn.edu

School of Mathematics, University of Minnesota, Minneapolis, MN 55455, USA