Collected trivialities on algebra derivations

Darij Grinberg

October 22, 2020

Contents

***

The purpose of this note is to state and prove in detail some folklore properties of derivations on $\mathbf{k}$-algebras. It contains no deep result or complicated proofs; its length is chiefly due to the level of detail and slow pacing.

1. Derivations in general

1.1. Definitions and conventions

Let us first recall some definitions and set up some notations that will be used for the rest of this note.

Convention 1.1. In the following, $\mathbb{N}$ denotes the set $\{0, 1, 2, \ldots\}$.

The word "ring" shall always mean an associative ring with unity.

We fix a commutative ring $\mathbf{k}$ (once and for all). If M and N are two $\mathbf{k}$-modules, then we let $\operatorname{Hom}(M, N)$ denote the $\mathbf{k}$-module consisting of all $\mathbf{k}$-module homomorphisms $M \to N$. (We use this notation even if M and N are equipped with some other structure; e.g., even if M and N are $\mathbf{k}$-algebras, we write $\operatorname{Hom}(M, N)$ for the $\mathbf{k}$-module homomorphisms $M \to N$, not the $\mathbf{k}$-algebra homomorphisms $M \to N$.) All tensor products are understood to be tensor products over $\mathbf{k}$ unless said otherwise.

A *magmatic* $\mathbf{k}$*-algebra* means a $\mathbf{k}$-module A equipped with a $\mathbf{k}$-bilinear map $m : A \times A \to A$. This map m is called the *multiplication* of the magmatic $\mathbf{k}$-algebra A. The multiplication of a magmatic $\mathbf{k}$-algebra A is often "written as multiplication"; i.e., one writes $a \cdot b$ (or, even shorter, ab) for $m((a, b))$ whenever a and b are two elements of A. (Often, a magmatic $\mathbf{k}$-algebra is called a *nonassociative* $\mathbf{k}$*-algebra*. However, the word "nonassociative" is slightly confusing here, since it does not mean that associativity must be violated; it only means that associativity is not required.)

Notice that any magmatic $\mathbf{k}$-algebra A automatically satisfies

$$\lambda(ab) = (\lambda a) b = a(\lambda b) \qquad \text{for every } \lambda \in \mathbf{k} \text{ and every } a \in A \text{ and } b \in A \tag{1}$$

(because its multiplication is $\mathbf{k}$-bilinear). Some authors use a different concept of $\mathbf{k}$-algebras, which is more general and does not always satisfy (1). In their notations, what we call "magmatic $\mathbf{k}$-algebra" is called "central magmatic $\mathbf{k}$-algebra".

A *unital magmatic* $\mathbf{k}$*-algebra* is defined to be a magmatic $\mathbf{k}$-algebra A equipped with an element $e \in A$ which satisfies

$$(ea = ae = a \qquad \text{for every } a \in A).$$

This element e is unique when it exists (this is easy to prove), and is called the *unity* of A. (Some authors use the word "identity", "unit" or "one" instead of "unity", but these words are sometimes ambiguous.) Since the unity of A is unique, we can afford not specifying it when defining a unital magmatic $\mathbf{k}$-algebra (as long as we make sure that it exists); we thus can say that some magmatic $\mathbf{k}$-algebra A "is unital", when we really mean that there exists an $e \in A$ such that A equipped with this e is a unital magmatic $\mathbf{k}$-algebra.

The unity of a unital magmatic $\mathbf{k}$-algebra A is denoted by 1_A, or by 1 when no confusion can arise.

Notice that any unital magmatic $\mathbf{k}$-algebra A automatically satisfies

$$\lambda a = (\lambda \cdot 1_A)\, a = a\, (\lambda \cdot 1_A) \qquad \text{for every } \lambda \in \mathbf{k} \text{ and every } a \in A.$$

(Indeed, this follows easily from (1).)

A magmatic $\mathbf{k}$-algebra A is said to be *associative* if it satisfies

$$(a\,(bc) = (ab)\,c \qquad \text{for every } a \in A,\ b \in A \text{ and } c \in A)\,.$$

We use the notation "$\mathbf{k}$-*algebra*" for "associative unital magmatic $\mathbf{k}$-algebra". Thus, of course, a magmatic $\mathbf{k}$-algebra is not always a $\mathbf{k}$-algebra.

The base ring $\mathbf{k}$ itself becomes a $\mathbf{k}$-algebra (equipped with its multiplication and its unity).

If A and B are two $\mathbf{k}$-algebras, then a $\mathbf{k}$-*algebra homomorphism from A to B* means a $\mathbf{k}$-module homomorphism $f : A \to B$ satisfying $f(1_A) = 1_B$ and

$$(f(ab) = f(a)\,f(b) \qquad \text{for every } a \in A \text{ and } b \in A)\,. \tag{2}$$

We notice that the condition $f(1_A) = 1_B$ does not follow from (2).

Every $\mathbf{k}$-algebra is a ring (when equipped with its multiplication and its unity), and every $\mathbf{k}$-algebra homomorphism is a ring homomorphism. Actually, if A and B are two $\mathbf{k}$-algebras, then a $\mathbf{k}$-algebra homomorphism $A \to B$ is the same as a $\mathbf{k}$-linear ring homomorphism $A \to B$. If A is a $\mathbf{k}$-algebra, then the map

$$\mathbf{k} \to A, \qquad \lambda \mapsto \lambda \cdot 1_A$$

is a ring homomorphism and a $\mathbf{k}$-algebra homomorphism. Thus, any left A-module M canonically becomes a $\mathbf{k}$-module (via this homomorphism); explicitly, its $\mathbf{k}$-module structure is given by

$$(\lambda m = (\lambda \cdot 1_A)\,m \qquad \text{for every } \lambda \in \mathbf{k} \text{ and } m \in M)\,.$$

Similarly, any right A-module M canonically becomes a $\mathbf{k}$-module.

The terminology we have introduced in Convention 1.1 is fairly standard in some parts of the mathematical world, and nonstandard in others. For instance, we use the word "$\mathbf{k}$-algebra" as an abbreviation for "associative unital magmatic $\mathbf{k}$-algebra"; some others use it as an abbreviation for "associative magmatic $\mathbf{k}$-algebra", whereas others use it for "magmatic $\mathbf{k}$-algebra". Some authors also say "unitary" instead of "unital".

Definition 1.2. Let A and B be $\mathbf{k}$-algebras. An $(A, B)_{\mathbf{k}}$-*bimodule* means a $\mathbf{k}$-module M equipped with a left A-module structure and a right B-module structure satisfying the following properties:

1. We have
$$(am)\,b = a\,(mb) \qquad \text{for any } a \in A,\ m \in M \text{ and } b \in B. \tag{3}$$
2. We have
$$(\lambda 1_A)\,m = m\,(\lambda 1_B) = \lambda m \qquad \text{for any } \lambda \in \mathbf{k} \text{ and } m \in M. \tag{4}$$

(In other words, the $\mathbf{k}$-module structure on M obtained from the left A-module structure on M is identical to the $\mathbf{k}$-module structure on M obtained from the right B-module structure on M, and also identical to the $\mathbf{k}$-module structure that was given on M.)

We shall abbreviate "$(A, B)_{\mathbf{k}}$-bimodule" as "(A, B)-bimodule", since the $\mathbf{k}$ is fixed.

The condition (3) in the definition of an (A, B)-bimodule allows us to write amb (without bracketing) for both $(am)\,b$ and $a\,(mb)$ (where M is an (A, B)-bimodule, and where $a \in A$, $m \in M$ and $b \in B$) without having to worry about ambiguity.

Notice that the notion of an (A, B)-bimodule depends on $\mathbf{k}$ (even if we suppress $\mathbf{k}$ from the notation), because $\mathbf{k}$ plays a role in the condition (4).[1] Various authors use a slightly different notion of an (A, B)-bimodule, in whose definition the $\mathbf{k}$ does not occur[2]; their notion of an (A, B)-bimodule is equivalent to our notion of an $(A, B)_{\mathbb{Z}}$-bimodule. Our definition is thus more general. My impression is that several authors use our general notion of an (A, B)-bimodule while only defining their less general one; this has led to some confusion ([CSA14]), and is the reason why I have done Definition 1.2 in this much detail.

Proposition 1.3. Let A be a $\mathbf{k}$-algebra. Then, equipping A with the natural left A-module structure on A (which is given by $am = a \cdot m$ for every $a \in A$ and $m \in A$) and with the natural right A-module structure on A (which is given by $mb = m \cdot b$ for every $b \in A$ and $m \in A$) yields an (A, A)-bimodule. This (A, A)-bimodule will be called the (A, A)-*bimodule* A. We consider A equipped with this (A, A)-bimodule structure by default.

(Let us point out that the axiom (3) is satisfied for the (A, A)-bimodule A because the $\mathbf{k}$-algebra A is associative.)

Remark 1.4. Let A and B be $\mathbf{k}$-algebras. Let M be an (A, B)-bimodule. Then, the maps

$$A \times M \to M, \qquad (a, m) \mapsto am$$

and

$$M \times B \to M, \qquad (m, b) \mapsto mb$$

are $\mathbf{k}$-bilinear. (This follows easily from (4).)

We can now define the notion of a derivation:

[1] For example, if $\mathbb{H}$ denotes the ring of quaternions, then $\mathbb{H}$ becomes a $(\mathbb{C}, \mathbb{C})$-bimodule when $\mathbf{k} = \mathbb{R}$ (by multiplication from the left and from the right), but not when $\mathbf{k} = \mathbb{C}$ (because (4) does not hold for $\lambda \in \mathbb{C}$ and $m \in \mathbb{H}$).

[2] More precisely: They require A and B to be rings (rather than $\mathbf{k}$-algebras); they require M to be an abelian group (rather than a $\mathbf{k}$-module); and they omit the condition (4).

Definition 1.5. Let A be a $\mathbf{k}$-algebra. Let M be an (A, A)-bimodule. A $\mathbf{k}$-linear map $f : A \to M$ is said to be a $\mathbf{k}$-*derivation* if it satisfies

$$\left(f(ab) = af(b) + f(a)\, b \qquad \text{for every } a \in A \text{ and } b \in A\right). \tag{5}$$

We shall often abbreviate "$\mathbf{k}$-derivation" as "derivation", since the $\mathbf{k}$ is fixed. The equality (5) is usually called the *Leibniz law*.

Remark 1.6. When authors speak of a "derivation", they usually mean a $\mathbf{k}$-derivation, where $\mathbf{k}$ is some base ring which is **usually** clear from the context. Sometimes, it is $\mathbb{Z}$; sometimes, it is whatever base ring their modules are defined over; sometimes it is A. Unfortunately, in some situations (e.g., field extensions), the base ring is ambiguous; forgetting to specify it is then a common source of mistakes.

We notice that the product ab on the left hand side of (5) is a product inside the $\mathbf{k}$-algebra A, whereas the "products" $af(b)$ and $f(a)\, b$ on the right hand side of (5) are defined using the left and right A-module structures on M.

Example 1.7. Let $\mathbf{k}[x]$ denote the polynomial ring in an indeterminate x over $\mathbf{k}$. Then, the $\mathbf{k}$-linear map $\partial_x : \mathbf{k}[x] \to \mathbf{k}[x]$ defined by

$$\left(\partial_x\left(x^n\right) = n x^{n-1} \qquad \text{for every } n \in \mathbb{N}\right)$$

(where nx^{n-1} is to be interpreted as 0 in the case when $n = 0$) is a derivation. So is the map $p \cdot \partial_x$ (which sends every $g \in \mathbf{k}[x]$ to $p \cdot \partial_x(g)$) for every $p \in \mathbf{k}[x]$. More generally, if M is any $(\mathbf{k}[x], \mathbf{k}[x])$-bimodule, and if p is a *central* element of M (that is, an element of M satisfying $fp = pf$ for every $f \in \mathbf{k}[x]$), then the map $p \cdot \partial_x : \mathbf{k}[x] \to M$ (which sends every $g \in \mathbf{k}[x]$ to $p \cdot \partial_x(g)$) is a derivation.

The map $\mathrm{id} : \mathbf{k}[x] \to \mathbf{k}[x]$ is not a derivation (unless the ring $\mathbf{k}$ is trivial), and the map $\partial_x^2 : \mathbf{k}[x] \to \mathbf{k}[x]$ is not a derivation (unless $2 = 0$ in $\mathbf{k}$, in which case $\partial_x^2 = 0$).

Example 1.8. Let A be a $\mathbb{Z}$-graded $\mathbf{k}$-algebra. Let $\mathcal{E} : A \to A$ be the $\mathbf{k}$-linear map defined by the following property: If $n \in \mathbb{Z}$, and if $a \in A$ is a homogeneous element of degree n, then $\mathcal{E}(a) = na$. It is easy to see that $\mathcal{E}$ is a derivation.

Example 1.9. Let A be a $\mathbf{k}$-algebra. Let M be an (A, A)-bimodule. Let $p \in M$. Define a map $\mathrm{ad}_p : A \to M$ by

$$\left(\mathrm{ad}_p(a) = pa - ap \qquad \text{for every } a \in A\right). \tag{6}$$

Then, the map ad_p is a derivation. It is called an *inner derivation*.

Remark 1.10. Many authors use a notion of "derivation" that is less general than ours. Namely, given a $\mathbf{k}$-algebra A, they define a *derivation of* A to be a $\mathbf{k}$-linear map $f : A \to A$ satisfying (5). This concept of a "derivation" is a particular case of our concept defined above. Namely, it is precisely what we call a derivation from the $\mathbf{k}$-algebra A to the (A, A)-bimodule A.

Remark 1.11. Let A and B be two $\mathbf{k}$-algebras. Let $\rho : A \to B$ and $\tau : A \to B$ be two $\mathbf{k}$-algebra homomorphisms. Some authors define a (ρ, τ)-*derivation* to be a $\mathbf{k}$-linear map $f : A \to B$ satisfying

$$(f(ab) = \rho(a) f(b) + f(a) \tau(b) \qquad \text{for every } a \in A \text{ and } b \in A).$$

This concept of a "(ρ, τ)-derivation" is, too, a particular case of the concept of a "derivation" defined in Definition 1.5. Namely, the left B-module B becomes a left A-module via the $\mathbf{k}$-algebra homomorphism $\rho : A \to B$. Furthermore, the right B-module B becomes a right A-module via the $\mathbf{k}$-algebra homomorphism $\tau : A \to B$. Thus we have defined a left A-module structure on B and a right A-module structure on B. These two structures, combined, make B into an (A, A)-bimodule. Now, a (ρ, τ)-derivation from A to B is the same as a derivation from A to this (A, A)-bimodule B.

1.2. Basic properties

The following fact is probably the simplest property of derivations:

Theorem 1.12. Let A be a $\mathbf{k}$-algebra. Let M be an (A, A)-bimodule. Let $f : A \to M$ be a derivation. Then, $f(1) = 0$. (Keep in mind that 1 denotes the unity 1_A of A here.)

Proof of Theorem 1.12. The map $f : A \to M$ is a derivation. In other words, f is a $\mathbf{k}$-linear map and satisfies (5) (by the definition of a "derivation"). Applying (5) to $a = 1$ and $b = 1$, we obtain $f(1 \cdot 1) = \underbrace{1 f(1)}_{= f(1)} + \underbrace{f(1) 1}_{= f(1)} = f(1) + f(1)$. Since $1 \cdot 1 = 1$, this rewrites as $f(1) = f(1) + f(1)$. Subtracting $f(1)$ from this equality, we obtain $0 = f(1)$. This proves Theorem 1.12. □

Remark 1.13. Theorem 1.12 is extremely basic; it holds even more generally when A is a unital magmatic $\mathbf{k}$-algebra rather than a $\mathbf{k}$-algebra. (Of course, in this case, the definition of a derivation should be generalized appropriately.)

Notice, however, that Theorem 1.12 does not generalize to algebras over semirings (because we cannot subtract over semirings). When defining derivations in this generality, it thus is advisable to include the equality $f(1) = 0$ as an axiom in the definition of a derivation.

The following fact (sometimes called the *generalized Leibniz law*, or simply the *Leibniz law* again) generalizes (5):

Theorem 1.14. Let A be a $\mathbf{k}$-algebra. Let M be an (A, A)-bimodule. Let $f : A \to M$ be a derivation. Let $n \in \mathbb{N}$, and let $a_1, a_2, \ldots, a_n \in A$. Then,

$$f(a_1 a_2 \cdots a_n) = \sum_{i=1}^{n} a_1 a_2 \cdots a_{i-1} f(a_i) a_{i+1} a_{i+2} \cdots a_n.$$

Proof of Theorem 1.14. We shall prove Theorem 1.14 by induction over n:

Induction base: Theorem 1.14 holds for $n = 0$ [3]. This completes the induction base.

Induction step: Let $N \in \mathbb{N}$. Assume that Theorem 1.14 holds for $n = N$. We must prove that Theorem 1.14 holds for $n = N + 1$.

Let $a_1, a_2, \ldots, a_{N+1}$ be $N + 1$ elements of A. We assumed that Theorem 1.14 holds for $n = N$. Hence, we can apply Theorem 1.14 to N and $a_1, a_2, \ldots, a_N$ instead of n and $a_1, a_2, \ldots, a_n$. As a result, we obtain

$$f(a_1 a_2 \cdots a_N) = \sum_{i=1}^{N} a_1 a_2 \cdots a_{i-1} f(a_i) a_{i+1} a_{i+2} \cdots a_N. \tag{7}$$

But f is a derivation. In other words, f is a $\mathbf{k}$-linear map and satisfies (5). Hence, we can apply (5) to $a_1 a_2 \cdots a_N$ and a_{N+1} instead of a and b. As a result,

[3] *Proof.* Let $a_1, a_2, \ldots, a_0$ be 0 elements of A. Then, $a_1 a_2 \cdots a_0 = (\text{empty product}) = 1$, and thus $f\left(\underbrace{a_1 a_2 \cdots a_0}_{=1}\right) = f(1) = 0$ (by Theorem 1.12). Comparing this with

$$\sum_{i=1}^{0} a_1 a_2 \cdots a_{i-1} f(a_i) a_{i+1} a_{i+2} \cdots a_0 = (\text{empty sum}) = 0,$$

we obtain $f(a_1 a_2 \cdots a_0) = \sum_{i=1}^{0} a_1 a_2 \cdots a_{i-1} f(a_i) a_{i+1} a_{i+2} \cdots a_0$. Thus, we have shown that Theorem 1.14 holds for $n = 0$.

we obtain

$$\begin{aligned}
&f\left(\left(a_1a_2\cdots a_N\right)a_{N+1}\right)\\
&= \underbrace{\left(a_1a_2\cdots a_N\right)f\left(a_{N+1}\right)}_{=a_1a_2\cdots a_Nf(a_{N+1})} + \underbrace{f\left(a_1a_2\cdots a_N\right)}_{=\sum\limits_{i=1}^{N} a_1a_2\cdots a_{i-1}f(a_i)a_{i+1}a_{i+2}\cdots a_N \text{ (by (7))}} a_{N+1}\\
&= a_1a_2\cdots a_Nf\left(a_{N+1}\right) + \underbrace{\left(\sum_{i=1}^{N} a_1a_2\cdots a_{i-1}f\left(a_i\right)a_{i+1}a_{i+2}\cdots a_N\right)a_{N+1}}_{=\sum\limits_{i=1}^{N}(a_1a_2\cdots a_{i-1}f(a_i)a_{i+1}a_{i+2}\cdots a_N)a_{N+1}}\\
&= a_1a_2\cdots a_Nf\left(a_{N+1}\right) + \sum_{i=1}^{N} \underbrace{\left(a_1a_2\cdots a_{i-1}f\left(a_i\right)a_{i+1}a_{i+2}\cdots a_N\right)a_{N+1}}_{\substack{=a_1a_2\cdots a_{i-1}f(a_i)a_{i+1}a_{i+2}\cdots a_Na_{N+1}\\ =a_1a_2\cdots a_{i-1}f(a_i)a_{i+1}a_{i+2}\cdots a_{N+1}}}\\
&= a_1a_2\cdots a_Nf\left(a_{N+1}\right) + \sum_{i=1}^{N} a_1a_2\cdots a_{i-1}f\left(a_i\right)a_{i+1}a_{i+2}\cdots a_{N+1}.
\end{aligned}$$

Compared with

$$\begin{aligned}
&\sum_{i=1}^{N+1} a_1a_2\cdots a_{i-1}f\left(a_i\right)a_{i+1}a_{i+2}\cdots a_{N+1}\\
&= \underbrace{a_1a_2\cdots a_{(N+1)-1}}_{=a_1a_2\cdots a_N} f\left(a_{N+1}\right) \underbrace{a_{(N+1)+1}a_{(N+1)+2}\cdots a_{N+1}}_{=(\text{empty product})=1}\\
&\qquad + \sum_{i=1}^{N} a_1a_2\cdots a_{i-1}f\left(a_i\right)a_{i+1}a_{i+2}\cdots a_{N+1}\\
&\qquad (\text{here, we have split off the addend for } i = N+1 \text{ from the sum})\\
&= a_1a_2\cdots a_N \underbrace{f\left(a_{N+1}\right)1}_{=f(a_{N+1})} + \sum_{i=1}^{N} a_1a_2\cdots a_{i-1}f\left(a_i\right)a_{i+1}a_{i+2}\cdots a_{N+1}\\
&= a_1a_2\cdots a_Nf\left(a_{N+1}\right) + \sum_{i=1}^{N} a_1a_2\cdots a_{i-1}f\left(a_i\right)a_{i+1}a_{i+2}\cdots a_{N+1},
\end{aligned}$$

this shows that

$$f\left(a_1a_2\cdots a_{N+1}\right) = \sum_{i=1}^{N+1} a_1a_2\cdots a_{i-1}f\left(a_i\right)a_{i+1}a_{i+2}\cdots a_{N+1}. \tag{8}$$

Now, let us forget that we fixed $a_1, a_2, \ldots, a_{N+1}$. We thus have shown that (8) holds for any $a_1, a_2, \ldots, a_{N+1} \in A$. In other words, Theorem 1.14 holds for $n = N+1$. This completes the induction step. Thus, the proof of Theorem 1.14 is complete. □

Let us gather some consequences of Theorem 1.14:

Corollary 1.15. Let A be a $\mathbf{k}$-algebra. Let M be an (A, A)-bimodule. Let $f : A \to M$ be a derivation. Let $n \in \mathbb{N}$, and let $a \in A$. Then,

$$f\left(a^n\right) = \sum_{i=1}^{n} a^{i-1} f\left(a\right) a^{n-i}.$$

Proof of Corollary 1.15. Set $a_k = a$ in Theorem 1.14. □

For the next corollary, we need another definition:

Definition 1.16. Let A be a $\mathbf{k}$-algebra. Let M be an (A, A)-bimodule. The (A, A)-bimodule M is said to be *symmetric* if it satisfies

$$\left(am = ma \qquad \text{for every } a \in A \text{ and } m \in M\right). \tag{9}$$

Remark 1.17. If A is a commutative ring, then the symmetric (A, A)-bimodules can be identified with the A-modules. Namely, a symmetric (A, A)-bimodule structure is the same as an A-module structure, and conversely, any A-module structure can be interpreted both as a left A-module structure and as a right A-module structure, and the latter two structures combined yield a symmetric (A, A)-bimodule structure.

Corollary 1.18. Let A be a $\mathbf{k}$-algebra. Let M be a symmetric (A, A)-bimodule. Let $f : A \to M$ be a derivation. Let $n \in \mathbb{N}$, and let $a \in A$. Then,

$$f\left(a^n\right) = nf\left(a\right) a^{n-1}.$$

Here, $nf\left(a\right) a^{n-1}$ is to be understood as 0 if $n = 0$.

Proof of Corollary 1.18. Corollary 1.15 yields

$$\begin{aligned} f\left(a^n\right) &= \sum_{i=1}^{n} \underbrace{a^{i-1} f\left(a\right)}_{\substack{= f(a) a^{i-1} \\ \text{(since } M \text{ is symmetric)}}} \underbrace{a^{n-i}}_{= a^{(i-1)+(n-i)} = a^{n-1}} = \sum_{i=1}^{n} f\left(a\right) \underbrace{a^{i-1} a^{n-i}}_{= a^{(i-1)+(n-i)} = a^{n-1}} \\ &= \sum_{i=1}^{n} f\left(a\right) a^{n-1} = nf\left(a\right) a^{n-1}. \end{aligned}$$

□

The following corollary can be regarded as a "chain rule" for derivations:

Corollary 1.19. Let A be a $\mathbf{k}$-algebra. Let M be a symmetric (A, A)-bimodule. Let $f : A \to M$ be a derivation. Let $p \in \mathbf{k}[x]$ be a polynomial. Let $a \in A$. Then,

$$f(p(a)) = p'(a) \cdot f(a).$$

Here, p' denotes the (usual) derivative of p in $\mathbf{k}[x]$.

Proof of Corollary 1.19. Write the polynomial p in the form $p = \sum_{n=0}^{N} p_n x^n$ for some $N \in \mathbb{N}$ and some $(p_0, p_1, \ldots, p_N) \in \mathbf{k}^{N+1}$. Then, the definition of p' shows that $p' = \sum_{n=1}^{N} n p_n x^{n-1}$. Substituting a for x in this equality, we obtain $p'(a) = \sum_{n=1}^{N} n p_n a^{n-1}$.

We have $f\left(\underbrace{a^0}_{=1}\right) = f(1) = 0$ (by Theorem 1.12), and thus $p_0 \underbrace{f\left(a^0\right)}_{=0} = 0$.

On the other hand, substituting a for x in the equality $p = \sum_{n=0}^{N} p_n x^n$, we obtain $p(a) = \sum_{n=0}^{N} p_n a^n$. Applying the map f to both sides of this equality, we obtain

$$\begin{aligned}
f(p(a)) &= f\left(\sum_{n=0}^{N} p_n a^n\right) = \sum_{n=0}^{N} p_n f(a^n) \qquad \text{(since the map } f \text{ is } \mathbf{k}\text{-linear)} \\
&= \underbrace{p_0 f\left(a^0\right)}_{=0} + \sum_{n=1}^{N} p_n f(a^n) \\
&\qquad \text{(here, we have split off the addend for } n = 0 \text{ from the sum)} \\
&= \sum_{n=1}^{N} p_n \underbrace{f(a^n)}_{\substack{=nf(a)a^{n-1} \\ \text{(by Corollary 1.18)}}} = \sum_{n=1}^{N} \underbrace{p_n n}_{=np_n} \underbrace{f(a)\, a^{n-1}}_{\substack{=a^{n-1}f(a) \\ \text{(because (9) (applied to } a^{n-1} \text{ and } f(a) \\ \text{instead of } a \text{ and } m) \text{ shows} \\ \text{that } a^{n-1}f(a)=f(a)a^{n-1})}} \\
&= \sum_{n=1}^{N} n p_n a^{n-1} f(a) = \underbrace{\left(\sum_{n=1}^{N} n p_n a^{n-1}\right)}_{=p'(a)} \cdot f(a) = p'(a) \cdot f(a).
\end{aligned}$$

This proves Corollary 1.19. □

Definition 1.20. Let A be a $\mathbf{k}$-algebra.

(a) If $I_1, I_2, \ldots, I_n$ are some two-sided ideals of A, then $I_1 I_2 \cdots I_n$ denotes the two-sided ideal of A generated by

$$\{u_1 u_2 \cdots u_n \mid (u_1, u_2, \ldots, u_n) \in I_1 \times I_2 \times \cdots \times I_n\}.$$

(In particular, for $n = 0$, this means that $I_1 I_2 \cdots I_n$ is the two-sided ideal of A generated by the empty product. Since the empty product is defined to be 1_A, this ideal is therefore A.)

(b) If I is a two-sided ideal of A, and if $n \in \mathbb{N}$, then I^n is defined to be $\underbrace{II\cdots I}_{n \text{ times}}$.

[**Note:** This notation is not really standard, particularly as far as the meaning of an empty product of ideals is concerned.]

(c) If S is any set, and if $n \in \mathbb{N}$, then we use the notation $S^{\times n}$ for the Cartesian product $\underbrace{S \times S \times \cdots \times S}_{n \text{ times}}$. (This is commonly denoted by S^n, but this notation would conflict with the one in Definition 1.20 **(b)**.)

Proposition 1.21. Let A be a $\mathbf{k}$-algebra. Let I be a two-sided ideal of A. Let $f : A \to A$ be a derivation. Then, $f\left(I^{n+1}\right) \subseteq I^n$ for every $n \in \mathbb{N}$.

Proof of Proposition 1.21. It is not hard to derive Proposition 1.21 from Theorem 1.14. We shall proceed differently, however.

We shall prove Proposition 1.21 by induction over n:

Induction base: We have $I^0 = A$. Now, $f\left(I^{0+1}\right) \subseteq A = I^0$. In other words, Proposition 1.21 holds for $n = 0$. This completes the induction base.

Induction step: Let $N \in \mathbb{N}$. Assume that Proposition 1.21 holds for $n = N$. We must prove that Proposition 1.21 holds for $n = N + 1$.

Proposition 1.21 holds for $n = N$. In other words, we have $f\left(I^{N+1}\right) \subseteq I^N$.

Now, let $r \in f\left(I^{(N+1)+1}\right)$. Thus, there exists a $q \in I^{(N+1)+1}$ such that $r = f(q)$. Consider this q. We have $q \in I^{(N+1)+1} = I^{N+1} \underbrace{I^1}_{=I} = I^{N+1} I$. But it is easy to see the following general fact: If U and V are two two-sided ideals of A, then every element of UV has the form $\sum_{i=1}^{m} u_i v_i$ for some $m \in \mathbb{N}$ and some two m-tuples $(u_1, u_2, \ldots, u_m) \in U^{\times m}$ and $(v_1, v_2, \ldots, v_m) \in V^{\times m}$ [4]. Applying this to $U = I^{N+1}$ and $V = I$ and to the element q of $I^{N+1} I$, we conclude that the element

[4] *Proof.* Let x be an element of UV. We need to prove that x has the form $\sum_{i=1}^{m} u_i v_i$ for some $m \in \mathbb{N}$ and some two m-tuples $(u_1, u_2, \ldots, u_m) \in U^{\times m}$ and $(v_1, v_2, \ldots, v_m) \in V^{\times m}$.

The ideal UV is the two-sided ideal generated by the set $\{uv \mid (u, v) \in U \times V\}$. Thus, the ideal UV is the set of all sums of the form $\sum_{i=1}^{p} a_i u_i' v_i' b_i$ with $p \in \mathbb{N}$, $(a_1, a_2, \ldots, a_p) \in A^{\times p}$,

q has the form $\sum\limits_{i=1}^{m} u_i v_i$ for some $m \in \mathbb{N}$ and some two m-tuples $(u_1, u_2, \ldots, u_m) \in \left(I^{N+1}\right)^{\times m}$ and $(v_1, v_2, \ldots, v_m) \in I^{\times m}$. Consider this m and these $(u_1, u_2, \ldots, u_m)$ and $(v_1, v_2, \ldots, v_m)$. Clearly, $u_i \in I^{N+1}$ and $v_i \in I$ for every $i \in \{1, 2, \ldots, m\}$. Hence, for every $i \in \{1, 2, \ldots, m\}$, we have $f\left(\underbrace{u_i}_{\in I^{N+1}}\right) \in f\left(I^{N+1}\right) \subseteq I^N$ and thus

$$\underbrace{f(u_i)}_{\in I^N} \underbrace{v_i}_{\in I} \in I^N I = I^{N+1}. \tag{10}$$

The map f is a derivation. Thus, it is $\mathbf{k}$-linear and satisfies (5). Applying the map f to both sides of the equality $q = \sum\limits_{i=1}^{m} u_i v_i$, we obtain

$$f(q) = f\left(\sum_{i=1}^{m} u_i v_i\right) = \sum_{i=1}^{m} \underbrace{f(u_i v_i)}_{\substack{= u_i f(v_i) + f(u_i) v_i \\ \text{(by (5), applied to } a = u_i \text{ and } b = v_i\text{)}}} \qquad \text{(since the map } f \text{ is } \mathbf{k}\text{-linear)}$$

$$= \sum_{i=1}^{m} \left(\underbrace{u_i f(v_i)}_{\substack{\in I^{N+1} \\ \text{(since } u_i \in I^{N+1}\text{, and} \\ \text{since } I^{N+1} \text{ is an ideal)}}} + \underbrace{f(u_i) v_i}_{\substack{\in I^{N+1} \\ \text{(by (10))}}} \right) \in \sum_{i=1}^{m} \left(I^{N+1} + I^{N+1} \right)$$

$$\subseteq I^{N+1} \qquad \left(\text{since } I^{N+1} \text{ is an ideal of } A\right).$$

Thus, $r = f(q) \in I^{N+1}$.

Now let us forget that we fixed r. We thus have proven that $r \in I^{N+1}$ for every $r \in f\left(I^{(N+1)+1}\right)$. In other words, $f\left(I^{(N+1)+1}\right) \subseteq I^{N+1}$. In other words, Proposition 1.21 holds for $n = N + 1$. This completes the induction step. Proposition 1.21 is thus proven by induction. □

$\left(u_1', u_2', \ldots, u_p'\right) \in U^{\times p}$, $\left(v_1', v_2', \ldots, v_p'\right) \in V^{\times p}$ and $(b_1, b_2, \ldots, b_p) \in A^{\times p}$. In particular, x therefore has this form (since $x \in UV$). So let us write x in the form $x = \sum\limits_{i=1}^{p} a_i u_i' v_i' b_i$ with $p \in \mathbb{N}$, $(a_1, a_2, \ldots, a_p) \in A^{\times p}$, $\left(u_1', u_2', \ldots, u_p'\right) \in U^{\times p}$, $\left(v_1', v_2', \ldots, v_p'\right) \in V^{\times p}$ and $(b_1, b_2, \ldots, b_p) \in A^{\times p}$. For every $i \in \{1, 2, \ldots, p\}$, we have $a_i u_i' \in U$ (since $u_i' \in U$ and since U is an ideal) and $v_i' b_i \in V$ (since $v_i' \in V$ and since V is an ideal). Hence, x has the form $\sum\limits_{i=1}^{m} u_i v_i$ for some $m \in \mathbb{N}$ and some two m-tuples $(u_1, u_2, \ldots, u_m) \in U^{\times m}$ and $(v_1, v_2, \ldots, v_m) \in V^{\times m}$ (namely, for $m = p$, $u_i = a_i u_i'$ and $v_i = v_i' b_i$). This is what we wanted to prove.

1.3. The module of derivations

We have so far studied properties of a single derivation. Let us next consider the set of all derivations.

Definition 1.22. Let A be a $\mathbf{k}$-algebra. Let M be an (A, A)-bimodule. We let $\operatorname{Der}(A, M)$ denote the set of all derivations from A to M.

Proposition 1.23. Let A be a $\mathbf{k}$-algebra. Let M be an (A, A)-bimodule. Then, $\operatorname{Der}(A, M)$ is a $\mathbf{k}$-submodule of $\operatorname{Hom}(A, M)$.

Proof of Proposition 1.23. All derivations are $\mathbf{k}$-linear; hence, $\operatorname{Der}(A, M) \subseteq \operatorname{Hom}(A, M)$. Also, the equation (5) is $\mathbf{k}$-linear in d. Thus, any $\mathbf{k}$-linear combination of derivations from A to M is a derivation from A to M. In other words, $\operatorname{Der}(A, M)$ is a $\mathbf{k}$-submodule of $\operatorname{Hom}(A, M)$. □

A more interesting property of $\operatorname{Der}(A, M)$ holds in the case when $M = A$:

Theorem 1.24. Let A be a $\mathbf{k}$-algebra. Let $f \in \operatorname{Der}(A, A)$ and $g \in \operatorname{Der}(A, A)$. (See Proposition 1.3 for the definition of the (A, A)-bimodule A that is used here.) Then, $f \circ g - g \circ f \in \operatorname{Der}(A, A)$.

Example 1.25. Let A be a $\mathbf{k}$-algebra. Let $x \in A$ and $y \in A$. As we know from Example 1.9, for every $p \in A$, we can define a map $\operatorname{ad}_p : A \to A$ by

$$\left(\operatorname{ad}_p(a) = pa - ap \qquad \text{for every } a \in A\right).$$

Thus, we have two maps $\operatorname{ad}_x : A \to A$ and $\operatorname{ad}_y : A \to A$. Example 1.9 shows that these two maps are derivations, i.e., belong to $\operatorname{Der}(A, A)$. Hence, Theorem 1.24 (applied to $f = \operatorname{ad}_x$ and $g = \operatorname{ad}_y$) yields that $\operatorname{ad}_x \circ \operatorname{ad}_y - \operatorname{ad}_y \circ \operatorname{ad}_x \in \operatorname{Der}(A, A)$. This can also be proven in a simpler way: A short calculation confirms that $\operatorname{ad}_x \circ \operatorname{ad}_y - \operatorname{ad}_y \circ \operatorname{ad}_x = \operatorname{ad}_{xy-yx} \in \operatorname{Der}(A, A)$ (again by Example 1.9).

Remark 1.26. Theorem 1.24 is usually worded in the language of Lie algebras. Namely, let A be a $\mathbf{k}$-algebra. We shall use the standard notation $\operatorname{End} A$ for the $\mathbf{k}$-algebra $\operatorname{Hom}(A, A)$ (where the multiplication is given by the composition of maps, and where the unity is the identity map id_A). Then, we can define a Lie algebra structure on the $\mathbf{k}$-module $\operatorname{End} A$ by setting

$$\left([f, g] = f \circ g - g \circ f \qquad \text{for every } f \in \operatorname{End} A \text{ and } g \in \operatorname{End} A\right).$$

This Lie algebra is denoted by $(\operatorname{End} A)^-$ or by $\mathfrak{gl}(A)$. Now, Theorem 1.24 states that $\operatorname{Der}(A, A)$ is a Lie subalgebra of this Lie algebra $(\operatorname{End} A)^-$. When f and g are two elements of $\operatorname{End} A$, the element $[f, g]$ is called the *commutator* of f and g.

Rather than proving Theorem 1.24, we shall prove a more general result:

Theorem 1.27. Let A be a $\mathbf{k}$-algebra. Let M be an (A, A)-bimodule. Let $f : A \to M$ and $h : A \to A$ be two derivations. Let $g : M \to M$ be a $\mathbf{k}$-linear map. Assume that

$$(g(am) = ag(m) + h(a)\, m \qquad \text{for every } a \in A \text{ and } m \in M) \tag{11}$$

and

$$(g(mb) = mh(b) + g(m)\, b \qquad \text{for every } b \in A \text{ and } m \in M). \tag{12}$$

Then, $f \circ h - g \circ f \in \operatorname{Der}(A, M)$.

Proof of Theorem 1.27. Let $a \in A$ and $b \in A$. Then,

$$\begin{aligned}
(f \circ h)(ab) &= f\left(\underbrace{h(ab)}_{\substack{=ah(b)+h(a)b \\ \text{(since } h \text{ is a derivation)}}}\right) = f(ah(b) + h(a)\, b) \\
&= \underbrace{f(ah(b))}_{\substack{=af(h(b))+f(a)h(b) \\ \text{(since } f \text{ is a derivation)}}} + \underbrace{f(h(a)\, b)}_{\substack{=h(a)f(b)+f(h(a))b \\ \text{(since } f \text{ is a derivation)}}} \qquad \text{(since the map } f \text{ is } \mathbf{k}\text{-linear)} \\
&= (af(h(b)) + f(a)\, h(b)) + (h(a)\, f(b) + f(h(a))\, b). \qquad (13)
\end{aligned}$$

On the other hand,

$$\begin{aligned}
(g \circ f)(ab) &= g\left(\underbrace{f(ab)}_{\substack{=af(b)+f(a)b \\ \text{(since } f \text{ is a derivation)}}}\right) = g(af(b) + f(a)\, b) \\
&= \underbrace{g(af(b))}_{\substack{=ag(f(b))+h(a)f(b) \\ \text{(by (11), applied to} \\ f(b) \text{ instead of } m)}} + \underbrace{g(f(a)\, b)}_{\substack{=f(a)h(b)+g(f(a))b \\ \text{(by (12), applied to} \\ f(a) \text{ instead of } m)}} \qquad \text{(since the map } g \text{ is } \mathbf{k}\text{-linear)} \\
&= (ag(f(b)) + h(a)\, f(b)) + (f(a)\, h(b) + g(f(a))\, b).
\end{aligned}$$

Subtracting this equality from (13), we obtain

$$\begin{aligned}
&(f \circ h)(ab) - (g \circ f)(ab) \\
&= ((af(h(b)) + f(a)\, h(b)) + (h(a)\, f(b) + f(h(a))\, b)) \\
&\qquad - ((ag(f(b)) + h(a)\, f(b)) + (f(a)\, h(b) + g(f(a))\, b)) \\
&= af(h(b)) + f(h(a))\, b - ag(f(b)) - g(f(a))\, b \\
&= \underbrace{af(h(b)) - ag(f(b))}_{=a(f(h(b))-g(f(b)))} + \underbrace{f(h(a))\, b - g(f(a))\, b}_{=(f(h(a))-g(f(a)))b} \\
&= a\underbrace{(f(h(b)) - g(f(b)))}_{=(f\circ h - g\circ f)(b)} + \underbrace{(f(h(a)) - g(f(a)))}_{=(f\circ h - g\circ f)(a)} b \\
&= a(f \circ h - g \circ f)(b) + (f \circ h - g \circ f)(a)\, b.
\end{aligned}$$

Thus,

$$\begin{aligned}
(f \circ h - g \circ f)(ab) &= (f \circ h)(ab) - (g \circ f)(ab) \\
&= a(f \circ h - g \circ f)(b) + (f \circ h - g \circ f)(a)\, b.
\end{aligned}$$

Let us now forget that we fixed a and b. We thus have proven that

$$(f \circ h - g \circ f)(ab) = a(f \circ h - g \circ f)(b) + (f \circ h - g \circ f)(a)\, b$$

for every $a \in A$ and $b \in A$. Thus, $f \circ h - g \circ f$ is a derivation from A to M (since $f \circ h - g \circ f$ is $\mathbf{k}$-linear). In other words, $f \circ h - g \circ f \in \operatorname{Der}(A, M)$. This proves Theorem 1.27. □

Proof of Theorem 1.24. Set $M = A$ and $h = g$. Recall that g is a derivation; thus, $g(ab) = ag(b) + g(a)\, b$ for any $a \in A$ and $b \in A$. Hence, the equalities (11) and (12) hold (because $h = g$). Theorem 1.27 thus shows that $f \circ h - g \circ f \in \operatorname{Der}(A, M)$. Since $h = g$ and $M = A$, this rewrites as $f \circ g - g \circ f \in \operatorname{Der}(A, A)$. □

Let us state another particular case of Theorem 1.27.

Corollary 1.28. Let B be a $\mathbf{k}$-algebra. Let A be a $\mathbf{k}$-subalgebra of B. Thus, B becomes an (A, A)-bimodule. (The left and the right A-module structures on B are given by multiplication inside B.) Let $f : A \to B$ and $g : B \to B$ be two derivations such that $g(A) \subseteq A$. Then, $f \circ (g\mid_A) - g \circ f : A \to B$ is a derivation.

This corollary is sometimes useful (e.g., it appears in [EtiGri12, Proposition 4.6.22][5], and is used there).

[5] More precisely, the particular case of Corollary 1.28 when $\mathbf{k} = \mathbb{C}$ appears in [EtiGri12, Proposition 4.6.22]. But the general case is proven in the same way as this particular case.

Notice that Theorem 1.24 (in the particular case when $\mathbf{k} = \mathbb{C}$) is [EtiGri12, Proposition 4.6.21].

Proof of Corollary 1.28. The map $g\mid_A\colon A \to A$ is a derivation (since $g : B \to B$ is a derivation). Moreover, the equalities (11) and (12) hold for $M = B$ and $h = g\mid_A$ (again because $g : B \to B$ is a derivation). Thus, Theorem 1.27 (applied to $M = B$ and $h = g\mid_A$) shows that $f \circ (g\mid_A) - g \circ f \in \operatorname{Der}(A, B)$. □

Let us now prove another fact about derivations.[6]

Proposition 1.29. Let A be a $\mathbf{k}$-algebra. Let M be an (A, A)-bimodule. Let $d : A \to M$ and $e : A \to M$ be two derivations. Let S be a subset of A which generates A as a $\mathbf{k}$-algebra. Assume that $d\mid_S = e\mid_S$. Then, $d = e$.

Proposition 1.29 shows that a derivation $A \to M$ is uniquely determined by its restriction to a generating set of the $\mathbf{k}$-algebra A. Of course, this does not yield that any map from such a generating set can be extended to a derivation $A \to M$ (though we will see some situations when such extensions are possible later[7]).

Proof of Proposition 1.29. We have $d\mid_S = e\mid_S$. In other words, $d(s) = e(s)$ for every $s \in S$.

Let z be the map $d - e : A \to M$. Then, $z(s) = 0$ for every $s \in S$ (since $d(s) = e(s)$ for every $s \in S$). In other words, $S \subseteq \operatorname{Ker} z$.

The maps d and e are two derivations, and thus belong to $\operatorname{Der}(A, M)$. Hence, $d - e \in \operatorname{Der}(A, M)$ as well (since Proposition 1.23 shows that $\operatorname{Der}(A, M)$ is a $\mathbf{k}$-submodule of $\operatorname{Hom}(A, M)$). Hence, $z = d - e \in \operatorname{Der}(A, M)$. In other words, z is a derivation from A to M. Hence, $z(1) = 0$ (by Theorem 1.12, applied to $f = z$).

Also, z is a derivation and thus $\mathbf{k}$-linear. Hence, $\operatorname{Ker} z$ is a $\mathbf{k}$-submodule of A. This $\mathbf{k}$-submodule $\operatorname{Ker} z$ furthermore satisfies $1 \in \operatorname{Ker} z$ (since $z(1) = 0$) and

$$(ab \in \operatorname{Ker} z \text{ for every } a \in \operatorname{Ker} z \text{ and } b \in \operatorname{Ker} z)$$

[8]; therefore, this $\mathbf{k}$-submodule $\operatorname{Ker} z$ is a $\mathbf{k}$-subalgebra of A. More precisely, it is a $\mathbf{k}$-subalgebra of A which contains S as a subset (since $S \subseteq \operatorname{Ker} z$).

But recall that the subset S generates A as a $\mathbf{k}$-algebra. Hence, the smallest $\mathbf{k}$-subalgebra of A which contains S as a subset must be A itself. Hence, if B is any $\mathbf{k}$-subalgebra of A which contains S as a subset, then $A \subseteq B$. We can apply this to $B = \operatorname{Ker} z$ (since $\operatorname{Ker} z$ is a $\mathbf{k}$-subalgebra of A which contains S as a subset). Thus, we obtain $A \subseteq \operatorname{Ker} z$. Hence, $z = 0$. Since $z = d - e$, this rewrites as $d - e = 0$. Hence, $d = e$. Proposition 1.29 is thus proven. □

[6] Proposition 1.29 appears in [EtiGri12, Proposition 4.6.13].

[7] in Proposition 2.3 and Proposition 2.7

[8] *Proof.* Let $a \in \operatorname{Ker} z$ and $b \in \operatorname{Ker} z$. We have $z(a) = 0$ (since $a \in \operatorname{Ker} z$) and $z(b) = 0$ (since $b \in \operatorname{Ker} z$). But z is a derivation. Hence,

$$z(ab) = a\underbrace{z(b)}_{=0} + \underbrace{z(a)}_{=0}b = 0 + 0 = 0.$$

In other words, $ab \in \operatorname{Ker} z$, qed.

Remark 1.30. Again, much of what we have proven (e.g., Proposition 1.23, Theorem 1.24, Theorem 1.27 and Proposition 1.29) can be generalized to the situation when A is a magmatic $\mathbf{k}$-algebra (instead of being a $\mathbf{k}$-algebra), provided that the definition of a derivation is properly adjusted to this generality.

1.4. The k-algebra $\mathcal{R}_A(M)$

We shall now introduce a construction that will allow us to reduce questions about derivations to questions about $\mathbf{k}$-algebra homomorphisms.

Theorem 1.31. Let A be a $\mathbf{k}$-algebra. Let M be an (A, A)-bimodule. Consider the $\mathbf{k}$-module $A \oplus M$. Define a map $m : (A \oplus M) \times (A \oplus M) \to A \oplus M$ by

$$(m((a,p),(b,q)) = (ab, aq+pb) \qquad \text{for all } (a,p) \in A \oplus M \text{ and } (b,q) \in A \oplus M).$$

Then, the map m is $\mathbf{k}$-bilinear. Furthermore, the $\mathbf{k}$-module $A \oplus M$, equipped with the $\mathbf{k}$-bilinear map m, becomes a $\mathbf{k}$-algebra with unity $(1, 0)$.

Definition 1.32. Let A be a $\mathbf{k}$-algebra. Let M be an (A, A)-bimodule. Consider the $\mathbf{k}$-module $A \oplus M$, and the map m defined in Theorem 1.31. Theorem 1.31 shows that the $\mathbf{k}$-module $A \oplus M$, equipped with the $\mathbf{k}$-bilinear map m, becomes a $\mathbf{k}$-algebra with unity $(1, 0)$. This $\mathbf{k}$-algebra will be denoted by $\mathcal{R}_A(M)$. Thus, its multiplication satisfies

$$(a,p) \cdot (b,q) = m((a,p),(b,q)) = (ab, aq+pb) \qquad (14)$$
$$\text{for all } (a,p) \in A \oplus M \text{ and } (b,q) \in A \oplus M.$$

Proof of Theorem 1.31. The map m is $\mathbf{k}$-bilinear[9]. Thus, the $\mathbf{k}$-module $A \oplus M$, equipped with the $\mathbf{k}$-bilinear map m, is a magmatic $\mathbf{k}$-algebra. We denote this magmatic $\mathbf{k}$-algebra by $\mathbf{R}$. Thus, $\mathbf{R} = A \oplus M$ as a $\mathbf{k}$-module, and the multiplication on $\mathbf{R}$ is the map m. Thus, the multiplication on $\mathbf{R}$ is given by

$$(a,p)(b,q) = m((a,p),(b,q)) = (ab, aq+pb) \qquad (15)$$
$$\text{for all } (a,p) \in A \oplus M \text{ and } (b,q) \in A \oplus M.$$

[9]The proof of this is just a straightforward application of the facts that

- the multiplication of the $\mathbf{k}$-algebra A (that is, the map $A \times A \to A$, $(a,b) \mapsto ab$) is $\mathbf{k}$-bilinear;
- the map $A \times M \to M$, $(a,n) \mapsto an$ is $\mathbf{k}$-bilinear;
- the map $M \times A \to M$, $(n,a) \mapsto na$ is $\mathbf{k}$-bilinear.

The magmatic $\mathbf{k}$-algebra $\mathbf{R}$ is associative[10]. Furthermore, straightforward computations show that

$$(1,0)\,a = a\,(1,0) = a \qquad \text{for every } a \in \mathbf{R}.$$

Thus, the magmatic $\mathbf{k}$-algebra $\mathbf{R}$, equipped with the element $(1,0)$, becomes a unital magmatic $\mathbf{k}$-algebra. Since it is also associative, it is therefore a $\mathbf{k}$-algebra with unity $(1,0)$. □

Definition 1.33. Let A be a $\mathbf{k}$-algebra. Let M be an (A,A)-bimodule. Recall that $\mathcal{R}_A(M) = A \oplus M$ as a $\mathbf{k}$-module.

Let $\mathbf{projl}_{A,M} : A \oplus M \to A$ and $\mathbf{projr}_{A,M} : A \oplus M \to M$ be the canonical projections from the direct sum $A \oplus M$ to its two addends. These two projections $\mathbf{projl}_{A,M}$ and $\mathbf{projr}_{A,M}$ are two surjective $\mathbf{k}$-linear maps. Explicitly, they are given by

$$\left(\mathbf{projl}_{A,M}((a,m)) = a \qquad \text{for every } (a,m) \in A \oplus M\right)$$

[10] *Proof.* We want to show that the magmatic $\mathbf{k}$-algebra $\mathbf{R}$ is associative. In order to do so, it is clearly enough to show that $x\,(yz) = (xy)\,z$ for all $x \in \mathbf{R}$, $y \in \mathbf{R}$ and $z \in \mathbf{R}$. So let us do this.

Fix $x \in \mathbf{R}$, $y \in \mathbf{R}$ and $z \in \mathbf{R}$. The elements x, y and z belong to $A \oplus M$, and therefore can be written in the forms $x = (a,p)$, $y = (b,q)$ and $z = (c,r)$ for some elements a, b and c of A and some elements p, q and r of M. Consider these a, b and c and these p, q and r.

We have

$$\underbrace{y}_{=(b,q)}\,\underbrace{z}_{=(c,r)} = (b,q)\,(c,r) = (bc, br+qc)$$

(by (15)) and thus

$$\underbrace{x}_{=(a,p)}\,\underbrace{(yz)}_{=(bc,br+qc)} = (a,p)\,(bc,br+qc) = \left(a\,(bc)\,,\ \underbrace{a\,(br+qc)}_{=a(br)+a(qc)} + p\,(bc)\right)$$

(by (15), applied to $(bc, br+qc)$ instead of (b,q))

$$= (a\,(bc)\,,\ a\,(br) + a\,(qc) + p\,(bc))\,. \tag{16}$$

A similar computation shows that

$$(xy)\,z = ((ab)\,c,\ (ab)\,r + (aq)\,c + (pb)\,c)\,. \tag{17}$$

But $(ab)\,c = a\,(bc)$ (since A is associative), and $(ab)\,r = a\,(br)$ (since M is a left A-module), and $(aq)\,c = a\,(qc)$ (by (3), applied to A, q and c instead of B, m and b), and $(pb)\,c = p\,(bc)$ (since M is a right A-module). Hence, (17) becomes

$$(xy)\,z = \left(\underbrace{(ab)\,c}_{=a(bc)},\ \underbrace{(ab)\,r}_{=a(br)} + \underbrace{(aq)\,c}_{=a(qc)} + \underbrace{(pb)\,c}_{=p(bc)}\right) = (a\,(bc)\,,\ a\,(br) + a\,(qc) + p\,(bc))$$

$$= x\,(yz) \qquad \text{(by (16))}\,.$$

Thus, $x\,(yz) = (xy)\,z$ is proven, qed.

and

$$\left(\mathbf{projr}_{A,M}\left((a,m)\right) = m \qquad \text{for every } (a,m) \in A \oplus M\right).$$

The map $\mathbf{projl}_{A,M}$ is a $\mathbf{k}$-linear map $A \oplus M \to A$, therefore a $\mathbf{k}$-linear map $\mathcal{R}_A(M) \to A$ (since $\mathcal{R}_A(M) = A \oplus M$).

The map $\mathbf{projr}_{A,M}$ is a $\mathbf{k}$-linear map $A \oplus M \to M$, therefore a $\mathbf{k}$-linear map $\mathcal{R}_A(M) \to M$ (since $\mathcal{R}_A(M) = A \oplus M$).

Let $\mathbf{inl}_{A,M} : A \to A \oplus M$ and $\mathbf{inr}_{A,M} : M \to A \oplus M$ be the canonical injections of the two $\mathbf{k}$-modules A and M into their direct sum $A \oplus M$. These two injections $\mathbf{inl}_{A,M}$ and $\mathbf{inr}_{A,M}$ are two injective $\mathbf{k}$-linear maps. Explicitly, they are given by

$$\left(\mathbf{inl}_{A,M}(a) = (a,0) \qquad \text{for every } a \in A\right)$$

and

$$\left(\mathbf{inr}_{A,M}(m) = (0,m) \qquad \text{for every } m \in M\right).$$

The map $\mathbf{inl}_{A,M}$ is a $\mathbf{k}$-linear map $A \to A \oplus M$, therefore a $\mathbf{k}$-linear map $A \to \mathcal{R}_A(M)$ (since $\mathcal{R}_A(M) = A \oplus M$).

The map $\mathbf{inr}_{A,M}$ is a $\mathbf{k}$-linear map $M \to A \oplus M$, therefore a $\mathbf{k}$-linear map $M \to \mathcal{R}_A(M)$ (since $\mathcal{R}_A(M) = A \oplus M$).

The basic properties of direct sums of $\mathbf{k}$-modules (known from linear algebra) now show:

Proposition 1.34. Let A be a $\mathbf{k}$-algebra. Let M be an (A,A)-bimodule. Then,

$$\mathbf{projl}_{A,M} \circ \mathbf{inl}_{A,M} = \mathrm{id}_A; \tag{18}$$

$$\mathbf{projl}_{A,M} \circ \mathbf{inr}_{A,M} = 0; \tag{19}$$

$$\mathbf{projr}_{A,M} \circ \mathbf{inl}_{A,M} = 0; \tag{20}$$

$$\mathbf{projr}_{A,M} \circ \mathbf{inr}_{A,M} = \mathrm{id}_M; \tag{21}$$

$$\mathbf{inl}_{A,M} \circ \mathbf{projl}_{A,M} + \mathbf{inr}_{A,M} \circ \mathbf{projr}_{A,M} = \mathrm{id}_{\mathcal{R}_A(M)}. \tag{22}$$

Proposition 1.35. Let A be a $\mathbf{k}$-algebra. Let M be an (A,A)-bimodule.

(a) The map $\mathbf{projl}_{A,M} : \mathcal{R}_A(M) \to A$ is a $\mathbf{k}$-algebra homomorphism from $\mathcal{R}_A(M)$ to A.

(b) The map $\mathbf{inl}_{A,M} : A \to \mathcal{R}_A(M)$ is a $\mathbf{k}$-algebra homomorphism from A to $\mathcal{R}_A(M)$.

Proof of Proposition 1.35. This is really straightforward. □

Remark 1.36. Let A be a $\mathbf{k}$-algebra, and let M be an (A,A)-bimodule. An alternative realization of $\mathcal{R}_A(M)$ (i.e., a $\mathbf{k}$-algebra that is canonically isomorphic to $\mathcal{R}_A(M)$) can be constructed as follows:

Let $\begin{pmatrix} A & M \\ 0 & A \end{pmatrix}$ denote the $\mathbf{k}$-module of all 2×2-matrices of the form $\begin{pmatrix} a & m \\ 0 & b \end{pmatrix}$ with $a \in A$, $m \in M$ and $b \in A$. We can make this $\mathbf{k}$-module $\begin{pmatrix} A & M \\ 0 & A \end{pmatrix}$ into a ring by defining the product of two matrices in the usual way:

$$\begin{pmatrix} a & m \\ 0 & b \end{pmatrix} \begin{pmatrix} a' & m' \\ 0 & b' \end{pmatrix} = \begin{pmatrix} aa' & am' + mb' \\ 0 & bb' \end{pmatrix}.$$

The set of all $\begin{pmatrix} a & m \\ 0 & b \end{pmatrix} \in \begin{pmatrix} A & M \\ 0 & A \end{pmatrix}$ with $a = b$ is easily seen to be a $\mathbf{k}$-subalgebra of this $\mathbf{k}$-algebra. This $\mathbf{k}$-subalgebra is canonically isomorphic to $\mathcal{R}_A(M)$; the isomorphism sends the element $\begin{pmatrix} a & m \\ 0 & a \end{pmatrix}$ of this $\mathbf{k}$-subalgebra to the element (a, m) of $\mathcal{R}_A(M)$.

This realization of $\mathcal{R}_A(M)$ is particularly suited for generalization; but we shall not go further in this direction.

Definition 1.37. Let A be a $\mathbf{k}$-algebra. Let M be an (A, A)-bimodule. We let $\mathrm{Hom}_{\mathcal{R}}(A, M)$ denote the subset

$$\left\{ \varphi \in \mathrm{Hom}(A, \mathcal{R}_A(M)) \mid \mathbf{projl}_{A,M} \circ \varphi = \mathrm{id}_A \right\}$$

of $\mathrm{Hom}(A, \mathcal{R}_A(M))$.

(**Warning:** This subset $\mathrm{Hom}_{\mathcal{R}}(A, M)$ is not a $\mathbf{k}$-submodule.)

Proposition 1.38. Let A be a $\mathbf{k}$-algebra. Let M be an (A, A)-bimodule.

(a) For every $f \in \mathrm{Hom}(A, M)$, we have $\mathbf{inl}_{A,M} + \mathbf{inr}_{A,M} \circ f \in \mathrm{Hom}_{\mathcal{R}}(A, M)$.

(b) For every $F \in \mathrm{Hom}_{\mathcal{R}}(A, M)$, we have $\mathbf{projr}_{A,M} \circ F \in \mathrm{Hom}(A, M)$.

We delay the (simple) proof of this proposition until later, as we would like to first explain what its purpose is.

Definition 1.39. Let A be a $\mathbf{k}$-algebra. Let M be an (A, A)-bimodule.

We define a map $\mathrm{dth}_{A,M} : \mathrm{Hom}(A, M) \to \mathrm{Hom}_{\mathcal{R}}(A, M)$ by

$$\left(\mathrm{dth}_{A,M}(f) = \mathbf{inl}_{A,M} + \mathbf{inr}_{A,M} \circ f \qquad \text{for every } f \in \mathrm{Hom}(A, M)\right).$$

This map is well-defined, because for every $f \in \mathrm{Hom}(A, M)$, we have $\mathbf{inl}_{A,M} + \mathbf{inr}_{A,M} \circ f \in \mathrm{Hom}_{\mathcal{R}}(A, M)$ (according to Proposition 1.38 **(a)**).

We define a map $\mathrm{htd}_{A,M} : \mathrm{Hom}_{\mathcal{R}}(A, M) \to \mathrm{Hom}(A, M)$ by

$$\left(\mathrm{htd}_{A,M}(F) = \mathbf{projr}_{A,M} \circ F \qquad \text{for every } F \in \mathrm{Hom}_{\mathcal{R}}(A, M)\right).$$

This map is well-defined, because for every $F \in \operatorname{Hom}_{\mathcal{R}}(A, M)$, we have $\mathbf{projr}_{A,M} \circ F \in \operatorname{Hom}(A, M)$ (according to Proposition 1.38 **(b)**).

Theorem 1.40. Let A be a $\mathbf{k}$-algebra. Let M be an (A, A)-bimodule.
(a) If $f \in \operatorname{Hom}(A, M)$ and $a \in A$, then $(\operatorname{dth}_{A,M}(f))(a) = (a, f(a))$.
(b) The maps $\operatorname{dth}_{A,M}$ and $\operatorname{htd}_{A,M}$ are mutually inverse.
(c) Let $f \in \operatorname{Hom}(A, M)$. Then, f is a derivation if and only if $\operatorname{dth}_{A,M}(f)$ is a $\mathbf{k}$-algebra homomorphism from A to $\mathcal{R}_A(M)$.

Theorem 1.40 shows that the derivations from a $\mathbf{k}$-algebra A to an (A, A)-bimodule M are in 1-to-1 correspondence with $\mathbf{k}$-algebra homomorphisms from A to $\mathcal{R}_A(M)$ that lie in $\operatorname{Hom}_{\mathcal{R}}(A, M)$. (This correspondence is given by the maps $\operatorname{dth}_{A,M}$ and $\operatorname{htd}_{A,M}$; it also explains the names that we gave these maps[11].) This allows deriving properties of derivations from properties of $\mathbf{k}$-algebra homomorphisms. For instance, we could have used this tactic to derive Proposition 1.29 from the analogous property of $\mathbf{k}$-algebra homomorphisms.

As promised, let us now prove Proposition 1.38 and Theorem 1.40. These proofs are again rather straightforward.

Proof of Proposition 1.38. **(a)** Let $f \in \operatorname{Hom}(A, M)$. We must prove that $\mathbf{inl}_{A,M} + \mathbf{inr}_{A,M} \circ f \in \operatorname{Hom}_{\mathcal{R}}(A, M)$.

We have

$$\begin{aligned}&\mathbf{projl}_{A,M} \circ (\mathbf{inl}_{A,M} + \mathbf{inr}_{A,M} \circ f) \\ &= \underbrace{\mathbf{projl}_{A,M} \circ \mathbf{inl}_{A,M}}_{\substack{=\mathrm{id}_A \\ \text{(by (18))}}} + \underbrace{\mathbf{projl}_{A,M} \circ \mathbf{inr}_{A,M}}_{\substack{=0 \\ \text{(by (19))}}} \circ f = \mathrm{id}_A + \underbrace{0 \circ f}_{=0} = \mathrm{id}_A .\end{aligned}$$

Thus, $\mathbf{inl}_{A,M} + \mathbf{inr}_{A,M} \circ f$ is an element of $\operatorname{Hom}(A, \mathcal{R}_A(M))$ (because $\mathbf{inl}_{A,M} + \mathbf{inr}_{A,M} \circ f$ is $\mathbf{k}$-linear) and satisfies $\mathbf{projl}_{A,M} \circ (\mathbf{inl}_{A,M} + \mathbf{inr}_{A,M} \circ f) = \mathrm{id}_A$. In other words, $\mathbf{inl}_{A,M} + \mathbf{inr}_{A,M} \circ f$ is an element φ of $\operatorname{Hom}(A, \mathcal{R}_A(M))$ satisfying $\mathbf{projl}_{A,M} \circ \varphi = \mathrm{id}_A$. In other words,

$$\mathbf{inl}_{A,M} + \mathbf{inr}_{A,M} \circ f \in \left\{ \varphi \in \operatorname{Hom}(A, \mathcal{R}_A(M)) \mid \mathbf{projl}_{A,M} \circ \varphi = \mathrm{id}_A \right\} = \operatorname{Hom}_{\mathcal{R}}(A, M)$$

(since $\operatorname{Hom}_{\mathcal{R}}(A, M)$ was defined as $\left\{ \varphi \in \operatorname{Hom}(A, \mathcal{R}_A(M)) \mid \mathbf{projl}_{A,M} \circ \varphi = \mathrm{id}_A \right\}$). This proves Proposition 1.38 **(a)**.

(b) Let $F \in \operatorname{Hom}_{\mathcal{R}}(A, M)$. We must prove that $\mathbf{projr}_{A,M} \circ F \in \operatorname{Hom}(A, M)$.

But this follows immediately from $F \in \operatorname{Hom}_{\mathcal{R}}(A, M) \subseteq \operatorname{Hom}(A, \mathcal{R}_A(M))$. Thus, Proposition 1.38 **(b)** is proven.

□

[11] Namely, $\operatorname{dth}_{A,M}$ is short for "derivation to homomorphism", and $\operatorname{htd}_{A,M}$ for "homomorphism to derivation". Of course, the map $\operatorname{dth}_{A,M}$ is defined not only on derivations, and $\operatorname{htd}_{A,M}$ not only on homomorphisms, but this should motivate the names well enough.

Proof of Theorem 1.40. **(a)** Let $f \in \operatorname{Hom}(A, M)$ and $a \in A$. We must show that $(\operatorname{dth}_{A,M}(f))(a) = (a, f(a))$.

The definition of $\operatorname{dth}_{A,M}$ shows that $\operatorname{dth}_{A,M}(f) = \mathbf{inl}_{A,M} + \mathbf{inr}_{A,M} \circ f$. Hence,

$$\underbrace{(\operatorname{dth}_{A,M}(f))}_{=\mathbf{inl}_{A,M}+\mathbf{inr}_{A,M}\circ f}(a) = (\mathbf{inl}_{A,M} + \mathbf{inr}_{A,M} \circ f)(a)$$

$$= \underbrace{\mathbf{inl}_{A,M}(a)}_{\substack{=(a,0) \\ \text{(by the definition of } \mathbf{inl}_{A,M}\text{)}}} + \underbrace{(\mathbf{inr}_{A,M} \circ f)(a)}_{\substack{=\mathbf{inr}_{A,M}(f(a))=(0,f(a)) \\ \text{(by the definition of } \mathbf{inr}_{A,M}\text{)}}}$$

$$= (a, 0) + (0, f(a)) = \left(\underbrace{a+0}_{=a}, \underbrace{0+f(a)}_{=f(a)} \right) = (a, f(a)).$$

This proves Theorem 1.40 **(a)**.

(b) Every $F \in \operatorname{Hom}_{\mathcal{R}}(A, M)$ satisfies $(\operatorname{dth}_{A,M} \circ \operatorname{htd}_{A,M})(F) = \operatorname{id}_{\operatorname{Hom}_{\mathcal{R}}(A,M)}(F)$ [12]. In other words, we have $\operatorname{dth}_{A,M} \circ \operatorname{htd}_{A,M} = \operatorname{id}_{\operatorname{Hom}_{\mathcal{R}}(A,M)}$.

On the other hand, every $f \in \operatorname{Hom}(A, M)$ satisfies $(\operatorname{htd}_{A,M} \circ \operatorname{dth}_{A,M})(f) = \operatorname{id}_{\operatorname{Hom}(A,M)}(f)$ [13]. In other words, we have $\operatorname{htd}_{A,M} \circ \operatorname{dth}_{A,M} = \operatorname{id}_{\operatorname{Hom}(A,M)}$.

---

[12] *Proof.* Let $F \in \operatorname{Hom}_{\mathcal{R}}(A, M)$. We must prove that $(\operatorname{dth}_{A,M} \circ \operatorname{htd}_{A,M})(F) = \operatorname{id}_{A,\mathcal{R}_A(M)}(F)$.

We have

$$F \in \operatorname{Hom}_{\mathcal{R}}(A, M) = \left\{ \varphi \in \operatorname{Hom}(A, \mathcal{R}_A(M)) \mid \mathbf{projl}_{A,M} \circ \varphi = \operatorname{id}_A \right\}$$

(by the definition of $\operatorname{Hom}_{\mathcal{R}}(A, M)$). In other words, F is an element of $\operatorname{Hom}(A, \mathcal{R}_A(M))$ and satisfies $\mathbf{projl}_{A,M} \circ F = \operatorname{id}_A$.

On the other hand, $\operatorname{htd}_{A,M}(F) = \mathbf{projr}_{A,M} \circ F$ (by the definition of $\operatorname{htd}_{A,M}$). Now,

$$(\operatorname{dth}_{A,M} \circ \operatorname{htd}_{A,M})(F) = \operatorname{dth}_{A,M}\left(\underbrace{\operatorname{htd}_{A,M}(F)}_{=\mathbf{projr}_{A,M}\circ F} \right) = \operatorname{dth}_{A,M}\left(\mathbf{projr}_{A,M} \circ F \right)$$

$$= \underbrace{\mathbf{inl}_{A,M}}_{=\mathbf{inl}_{A,M}\circ \operatorname{id}_A} + \underbrace{\mathbf{inr}_{A,M} \circ \left(\mathbf{projr}_{A,M} \circ F \right)}_{=\mathbf{inr}_{A,M}\circ\mathbf{projr}_{A,M}\circ F} \qquad \text{(by the definition of } \operatorname{dth}_{A,M}\text{)}$$

$$= \mathbf{inl}_{A,M} \circ \underbrace{\operatorname{id}_A}_{=\mathbf{projl}_{A,M}\circ F} + \mathbf{inr}_{A,M} \circ \mathbf{projr}_{A,M} \circ F$$

$$= \mathbf{inl}_{A,M} \circ \mathbf{projl}_{A,M} \circ F + \mathbf{inr}_{A,M} \circ \mathbf{projr}_{A,M} \circ F$$

$$= \underbrace{\left(\mathbf{inl}_{A,M} \circ \mathbf{projl}_{A,M} + \mathbf{inr}_{A,M} \circ \mathbf{projr}_{A,M} \right)}_{\substack{=\operatorname{id}_{\mathcal{R}_A(M)} \\ \text{(by (22))}}} \circ F$$

$$= \operatorname{id}_{\mathcal{R}_A(M)} \circ F = F = \operatorname{id}_{\operatorname{Hom}_{\mathcal{R}}(A,M)}(F),$$

qed.

[13] *Proof.* Let $f \in \operatorname{Hom}(A, M)$. We must prove that $(\operatorname{htd}_{A,M} \circ \operatorname{dth}_{A,M})(f) = \operatorname{id}_{\operatorname{Hom}(A,M)}(f)$.

Combining $\operatorname{htd}_{A,M} \circ \operatorname{dth}_{A,M} = \operatorname{id}_{\operatorname{Hom}(A,M)}$ and $\operatorname{htd}_{A,M} \circ \operatorname{dth}_{A,M} = \operatorname{id}_{\operatorname{Hom}(A,M)}$, we conclude that the maps $\operatorname{dth}_{A,M}$ and $\operatorname{htd}_{A,M}$ are mutually inverse. Theorem 1.40 **(b)** is thus proven.

(c) Set $\varphi = \operatorname{dth}_{A,M}(f)$. Thus, $\varphi = \operatorname{dth}_{A,M}(f) \in \operatorname{Hom}_{\mathcal{R}}(A, M)$. For every $a \in A$, we have

$$\underbrace{\varphi}_{=\operatorname{dth}_{A,M}(f)}(a) = (\operatorname{dth}_{A,M}(f))(a) = (a, f(a)). \tag{23}$$

We must prove that f is a derivation if and only if $\operatorname{dth}_{A,M}(f)$ is a $\mathbf{k}$-algebra homomorphism from A to $\mathcal{R}_A(M)$. In other words, we must prove that f is a derivation if and only if φ is a $\mathbf{k}$-algebra homomorphism from A to $\mathcal{R}_A(M)$ (because $\varphi = \operatorname{dth}_{A,M}(f)$). In other words, we must prove the following two statements:

Statement 1: If f is a derivation, then φ is a $\mathbf{k}$-algebra homomorphism from A to $\mathcal{R}_A(M)$.

Statement 2: If φ is a $\mathbf{k}$-algebra homomorphism from A to $\mathcal{R}_A(M)$, then f is a derivation.

Proof of Statement 1: Assume that f is a derivation. We must prove that φ is a $\mathbf{k}$-algebra homomorphism from A to $\mathcal{R}_A(M)$.

Recall that $(1, 0)$ is the unity of the $\mathbf{k}$-algebra $\mathcal{R}_A(M)$. In other words, $(1, 0) = 1_{\mathcal{R}_A(M)}$.

Theorem 1.12 shows that $f(1) = 0$. Now, (23) (applied to $a = 1$) shows that

$$\varphi(1) = \left(1, \underbrace{f(1)}_{=0}\right) = (1, 0) = 1_{\mathcal{R}_A(M)}.$$

Also, $\varphi = \operatorname{dth}_{A,M}(f) \in \operatorname{Hom}_{\mathcal{R}}(A, M) \subseteq \operatorname{Hom}(A, \mathcal{R}_A(M))$. Thus, the map φ is $\mathbf{k}$-linear.

The definition of $\operatorname{dth}_{A,M}$ shows that $\operatorname{dth}_{A,M}(f) = \mathbf{inl}_{A,M} + \mathbf{inr}_{A,M} \circ f$. Now,

$$\begin{aligned}
(\operatorname{htd}_{A,M} \circ \operatorname{dth}_{A,M})(f) &= \operatorname{htd}_{A,M}\left(\underbrace{\operatorname{dth}_{A,M}(f)}_{=\mathbf{inl}_{A,M}+\mathbf{inr}_{A,M}\circ f}\right) = \operatorname{htd}_{A,M}(\mathbf{inl}_{A,M} + \mathbf{inr}_{A,M} \circ f) \\
&= \mathbf{projr}_{A,M} \circ (\mathbf{inl}_{A,M} + \mathbf{inr}_{A,M} \circ f) \qquad \text{(by the definition of } \operatorname{htd}_{A,M}) \\
&= \mathbf{projr}_{A,M} \circ \mathbf{inl}_{A,M} + \underbrace{\mathbf{projr}_{A,M} \circ (\mathbf{inr}_{A,M} \circ f)}_{=\mathbf{projr}_{A,M}\circ\mathbf{inr}_{A,M}\circ f} \\
&\qquad \left(\begin{array}{c}\text{since composition of } \mathbf{k}\text{-linear maps is } \mathbf{k}\text{-bilinear (and since} \\ \text{the three maps } \mathbf{projr}_{A,M},\ \mathbf{inl}_{A,M} \text{ and } \mathbf{inr}_{A,M} \circ f \text{ are } \mathbf{k}\text{-linear)}\end{array}\right) \\
&= \underbrace{\mathbf{projr}_{A,M} \circ \mathbf{inl}_{A,M}}_{\substack{=0 \\ \text{(by (20))}}} + \underbrace{\mathbf{projr}_{A,M} \circ \mathbf{inr}_{A,M}}_{\substack{=\operatorname{id}_M \\ \text{(by (21))}}} \circ f \\
&= \operatorname{id}_M \circ f = f = \operatorname{id}_{\operatorname{Hom}(A,M)}(f),
\end{aligned}$$

qed.

Furthermore, $\varphi(ab) = \varphi(a)\,\varphi(b)$ for every $a \in A$ and $b \in A$ [14]. This (combined with the fact that φ is a $\mathbf{k}$-linear map and the fact that $\varphi(1) = 1_{\mathcal{R}_A(M)}$) shows that φ is a $\mathbf{k}$-algebra homomorphism from A to $\mathcal{R}_A(M)$. Statement 1 is thus proven.

Proof of Statement 2: Assume that φ is a $\mathbf{k}$-algebra homomorphism from A to $\mathcal{R}_A(M)$. We must prove that f is a derivation.

We have $f \in \operatorname{Hom}(A, M)$. In other words, f is a $\mathbf{k}$-linear map from A to M.

Now, let $a \in A$ and $b \in A$. We shall show that $f(ab) = af(b) + f(a)\,b$.

Indeed, we have $\varphi(a) = (a, f(a))$ (by (23)) and $\varphi(b) = (b, f(b))$ (by (23), applied to b instead of a). Furthermore, (23) (applied to ab instead of a) shows that $\varphi(ab) = (ab, f(ab))$. Thus,

$$\begin{aligned}(ab, f(ab)) = \varphi(ab) &= \underbrace{\varphi(a)}_{=(a,f(a))} \cdot \underbrace{\varphi(b)}_{=(b,f(b))} && \text{(since } \varphi \text{ is a } \mathbf{k}\text{-algebra homomorphism)} \\ &= (a, f(a)) \cdot (b, f(b)) \\ &= (ab, af(b) + f(a)\,b) && \left(\begin{array}{c}\text{by (14), applied to } (a, f(a)) \text{ and } (b, f(b)) \\ \text{instead of } (a, p) \text{ and } (b, q)\end{array}\right).\end{aligned}$$

In other words, $ab = ab$ and $f(ab) = af(b) + f(a)\,b$.

Now, let us forget that we fixed a and b. We thus have proven that $f(ab) = af(b) + f(a)\,b$ for every $a \in A$ and $b \in A$. In other words, f is a derivation. This proves Statement 2.

Now, Statement 1 and Statement 2 are both proven. As we have said above, this concludes the proof of Theorem 1.40 **(c)**. □

Theorem 1.40 is a classical folklore result. In the particular case when A is a commutative $\mathbf{k}$-algebra and M is a symmetric (A, A)-bimodule, it is explicitly stated in [Becker14] (where $\mathcal{R}_A(M)$ is denoted by $M \rtimes A$, and called a *square zero extension*[15]), but it was widely used before.

[14] *Proof.* Let $a \in A$ and $b \in A$. Then, $\varphi(a) = (a, f(a))$ (by (23)) and $\varphi(b) = (b, f(b))$ (by (23), applied to b instead of a). Furthermore, (23) (applied to ab instead of a) shows that

$$\varphi(ab) = \left(ab, \underbrace{f(ab)}_{\substack{=af(b)+f(a)b \\ \text{(by (5))}}}\right) = (ab, af(b) + f(a)\,b).$$

Comparing this with

$$\begin{aligned}\varphi(a)\,\varphi(b) &= \underbrace{\varphi(a)}_{=(a,f(a))} \cdot \underbrace{\varphi(b)}_{=(b,f(b))} = (a, f(a)) \cdot (b, f(b)) \\ &= (ab, af(b) + f(a)\,b) && \left(\begin{array}{c}\text{by (14), applied to } (a, f(a)) \text{ and } (b, f(b)) \\ \text{instead of } (a, p) \text{ and } (b, q)\end{array}\right),\end{aligned}$$

we obtain $\varphi(ab) = \varphi(a)\,\varphi(b)$. Qed.

[15] This name has a slightly more general meaning in other sources.

The following is a sufficient criterion for $\mathcal{R}_A(M)$ to be commutative[16]:

Proposition 1.41. Let A be a commutative $\mathbf{k}$-algebra. Let M be a symmetric (A, A)-bimodule. Then, the $\mathbf{k}$-algebra $\mathcal{R}_A(M)$ is commutative.

Proof of Proposition 1.41. Let $x \in \mathcal{R}_A(M)$ and $y \in \mathcal{R}_A(M)$.

We have $x \in \mathcal{R}_A(M)$. Then, $x \in \mathcal{R}_A(M) = A \oplus M$. Thus, we can write x in the form $x = (a, p)$ for some $a \in A$ and $p \in M$. Consider these a and p.

We have $y \in \mathcal{R}_A(M)$. Then, $y \in \mathcal{R}_A(M) = A \oplus M$. Thus, we can write y in the form $y = (b, q)$ for some $b \in A$ and $q \in M$. Consider these b and q.

Since the (A, A)-bimodule M is symmetric, we have $aq = qa$ and $bp = pb$. Finally, $ab = ba$ (since A is commutative).

Now,

$$xy = \underbrace{x}_{=(a,p)} \cdot \underbrace{y}_{=(b,q)} = (a,p)\cdot(b,q) = \left(\underbrace{ab}_{=ba}, \underbrace{aq}_{=qa} + \underbrace{pb}_{=bp}\right) \qquad \text{(by (14))}$$

$$= \left(ba, \underbrace{qa + bp}_{=bp+qa}\right) = (ba, bp + qa).$$

Comparing this with

$$yx = \underbrace{y}_{=(b,q)} \cdot \underbrace{x}_{=(a,p)} = (b,q)\cdot(a,p) = (ba, bp+qa)$$
$$\left(\begin{array}{c}\text{by (14), applied to } (b,q) \text{ and } (a,p)\\ \text{instead of } (a,p) \text{ and } (b,q)\end{array}\right),$$

we obtain $xy = yx$.

Let us now forget that we fixed x and y. We thus have proven that $xy = yx$ for every $x \in \mathcal{R}_A(M)$ and $y \in \mathcal{R}_A(M)$. In other words, the $\mathbf{k}$-algebra $\mathcal{R}_A(M)$ is commutative. □

1.5. Compositions and tensor products

The following two (almost trivial) facts show how derivations can be obtained by composing derivations with other maps:

Proposition 1.42. Let A and B be two $\mathbf{k}$-algebras. Let $f : A \to B$ be a $\mathbf{k}$-algebra homomorphism. Let M be a (B, B)-bimodule. Let $d : B \to M$ be a derivation.

We consider the (B, B)-bimodule M as an (A, A)-bimodule (via the $\mathbf{k}$-algebra homomorphism $f : A \to B$). Then, $d \circ f : A \to M$ is a derivation.

[16] Actually, it is a necessary criterion as well (i.e., if $\mathcal{R}_A(M)$ is commutative, then A is commutative and M is symmetric), but we will not have use for this fact.

Proof of Proposition 1.42. We know that d is a derivation. In other words, we have

$$\left(d\left(ab\right) = ad\left(b\right) + d\left(a\right)b \qquad \text{for every } a \in B \text{ and } b \in B\right). \tag{24}$$

Now, let $a \in A$ and $b \in A$. Then, $f\left(ab\right) = f\left(a\right)f\left(b\right)$ (since f is a $\mathbf{k}$-algebra homomorphism). Now,

$$\left(d \circ f\right)\left(ab\right) = d\left(\underbrace{f\left(ab\right)}_{=f(a)f(b)}\right) = d\left(f\left(a\right)f\left(b\right)\right) = f\left(a\right)d\left(f\left(b\right)\right) + d\left(f\left(a\right)\right)f\left(b\right) \tag{25}$$

(by (24), applied to $f\left(a\right)$ and $f\left(b\right)$ instead of a and b).

On the other hand, recall that the (B, B)-module M was made into an (A, A)-bimodule via the $\mathbf{k}$-algebra homomorphism $f : A \to B$. Hence, $xu = f\left(x\right)u$ for every $x \in A$ and $u \in M$. Applying this to $x = a$ and $u = d\left(f\left(b\right)\right)$, we obtain $ad\left(f\left(b\right)\right) = f\left(a\right)d\left(f\left(b\right)\right)$. Similarly, we obtain $d\left(f\left(a\right)\right)b = d\left(f\left(a\right)\right)f\left(b\right)$. Now,

$$\begin{aligned} & a\underbrace{\left(d \circ f\right)\left(b\right)}_{=d(f(b))} + \underbrace{\left(d \circ f\right)\left(a\right)}_{=d(f(a))}b \\ &= \underbrace{ad\left(f\left(b\right)\right)}_{=f(a)d(f(b))} + \underbrace{d\left(f\left(a\right)\right)b}_{=d(f(a))f(b)} = f\left(a\right)d\left(f\left(b\right)\right) + d\left(f\left(a\right)\right)f\left(b\right). \end{aligned}$$

Comparing this with (25), we obtain $\left(d \circ f\right)\left(ab\right) = a\left(d \circ f\right)\left(b\right) + \left(d \circ f\right)\left(a\right)b$.

Let us now forget that we fixed a and b. We thus have shown that $\left(d \circ f\right)\left(ab\right) = a\left(d \circ f\right)\left(b\right) + \left(d \circ f\right)\left(a\right)b$ for all $a \in A$ and $b \in A$. In other words, $d \circ f$ is a derivation (by the definition of a "derivation", since $d \circ f$ is a $\mathbf{k}$-linear map). This proves Proposition 1.42. □

Proposition 1.43. Let A be a $\mathbf{k}$-algebra. Let M and N be two (A, A)-bimodules. Let $f : M \to N$ be an (A, A)-bimodule homomorphism. Let $d : A \to M$ be a derivation. Then, $f \circ d : A \to N$ is a derivation.

Proof of Proposition 1.43. This is even more straightforward than the proof of Proposition 1.42, and thus left to the reader. □

A composition of two derivations is usually not a derivation (although we will see some properties of composition powers of derivations in the next section).

A tensor product of two derivations is usually not a derivation either. However, here is an example of how derivations interact with tensor products:

Proposition 1.44. Let A be a $\mathbf{k}$-algebra. Let M be an (A, A)-bimodule. Let $f : A \to M$ be a derivation. Let B be a $\mathbf{k}$-algebra. Clearly, the tensor product $M \otimes B$ of the (A, A)-bimodule M with the (B, B)-bimodule B is an $(A \otimes B, A \otimes B)$-bimodule. Similarly, $B \otimes M$ is an $(B \otimes A, B \otimes A)$-bimodule.

(a) The map $f \otimes \mathrm{id}_B : A \otimes B \to M \otimes B$ is a derivation.

(b) The map $\mathrm{id}_B \otimes f : B \otimes A \to B \otimes M$ is a derivation.

Proof of Proposition 1.44. **(a)** The map $f \otimes \mathrm{id}_B$ is clearly $\mathbf{k}$-linear (since it is the tensor product of the two $\mathbf{k}$-linear maps f and id_B).

Now, we are going to prove that

$$(f \otimes \mathrm{id}_B)(ab) = a\,(f \otimes \mathrm{id}_B)(b) + (f \otimes \mathrm{id}_B)(a)\,b \tag{26}$$

for every $a \in A \otimes B$ and $b \in A \otimes B$.

Proof of (26): Let $a \in A \otimes B$ and $b \in A \otimes B$. We need to prove the equality (26). But this equality is $\mathbf{k}$-linear in each of a and b (because the multiplication of the $\mathbf{k}$-algebra $A \otimes B$ is $\mathbf{k}$-bilinear, because the left and right actions of $A \otimes B$ on $M \otimes B$ are $\mathbf{k}$-bilinear, and because the map $f \otimes \mathrm{id}_B$ is $\mathbf{k}$-linear). Hence, we can WLOG assume that a and b are pure tensors (since the $\mathbf{k}$-module $A \otimes B$ is spanned by pure tensors). Assume this.

We know that a is a pure tensor. In other words, $a = x \otimes y$ for some $x \in A$ and $y \in B$. Consider these x and y.

We know that b is a pure tensor. In other words, $b = z \otimes w$ for some $z \in A$ and $w \in B$. Consider these z and w.

We have $\underbrace{a}_{=x\otimes y}\,\underbrace{b}_{=z\otimes w} = (x \otimes y)(z \otimes w) = xz \otimes yw$. Hence,

$$\begin{aligned}(f \otimes \mathrm{id}_B)\left(\underbrace{ab}_{=xz\otimes yw}\right) &= (f \otimes \mathrm{id}_B)(xz \otimes yw) = \underbrace{f(xz)}_{\substack{=xf(z)+f(x)z \\ \text{(since } f \text{ is a derivation)}}} \otimes \underbrace{\mathrm{id}_B(yw)}_{=yw} \\ &= (xf(z) + f(x)z) \otimes yw = xf(z) \otimes yw + f(x)z \otimes yw.\end{aligned}$$

Compared with

$$\begin{aligned}&\underbrace{a}_{=x\otimes y}(f \otimes \mathrm{id}_B)\left(\underbrace{b}_{=z\otimes w}\right) + (f \otimes \mathrm{id}_B)\left(\underbrace{a}_{=x\otimes y}\right)\underbrace{b}_{=z\otimes w} \\ &= (x \otimes y) \cdot \underbrace{(f \otimes \mathrm{id}_B)(z \otimes w)}_{=f(z)\otimes \mathrm{id}_B(w)} + \underbrace{(f \otimes \mathrm{id}_B)(x \otimes y)}_{=f(x)\otimes \mathrm{id}_B(y)} \cdot (z \otimes w) \\ &= (x \otimes y) \cdot \left(f(z) \otimes \underbrace{\mathrm{id}_B(w)}_{=w}\right) + \left(f(x) \otimes \underbrace{\mathrm{id}_B(y)}_{=y}\right) \cdot (z \otimes w) \\ &= \underbrace{(x \otimes y) \cdot (f(z) \otimes w)}_{=xf(z)\otimes yw} + \underbrace{(f(x) \otimes y) \cdot (z \otimes w)}_{=f(x)z\otimes yw} = xf(z) \otimes yw + f(x)z \otimes yw,\end{aligned}$$

this shows that $(f \otimes \mathrm{id}_B)(ab) = a\,(f \otimes \mathrm{id}_B)(b) + (f \otimes \mathrm{id}_B)(a)\,b$. Thus, (26) is proven.

We thus have shown that $(f \otimes \mathrm{id}_B)(ab) = a\,(f \otimes \mathrm{id}_B)(b) + (f \otimes \mathrm{id}_B)(a)\,b$ for every $a \in A \otimes B$ and $b \in A \otimes B$. Since $f \otimes \mathrm{id}_B$ is a $\mathbf{k}$-linear map, this shows

that $f \otimes \mathrm{id}_B$ is a derivation (by the definition of a "derivation"). This proves Proposition 1.44 **(a)**.

(b) The proof of Proposition 1.44 **(b)** is analogous to the proof of Proposition 1.44 **(a)** (indeed, Proposition 1.44 **(b)** differs from Proposition 1.44 **(a)** only in the order of the tensorands), and thus is left to the reader. (Alternatively, Proposition 1.44 **(b)** can be deduced from Proposition 1.44 **(a)** via the isomorphisms $A \otimes B \cong B \otimes A$ and $M \otimes B \cong B \otimes M$.) □

As an example of how Propositions 1.42 and 1.43 can be used, let us combine them with Proposition 1.29:

Corollary 1.45. Let A and B be two $\mathbf{k}$-algebras. Let $f : A \to B$ be a $\mathbf{k}$-algebra homomorphism. Let $d : A \to A$ and $e : B \to B$ be two derivations. Let S be a subset of A which generates A as a $\mathbf{k}$-algebra. Assume that $(f \circ d) \mid_S = (e \circ f) \mid_S$. Then, $f \circ d = e \circ f$.

Proof of Corollary 1.45. We make the (B, B)-bimodule B into an (A, A)-bimodule via the $\mathbf{k}$-algebra homomorphism $f : A \to B$. Proposition 1.42 (applied to B and e instead of M and d) shows that $e \circ f : A \to B$ is a derivation.

It is easy to see (using the definition of the (A, A)-bimodule structure on B and the fact that $f : A \to B$ is a $\mathbf{k}$-algebra homomorphism) that $f : A \to B$ is an (A, A)-bimodule homomorphism.

Hence, Proposition 1.43 (applied to $M = A$ and $N = B$) shows that $f \circ d : A \to B$ is a derivation. Now, Proposition 1.29 (applied to B, $f \circ d$ and $e \circ f$ instead of M, d and e) shows that $f \circ d = e \circ f$. This proves Corollary 1.45. □

1.6. Composition powers of derivations

If A is a $\mathbf{k}$-algebra and $f : A \to A$ is a derivation, then f^n (for $n \in \mathbb{N}$) is usually not a derivation. However, this does not mean that nothing can be said about $f^n(ab)$. The following result generalizes the Leibniz law:

Proposition 1.46. Let A be a $\mathbf{k}$-algebra. Let $f : A \to A$ be a derivation. Let $n \in \mathbb{N}$. Let $a \in A$ and $b \in A$. Then,

$$f^n(ab) = \sum_{k=0}^{n} \binom{n}{k} f^k(a) f^{n-k}(b).$$

Proof of Proposition 1.46. We shall prove Proposition 1.46 by induction over n:

Induction base: Proposition 1.46 is easily verified in the case when $n = 0$ (because $f^0 = \mathrm{id}$). This completes the induction base.

Induction step: Let N be a positive integer. Assume that Proposition 1.46 holds in the case when $n = N - 1$. We now must prove that Proposition 1.46 holds in the case when $n = N$.

Recall the recurrence relation for the binomial coefficients. It says that

$$\binom{a}{b} = \binom{a-1}{b} + \binom{a-1}{b-1} \tag{27}$$

for every $a \in \mathbb{Z}$ and every positive integer b. Hence, every positive integer k satisfies

$$\binom{N}{k} = \binom{N-1}{k} + \binom{N-1}{k-1} \tag{28}$$

(by (27), applied to $a = N$ and $b = k$).

We have assumed that Proposition 1.46 holds in the case when $n = N-1$. In other words, we have

$$f^{N-1}(ab) = \sum_{k=0}^{N-1} \binom{N-1}{k} f^k(a)\, f^{(N-1)-k}(b).$$

Now,

$$\begin{aligned}
&\underbrace{f^N}_{=f\circ f^{N-1}}(ab) \\
&= \left(f \circ f^{N-1}\right)(ab) = f\left(\underbrace{f^{N-1}(ab)}_{=\sum_{k=0}^{N-1}\binom{N-1}{k} f^k(a) f^{(N-1)-k}(b)}\right) \\
&= f\left(\sum_{k=0}^{N-1} \binom{N-1}{k} f^k(a)\, f^{(N-1)-k}(b)\right) \\
&= \sum_{k=0}^{N-1} \binom{N-1}{k} \underbrace{f\left(f^k(a)\, f^{(N-1)-k}(b)\right)}_{\substack{=f^k(a) f\left(f^{(N-1)-k}(b)\right) + f\left(f^k(a)\right) f^{(N-1)-k}(b) \\ \text{(since } f \text{ is a derivation)}}} \qquad \text{(since } f \text{ is a } \mathbf{k}\text{-linear map)} \\
&= \sum_{k=0}^{N-1} \binom{N-1}{k} \left(f^k(a)\, f\left(f^{(N-1)-k}(b)\right) + f\left(f^k(a)\right) f^{(N-1)-k}(b)\right) \\
&= \sum_{k=0}^{N-1} \binom{N-1}{k} f^k(a)\, f\left(f^{(N-1)-k}(b)\right) \\
&\qquad + \sum_{k=0}^{N-1} \binom{N-1}{k} f\left(f^k(a)\right) f^{(N-1)-k}(b).
\end{aligned} \tag{29}$$

But

$$\begin{aligned}
&\sum_{k=0}^{N-1} \binom{N-1}{k} f^k(a) \underbrace{f\left(f^{(N-1)-k}(b)\right)}_{\substack{=\left(f\circ f^{(N-1)-k}\right)(b) \\ =\left(f\circ f^{(N-k)-1}\right)(b) \\ =f^{N-k}(b)}} \\
&= \sum_{k=0}^{N-1} \binom{N-1}{k} f^k(a) f^{N-k}(b) \\
&= \underbrace{\binom{N-1}{0}}_{=1} f^0(a) f^{N-0}(b) + \sum_{k=1}^{N-1} \binom{N-1}{k} f^k(a) f^{N-k}(b) \\
&= f^0(a) f^{N-0}(b) + \sum_{k=1}^{N-1} \binom{N-1}{k} f^k(a) f^{N-k}(b)
\end{aligned}$$

and

$$\begin{aligned}
&\sum_{k=0}^{N-1} \binom{N-1}{k} f\left(f^k(a)\right) f^{(N-1)-k}(b) \\
&= \sum_{k=1}^{N} \binom{N-1}{k-1} \underbrace{f\left(f^{k-1}(a)\right)}_{\substack{=\left(f\circ f^{k-1}\right)(a) \\ =f^k(a)}} \underbrace{f^{(N-1)-(k-1)}}_{=f^{N-k}}(b) \\
&\qquad \text{(here, we have substituted } k-1 \text{ for } k \text{ in the sum)} \\
&= \sum_{k=1}^{N} \binom{N-1}{k-1} f^k(a) f^{N-k}(b) \\
&= \sum_{k=1}^{N-1} \binom{N-1}{k-1} f^k(a) f^{N-k}(b) + \underbrace{\binom{N-1}{N-1}}_{=1} f^N(a) f^{N-N}(b) \\
&\qquad \text{(here, we have split off the addend for } k=N \text{ from the sum)} \\
&= \sum_{k=1}^{N-1} \binom{N-1}{k-1} f^k(a) f^{N-k}(b) + f^N(a) f^{N-N}(b).
\end{aligned}$$

Hence, (29) becomes

$$
\begin{aligned}
&f^N(ab)\\
&= \underbrace{\sum_{k=0}^{N-1}\binom{N-1}{k} f^k(a)\, f\left(f^{(N-1)-k}(b)\right)}_{=f^0(a)f^{N-0}(b)+\sum_{k=1}^{N-1}\binom{N-1}{k}f^k(a)f^{N-k}(b)}\\
&\qquad + \underbrace{\sum_{k=0}^{N-1}\binom{N-1}{k} f\left(f^k(a)\right) f^{(N-1)-k}(b)}_{=\sum_{k=1}^{N-1}\binom{N-1}{k-1}f^k(a)f^{N-k}(b)+f^N(a)f^{N-N}(b)}\\
&= f^0(a)\, f^{N-0}(b) + \sum_{k=1}^{N-1}\binom{N-1}{k} f^k(a)\, f^{N-k}(b)\\
&\qquad + \sum_{k=1}^{N-1}\binom{N-1}{k-1} f^k(a)\, f^{N-k}(b) + f^N(a)\, f^{N-N}(b)\\
&= f^0(a)\, f^{N-0}(b)\\
&\qquad + \underbrace{\sum_{k=1}^{N-1}\binom{N-1}{k} f^k(a)\, f^{N-k}(b) + \sum_{k=1}^{N-1}\binom{N-1}{k-1} f^k(a)\, f^{N-k}(b)}_{=\sum_{k=1}^{N-1}\left(\binom{N-1}{k}+\binom{N-1}{k-1}\right)f^k(a)f^{N-k}(b)}\\
&\qquad + f^N(a)\, f^{N-N}(b)\\
&= f^0(a)\, f^{N-0}(b) + \sum_{k=1}^{N-1}\underbrace{\left(\binom{N-1}{k}+\binom{N-1}{k-1}\right)}_{=\binom{N}{k}\ \text{(by (28))}} f^k(a)\, f^{N-k}(b)\\
&\qquad + f^N(a)\, f^{N-N}(b)\\
&= f^0(a)\, f^{N-0}(b) + \sum_{k=1}^{N-1}\binom{N}{k} f^k(a)\, f^{N-k}(b) + f^N(a)\, f^{N-N}(b).
\end{aligned}
$$

Comparing this with

$$\begin{aligned}
&\sum_{k=0}^{N} \binom{N}{k} f^k(a)\, f^{N-k}(b) \\
&= \underbrace{\binom{N}{0}}_{=1} f^0(a)\, f^{N-0}(b) + \sum_{k=1}^{N-1} \binom{N}{k} f^k(a)\, f^{N-k}(b) + \underbrace{\binom{N}{N}}_{=1} f^N(a)\, f^{N-N}(b) \\
&\qquad \text{(here, we have split off the addends for } k=0 \text{ and for } k=N \text{ from the sum)} \\
&= f^0(a)\, f^{N-0}(b) + \sum_{k=1}^{N-1} \binom{N}{k} f^k(a)\, f^{N-k}(b) + f^N(a)\, f^{N-N}(b),
\end{aligned}$$

we obtain $f^N(ab) = \sum_{k=0}^{N} \binom{N}{k} f^k(a)\, f^{N-k}(b)$. In other words, Proposition 1.46 holds in the case when $n = N$. This completes the induction step. Thus, the induction proof of Proposition 1.46 is complete.

□

As a consequence of Proposition 1.46, we can easily see the following:

Corollary 1.47. Let A be a $\mathbf{k}$-algebra. Let p be a prime number such that $p1_A = 0$. Let $f : A \to A$ be a derivation. Then, $f^p : A \to A$ is a derivation.

Proof of Corollary 1.47. A well-known property of binomial coefficients[17] states that if q is a prime number, then every $k \in \{1, 2, \ldots, q-1\}$ satisfies $q \mid \binom{q}{k}$. Applying this to $q = p$, we obtain the following: Every $k \in \{1, 2, \ldots, p-1\}$ satisfies

$$p \mid \binom{p}{k}. \tag{30}$$

Hence, every $k \in \{1, 2, \ldots, p-1\}$ and every $a \in A$ satisfy

$$\binom{p}{k} a = 0 \tag{31}$$

[18].

[17] See, for example, [Grinbe16, Corollary 5.6] for a proof of this property (although what we call q is denoted by p in [Grinbe16, Corollary 5.6]).

[18] *Proof of (31):* Let $k \in \{1, 2, \ldots, p-1\}$ and $a \in A$. Then, (30) yields $p \mid \binom{p}{k}$. In other words, there exists an integer z such that $\binom{p}{k} = zp$. Consider this z. Now, $\underbrace{\binom{p}{k}}_{=zp} \underbrace{a}_{=1_A \cdot a} = z \underbrace{p1_A}_{=0} \cdot a = 0$. This proves (31).

Now, for every $a \in A$ and $b \in A$, we have

$$f^p(ab) = \sum_{k=0}^{p} \binom{p}{k} f^k(a) f^{p-k}(b) \qquad \text{(by Proposition 1.46, applied to } n = p\text{)}$$

$$= \underbrace{\binom{p}{0}}_{=1} \underbrace{f^0}_{=\mathrm{id}}(a) \underbrace{f^{p-0}}_{=f^p}(b) + \sum_{k=1}^{p-1} \underbrace{\binom{p}{k} f^k(a)}_{\substack{=0 \\ \text{(by (31), applied to} \\ f^k(a) \text{ instead of } a\text{)}}} f^{p-k}(b) + \underbrace{\binom{p}{p}}_{=1} f^p(a) \underbrace{f^{p-p}}_{=f^0=\mathrm{id}}(b)$$

(here, we have split off the addends for $k = 0$ and for $k = p$ from the sum)

$$= \underbrace{\mathrm{id}(a)}_{=a} f^p(b) + \underbrace{\sum_{k=1}^{p-1} 0 f^{p-k}(b)}_{=0} + f^p(a) \underbrace{\mathrm{id}(b)}_{=b} = a f^p(b) + f^p(a) b.$$

Hence, f^p is a derivation. This proves Corollary 1.47. □

Example 1.48. Let A be a $\mathbf{k}$-algebra. Let p be a prime number. Assume that $p1_A = 0$. Let $x \in A$. Then, Example 1.9 (applied to x instead of p) defines a derivation $\mathrm{ad}_x : A \to A$ which sends every $a \in A$ to $xa - ax$. Corollary 1.47 thus shows that $(\mathrm{ad}_x)^p$ is a derivation as well. It can be shown that it is actually an inner derivation: $(\mathrm{ad}_x)^p = \mathrm{ad}_{x^p}$. (Hint for the proof: Define two $\mathbf{k}$-linear maps $L_x : A \to A$, $a \mapsto xa$ and $R_x : A \to A$, $a \mapsto ax$, and prove that these two maps commute. Thus, the binomial formula can be used to compute $(L_x - R_x)^p$. Finally observe that $\mathrm{ad}_x = L_x - R_x$.)

1.7. A product formula for the Wronskian

A consequence of Proposition 1.46 is a certain property of the Wronskian determinant, which we shall now define:[19]

Definition 1.49. Let A be a commutative $\mathbf{k}$-algebra. Let $f : A \to A$ be a derivation. Let $n \in \mathbb{N}$. Let $a_1, a_2, \ldots, a_n \in A$. Then, the *$f$-Wronskian* of $a_1, a_2, \ldots, a_n$ is defined to be the determinant

$$\det\left(\left(f^{j-1}(a_i)\right)_{1 \le i \le n,\ 1 \le j \le n}\right)$$
$$= \det\begin{pmatrix} f^0(a_1) & f^1(a_1) & \cdots & f^{n-1}(a_1) \\ f^0(a_2) & f^1(a_2) & \cdots & f^{n-1}(a_2) \\ \vdots & \vdots & \ddots & \vdots \\ f^0(a_n) & f^1(a_n) & \cdots & f^{n-1}(a_n) \end{pmatrix} \in A.$$

This f-Wronskian is denoted by $W_f(a_1, a_2, \ldots, a_n)$.

[19] We are using the notation $(a_{i,j})_{1 \le i \le n,\ 1 \le j \le m}$ for the $n \times m$-matrix whose (i,j)-th entry is $a_{i,j}$.

Theorem 1.50. Let A be a commutative $\mathbf{k}$-algebra. Let $f : A \to A$ be a derivation. Let $n \in \mathbb{N}$. Let $a_1, a_2, \ldots, a_n \in A$. Let $a \in A$. Then,

$$W_f\left(aa_1, aa_2, \ldots, aa_n\right) = a^n W_f\left(a_1, a_2, \ldots, a_n\right).$$

Proof of Theorem 1.50. The definition of the f-Wronskian $W_f\left(a_1, a_2, \ldots, a_n\right)$ yields

$$W_f\left(a_1, a_2, \ldots, a_n\right) = \det\left(\left(f^{j-1}\left(a_i\right)\right)_{1 \leq i \leq n,\ 1 \leq j \leq n}\right). \tag{32}$$

The same argument (applied to aa_i instead of a_i) yields

$$W_f\left(aa_1, aa_2, \ldots, aa_n\right) = \det\left(\left(f^{j-1}\left(aa_i\right)\right)_{1 \leq i \leq n,\ 1 \leq j \leq n}\right). \tag{33}$$

Each $(i, j) \in \{1, 2, \ldots, n\} \times \{1, 2, \ldots, n\}$ satisfies $j - 1 \in \mathbb{N}$ (since $j \geq 1$) and thus

$$\begin{aligned} f^{j-1}\left(aa_i\right) &= \sum_{k=0}^{j-1} \binom{j-1}{k} f^k\left(a\right) f^{j-1-k}\left(a_i\right) \\ &\qquad \text{(by Proposition 1.46, applied to } n = j-1 \text{ and } b = a_i\text{)} \\ &= \sum_{k=1}^{j} \binom{j-1}{k-1} f^{k-1}\left(a\right) \underbrace{f^{j-1-(k-1)}}_{\substack{=f^{j-k} \\ \text{(since } j-1-(k-1)=j-1\text{)}}} \left(a_i\right) \\ &\qquad \text{(here, we have substituted } k-1 \text{ for } k \text{ in the sum)} \\ &= \sum_{k=1}^{j} \binom{j-1}{k-1} f^{k-1}\left(a\right) f^{j-k}\left(a_i\right). \end{aligned} \tag{34}$$

For each $(i, j) \in \{1, 2, \ldots, n\} \times \{1, 2, \ldots, n\}$, we define an element $c_{i,j}$ of A by

$$c_{i,j} = \begin{cases} \dbinom{j-1}{i-1} f^{j-i}\left(a\right), & \text{if } j \geq i; \\ 0, & \text{if } j < i. \end{cases} \tag{35}$$

Note that this is well-defined (because if $j \geq i$, then $j - i \in \mathbb{N}$ and therefore f^{j-i} is a well-defined map).

For each $i \in \{1, 2, \ldots, n\}$, we have $i \geq i$ and thus

$$\begin{aligned} c_{i,i} &= \underbrace{\binom{i-1}{i-1}}_{=1} \underbrace{f^{i-i}\left(a\right)}_{\substack{=f^0(a) \\ \text{(since } i-i=0\text{)}}} \qquad \text{(by the definition of } c_{i,i}\text{)} \\ &= \underbrace{f^0}_{=\mathrm{id}_A}\left(a\right) = a. \end{aligned} \tag{36}$$

Define two $n \times n$-matrices $B \in A^{n \times n}$ and $C \in A^{n \times n}$ by

$$B = \left(f^{j-1}\left(a_i\right) \right)_{1 \leq i \leq n,\ 1 \leq j \leq n} \qquad \text{and} \qquad C = \left(c_{i,j} \right)_{1 \leq i \leq n,\ 1 \leq j \leq n}.$$

Then, the definition of the product of two matrices shows that

$$BC = \left(\sum_{k=1}^{n} f^{k-1}\left(a_i\right) \cdot c_{k,j} \right)_{1 \leq i \leq n,\ 1 \leq j \leq n}. \tag{37}$$

However, each $(i,j) \in \{1,2,\ldots,n\} \times \{1,2,\ldots,n\}$ satisfies

$$\begin{aligned}
&\sum_{k=1}^{n} f^{k-1}\left(a_i\right) \cdot c_{k,j} \\
&= \sum_{k=1}^{j} f^{k-1}\left(a_i\right) \cdot \underbrace{c_{k,j}}_{\substack{= \binom{j-1}{k-1} f^{j-k}(a) \\ \text{(by the definition of } c_{k,j} \text{ (because } k \leq j\text{))}}} + \sum_{k=j+1}^{n} f^{k-1}\left(a_i\right) \cdot \underbrace{c_{k,j}}_{\substack{=0 \\ \text{(by the definition of } c_{k,j} \text{ (because } k > j\text{))}}} \\
&\qquad \left(\text{here, we have split the sum at } k = j, \text{ since } 1 \leq j \leq n\right) \\
&= \sum_{k=1}^{j} f^{k-1}\left(a_i\right) \cdot \binom{j-1}{k-1} f^{j-k}\left(a\right) + \underbrace{\sum_{k=j+1}^{n} f^{k-1}\left(a_i\right) \cdot 0}_{=0} \\
&= \sum_{k=1}^{j} f^{k-1}\left(a_i\right) \cdot \binom{j-1}{k-1} f^{j-k}\left(a\right) = \sum_{k=1}^{j} \binom{j-1}{k-1} f^{k-1}\left(a\right) f^{j-k}\left(a_i\right) \\
&= f^{j-1}\left(a a_i\right) \qquad \left(\text{by (34)}\right).
\end{aligned}$$

Hence,

$$\left(\sum_{k=1}^{n} f^{k-1}\left(a_i\right) \cdot c_{k,j} \right)_{1 \leq i \leq n,\ 1 \leq j \leq n} = \left(f^{j-1}\left(a a_i\right) \right)_{1 \leq i \leq n,\ 1 \leq j \leq n}.$$

Thus, (37) rewrites as

$$BC = \left(f^{j-1}\left(a a_i\right) \right)_{1 \leq i \leq n,\ 1 \leq j \leq n}. \tag{38}$$

The definition of $c_{i,j}$ shows that $c_{i,j} = 0$ whenever $j < i$. In other words, the $n \times n$-matrix $\left(c_{i,j}\right)_{1 \leq i \leq n,\ 1 \leq j \leq n}$ is upper-triangular. Since the determinant of an upper-triangular matrix is the product of its diagonal entries, we thus conclude that

$$\det\left(\left(c_{i,j}\right)_{1 \leq i \leq n,\ 1 \leq j \leq n} \right) = \prod_{i=1}^{n} \underbrace{c_{i,i}}_{\substack{=a \\ \text{(by (36))}}} = \prod_{i=1}^{n} a = a^n. \tag{39}$$

It is well-known that the determinant of the product of two square matrices equals the product of their determinants. In other words, if M and N are two $n \times n$-matrices (with entries in A), then $\det(MN) = \det M \cdot \det N$. Applying this to $M = B$ and $N = C$, we find

$$\begin{aligned}
\det(BC) &= \det \underbrace{B}_{=\left(f^{j-1}(a_i)\right)_{1\le i\le n,\ 1\le j\le n}} \cdot \det \underbrace{C}_{=\left(c_{i,j}\right)_{1\le i\le n,\ 1\le j\le n}} \\
&= \underbrace{\det\left(\left(f^{j-1}(a_i)\right)_{1\le i\le n,\ 1\le j\le n}\right)}_{=W_f(a_1,a_2,\ldots,a_n) \text{ (by (32))}} \cdot \underbrace{\det\left(\left(c_{i,j}\right)_{1\le i\le n,\ 1\le j\le n}\right)}_{=a^n \text{ (by (39))}} \\
&= W_f(a_1,a_2,\ldots,a_n)\cdot a^n = a^n W_f(a_1,a_2,\ldots,a_n).
\end{aligned}$$

Comparing this with

$$\begin{aligned}
\det(BC) &= \det\left(\left(f^{j-1}(aa_i)\right)_{1\le i\le n,\ 1\le j\le n}\right) \qquad \text{(by (38))} \\
&= W_f(aa_1,aa_2,\ldots,aa_n) \qquad \text{(by (33))},
\end{aligned}$$

we obtain $W_f(aa_1,aa_2,\ldots,aa_n) = a^n W_f(a_1,a_2,\ldots,a_n)$. This proves Theorem 1.50. □

Example 1.51. Let $\mathbf{k}[x]$ and ∂_x be as in Example 1.7. Recall that ∂_x is a derivation. Let $n \in \mathbb{N}$. Let $a_1, a_2, \ldots, a_n \in \mathbf{k}[x]$ be any n polynomials in $\mathbf{k}[x]$. Then, the ∂_x-Wronskian $W_{\partial_x}(a_1,a_2,\ldots,a_n)$ is just called the *Wronskian* of $a_1, a_2, \ldots, a_n$, and is denoted by $W(a_1,a_2,\ldots,a_n)$. Thus, Theorem 1.50 (applied to $A = \mathbf{k}[x]$ and $f = \partial_x$) yields that

$$W(aa_1,aa_2,\ldots,aa_n) = a^n W(a_1,a_2,\ldots,a_n) \qquad \text{for every polynomial } a \in \mathbf{k}[x].$$

2. Derivations from the tensor and symmetric algebras

2.1. The tensor algebra

We now turn to the topic of derivations from certain more specific algebras. We begin with the tensor algebra of a $\mathbf{k}$-module. Let us recall its definition:

Definition 2.1. Let V be a $\mathbf{k}$-module. For every $n \in \mathbb{N}$, we let $V^{\otimes n}$ denote the n-th tensor power of V (that is, the $\mathbf{k}$-module $\underbrace{V \otimes V \otimes \cdots \otimes V}_{n \text{ times } V}$). We let

$T(V)$ denote the tensor algebra of V. This is the $\mathbf{k}$-algebra whose underlying $\mathbf{k}$-module is $\bigoplus_{n \geq 0} V^{\otimes n}$, and whose multiplication is given by

$$(a_1 \otimes a_2 \otimes \cdots \otimes a_n) \cdot (b_1 \otimes b_2 \otimes \cdots \otimes b_m)$$
$$= a_1 \otimes a_2 \otimes \cdots \otimes a_n \otimes b_1 \otimes b_2 \otimes \cdots \otimes b_m$$
$$\text{for every } n \in \mathbb{N},\ m \in \mathbb{N},\ a_1, a_2, \ldots, a_n \in V \text{ and } b_1, b_2, \ldots, b_m \in V\ .$$

We let $\iota_{T,V}$ be the canonical inclusion map of V into $T(V)$. This map is the composition $V \xrightarrow{\cong} V^{\otimes 1} \xrightarrow{\text{inclusion}} \bigoplus_{n \geq 0} V^{\otimes n} = T(V)$. (Here, the "$T$" in "$\iota_{T,V}$" is not a variable, but stands for the letter "t" in "**t**ensor algebra".)

We have $\iota_{T,V}(V) = V^{\otimes 1}$. Some authors identify V with the $\mathbf{k}$-submodule $V^{\otimes 1}$ of $T(V)$ via this map $\iota_{T,V}$ (so that for every $n \in \mathbb{N}$ and $a_1, a_2, \ldots, a_n \in V$, we have $a_1 a_2 \cdots a_n = a_1 \otimes a_2 \otimes \cdots \otimes a_n$ in the $\mathbf{k}$-algebra $T(V)$). This identification is harmless since $\iota_{T,V}$ is injective; but we will not use this identification.

The following fact is well-known as the universal property of the tensor algebra:

Proposition 2.2. Let V be a $\mathbf{k}$-module. Let A be a $\mathbf{k}$-algebra. Let $f : V \to A$ be a $\mathbf{k}$-linear map. Then, there exists a unique $\mathbf{k}$-algebra homomorphism $F : T(V) \to A$ such that $F \circ \iota_{T,V} = f$.

Our goal is now to prove an analogue of this property for derivations instead of $\mathbf{k}$-algebra homomorphisms:

Proposition 2.3. Let V be a $\mathbf{k}$-module. Let M be a $(T(V), T(V))$-bimodule. Let $f : V \to M$ be a $\mathbf{k}$-linear map. Then, there exists a unique derivation $F : T(V) \to M$ such that $F \circ \iota_{T,V} = f$.

Proof of Proposition 2.3. Define a map $g : V \to \mathcal{R}_{T(V)}(M)$ by

$$g = \mathbf{inl}_{T(V),M} \circ \iota_{T,V} + \mathbf{inr}_{T(V),M} \circ f. \tag{40}$$

This map g is $\mathbf{k}$-linear (since the maps $\mathbf{inl}_{T(V),M}$, $\iota_{T,V}$, $\mathbf{inr}_{T(V),M}$ and f are $\mathbf{k}$-linear).

From $g = \mathbf{inl}_{T(V),M} \circ \iota_{T,V} + \mathbf{inr}_{T(V),M} \circ f$, we obtain

$$\begin{aligned}
&\mathbf{projl}_{T(V),M} \circ g \\
&= \mathbf{projl}_{T(V),M} \circ \left(\mathbf{inl}_{T(V),M} \circ \iota_{T,V} + \mathbf{inr}_{T(V),M} \circ f \right) \\
&= \mathbf{projl}_{T(V),M} \circ \left(\mathbf{inl}_{T(V),M} \circ \iota_{T,V} \right) + \mathbf{projl}_{T(V),M} \circ \left(\mathbf{inr}_{T(V),M} \circ f \right) \\
&\qquad \left(\text{since the maps } \mathbf{projl}_{T(V),M},\ \mathbf{inl}_{T(V),M} \circ \iota_{T,V} \text{ and } \mathbf{inr}_{T(V),M} \circ f \text{ are } \mathbf{k}\text{-linear} \right) \\
&= \underbrace{\mathbf{projl}_{T(V),M} \circ \mathbf{inl}_{T(V),M}}_{\substack{=\mathrm{id}_{T(V)} \\ \text{(by (18), applied to } A=T(V))}} \circ \iota_{T,V} + \underbrace{\mathbf{projl}_{T(V),M} \circ \mathbf{inr}_{T(V),M}}_{\substack{=0 \\ \text{(by (19), applied to } A=T(V))}} \circ f \\
&= \underbrace{\mathrm{id}_{T(V)} \circ \iota_{T,V}}_{=\iota_{T,V}} + \underbrace{0 \circ f}_{=0} = \iota_{T,V}.
\end{aligned}$$

Recall that $\mathrm{Hom}_{\mathcal{R}}\left(T\left(V\right), M\right)$ was defined by

$$\begin{aligned}
&\mathrm{Hom}_{\mathcal{R}}\left(T\left(V\right), M\right) \\
&= \left\{ \varphi \in \mathrm{Hom}\left(T\left(V\right), \mathcal{R}_{T(V)}\left(M\right) \right) \ \mid \ \mathbf{projl}_{T(V),M} \circ \varphi = \mathrm{id}_{T(V)} \right\}. \qquad (41)
\end{aligned}$$

Proposition 2.2 (applied to g and $\mathcal{R}_{T(V)}\left(M\right)$ instead of f and A) now shows that there exists a unique $\mathbf{k}$-algebra homomorphism $F : T\left(V\right) \to \mathcal{R}_{T(V)}\left(M\right)$ such that $F \circ \iota_{T,V} = g$. Let us denote this F by G. Thus, G is a $\mathbf{k}$-algebra homomorphism $T\left(V\right) \to \mathcal{R}_{T(V)}\left(M\right)$ and satisfies $G \circ \iota_{T,V} = g$.

Proposition 1.35 **(a)** (applied to $A = T\left(V\right)$) shows that the map $\mathbf{projl}_{T(V),M} : \mathcal{R}_{T(V)}\left(M\right) \to T\left(V\right)$ is a $\mathbf{k}$-algebra homomorphism from $\mathcal{R}_{T(V)}\left(M\right)$ to $T\left(V\right)$. Hence, the map $\mathbf{projl}_{T(V),M} \circ G$ is a $\mathbf{k}$-algebra homomorphism (since it is the composition of the two $\mathbf{k}$-algebra homomorphisms $\mathbf{projl}_{T(V),M}$ and G).

But Proposition 2.2 (applied to $T\left(V\right)$ and $\iota_{T,V}$ instead of A and f) shows that there exists a unique $\mathbf{k}$-algebra homomorphism $F : T\left(V\right) \to T\left(V\right)$ such that $F \circ \iota_{T,V} = \iota_{T,V}$. In particular, there exists **at most one** such homomorphism. In other words, if F_1 and F_2 are two $\mathbf{k}$-algebra homomorphisms $F : T\left(V\right) \to T\left(V\right)$ such that $F \circ \iota_{T,V} = \iota_{T,V}$, then

$$F_1 = F_2. \qquad (42)$$

Now, both $\mathrm{id}_{T(V)}$ and $\mathbf{projl}_{T(V),M} \circ G$ are $\mathbf{k}$-algebra homomorphisms $F : T\left(V\right) \to T\left(V\right)$ such that $F \circ \iota_{T,V} = \iota_{T,V}$ (indeed, this is clear for $\mathrm{id}_{T(V)}$, whereas for $\mathbf{projl}_{T(V),M} \circ G$ it follows from $\mathbf{projl}_{T(V),M} \circ \underbrace{G \circ \iota_{T,V}}_{=g} = \mathbf{projl}_{T(V),M} \circ g = \iota_{T,V}$). Thus, we can apply (42) to $F_1 = \mathbf{projl}_{T(V),M} \circ G$ and $F_2 = \mathrm{id}_{T(V)}$. As a result, we obtain $\mathbf{projl}_{T(V),M} \circ G = \mathrm{id}_{T(V)}$.

Now, G is an element of $\mathrm{Hom}\left(T\left(V\right), \mathcal{R}_{T(V)}\left(M\right) \right)$ and satisfies $\mathbf{projl}_{T(V),M} \circ$

$G = \mathrm{id}_{T(V)}$. In other words,

$$\begin{aligned} G &\in \left\{ \varphi \in \mathrm{Hom}\left(T\left(V\right), \mathcal{R}_{T(V)}\left(M\right) \right) \ \mid \ \mathbf{projl}_{T(V),M} \circ \varphi = \mathrm{id}_{T(V)} \right\} \\ &= \mathrm{Hom}_{\mathcal{R}}\left(T\left(V\right), M\right) \qquad \left(\text{by (41)}\right). \end{aligned}$$

Consider the maps $\mathrm{dth}_{T(V),M} : \mathrm{Hom}\left(T\left(V\right), M\right) \to \mathrm{Hom}_{\mathcal{R}}\left(T\left(V\right), M\right)$ and $\mathrm{htd}_{T(V),M} : \mathrm{Hom}_{\mathcal{R}}\left(T\left(V\right), M\right) \to \mathrm{Hom}\left(T\left(V\right), M\right)$ defined in Definition 1.39. Theorem 1.40 **(b)** (applied to $A = T\left(V\right)$) shows that the maps $\mathrm{dth}_{T(V),M}$ and $\mathrm{htd}_{T(V),M}$ are mutually inverse. In particular,

$$\mathrm{dth}_{T(V),M} \circ \mathrm{htd}_{T(V),M} = \mathrm{id}_{\mathrm{Hom}_{\mathcal{R}}(T(V),M)}.$$

We shall now prove the following two statements:

Statement 1: There exists **at least one** derivation $F : T\left(V\right) \to M$ such that $F \circ \iota_{T,V} = f$.

Statement 2: There exists **at most one** derivation $F : T\left(V\right) \to M$ such that $F \circ \iota_{T,V} = f$.

Proof of Statement 1: Recall that $G \in \mathrm{Hom}_{\mathcal{R}}\left(T\left(V\right), M\right)$.

Define an element $h \in \mathrm{Hom}\left(T\left(V\right), M\right)$ by $h = \mathrm{htd}_{T(V),M}\left(G\right)$. Thus,

$$h = \mathrm{htd}_{T(V),M}\left(G\right) = \mathbf{projr}_{T(V),M} \circ G \qquad \left(\text{by the definition of } \mathrm{htd}_{T(V),M}\right),$$

so that

$$\begin{aligned} &\underbrace{h}_{=\mathbf{projr}_{T(V),M} \circ G} \circ \iota_{T,V} \\ &= \mathbf{projr}_{T(V),M} \circ \underbrace{G \circ \iota_{T,V}}_{= g = \mathbf{inl}_{T(V),M} \circ \iota_{T,V} + \mathbf{inr}_{T(V),M} \circ f} \\ &= \mathbf{projr}_{T(V),M} \circ \left(\mathbf{inl}_{T(V),M} \circ \iota_{T,V} + \mathbf{inr}_{T(V),M} \circ f \right) \\ &= \mathbf{projr}_{T(V),M} \circ \left(\mathbf{inl}_{T(V),M} \circ \iota_{T,V} \right) + \mathbf{projr}_{T(V),M} \circ \left(\mathbf{inr}_{T(V),M} \circ f \right) \\ &\qquad \left(\text{since the maps } \mathbf{projr}_{T(V),M}, \ \mathbf{inl}_{T(V),M} \circ \iota_{T,V} \text{ and } \mathbf{inr}_{T(V),M} \circ f \text{ are } \mathbf{k}\text{-linear}\right) \\ &= \underbrace{\mathbf{projr}_{T(V),M} \circ \mathbf{inl}_{T(V),M}}_{\substack{=0 \\ \text{(by (20), applied to } A = T(V)\text{)}}} \circ \iota_{T,V} + \underbrace{\mathbf{projr}_{T(V),M} \circ \mathbf{inr}_{T(V),M}}_{\substack{=\mathrm{id}_M \\ \text{(by (21), applied to } A = T(V)\text{)}}} \circ f \\ &= \underbrace{0}_{=0} \circ \iota_{T,V} + \underbrace{\mathrm{id}_M \circ f}_{=f} = f. \end{aligned}$$

Also,

$$\operatorname{dth}_{T(V),M}\left(\underbrace{h}_{=\operatorname{htd}_{T(V),M}(G)}\right) = \operatorname{dth}_{T(V),M}\left(\operatorname{htd}_{T(V),M}(G)\right)$$
$$= \underbrace{\left(\operatorname{dth}_{T(V),M} \circ \operatorname{htd}_{T(V),M}\right)}_{=\operatorname{id}_{\operatorname{Hom}_{\mathcal{R}}(T(V),M)}}(G) = G.$$

Thus, $\operatorname{dth}_{T(V),M}(h)$ is a $\mathbf{k}$-algebra homomorphism (since G is a $\mathbf{k}$-algebra homomorphism).

But Theorem 1.40 **(c)** (applied to $T(V)$ and h instead of A and f) shows that h is a derivation if and only if $\operatorname{dth}_{T(V),M}(h)$ is a $\mathbf{k}$-algebra homomorphism from $T(V)$ to $\mathcal{R}_{T(V)}(M)$. Thus, h is a derivation (since we know that $\operatorname{dth}_{T(V),M}(h)$ is a $\mathbf{k}$-algebra homomorphism from $T(V)$ to $\mathcal{R}_{T(V)}(M)$).

So we know that h is a derivation from $T(V)$ to M and satisfies $h \circ \iota_{T,V} = f$. Thus, there exists **at least one** derivation $F : T(V) \to M$ such that $F \circ \iota_{T,V} = f$ (namely, $F = h$). This proves Statement 1.

Proof of Statement 2: Let φ_1 and φ_2 be two derivations $F : T(V) \to M$ such that $F \circ \iota_{T,V} = f$. We shall prove that $\varphi_1 = \varphi_2$.

We know that φ_1 is a derivation $F : T(V) \to M$ such that $F \circ \iota_{T,V} = f$. In other words, φ_1 is a derivation $T(V) \to M$ and satisfies $\varphi_1 \circ \iota_{T,V} = f$.

Let $\Phi_1 = \operatorname{dth}_{T(V),M}(\varphi_1)$ and $\Phi_2 = \operatorname{dth}_{T(V),M}(\varphi_2)$.

Theorem 1.40 **(c)** (applied to $T(V)$ and φ_1 instead of A and f) shows that φ_1 is a derivation if and only if $\operatorname{dth}_{T(V),M}(\varphi_1)$ is a $\mathbf{k}$-algebra homomorphism from $T(V)$ to $\mathcal{R}_{T(V)}(M)$. Thus, $\operatorname{dth}_{T(V),M}(\varphi_1)$ is a $\mathbf{k}$-algebra homomorphism from $T(V)$ to $\mathcal{R}_{T(V)}(M)$ (since we know that φ_1 is a derivation). In other words, Φ_1 is a $\mathbf{k}$-algebra homomorphism from $T(V)$ to $\mathcal{R}_{T(V)}(M)$ (since $\Phi_1 = \operatorname{dth}_{T(V),M}(\varphi_1)$). The same argument (but applied to φ_2 and Φ_2 instead of φ_1 and Φ_1) shows that Φ_2 is a $\mathbf{k}$-algebra homomorphism from $T(V)$ to $\mathcal{R}_{T(V)}(M)$.

Now,

$$\Phi_1 = \operatorname{dth}_{T(V),M}(\varphi_1) = \mathbf{inl}_{T(V),M} + \mathbf{inr}_{T(V),M} \circ \varphi_1$$
$$\left(\text{by the definition of } \operatorname{dth}_{T(V),M}\right).$$

Hence,

$$\begin{aligned}
&\underbrace{\Phi_1}_{=\mathbf{inl}_{T(V),M}+\mathbf{inr}_{T(V),M}\circ\varphi_1} \circ \iota_{T,V} \\
&= \left(\mathbf{inl}_{T(V),M} + \mathbf{inr}_{T(V),M} \circ \varphi_1\right) \circ \iota_{T,V} \\
&= \mathbf{inl}_{T(V),M} \circ \iota_{T,V} + \underbrace{\left(\mathbf{inr}_{T(V),M} \circ \varphi_1\right) \circ \iota_{T,V}}_{=\mathbf{inr}_{T(V),M}\circ\varphi_1\circ\iota_{T,V}} \\
&\qquad \left(\begin{array}{c} \text{since composition of } \mathbf{k}\text{-linear maps is } \mathbf{k}\text{-bilinear, and since} \\ \text{the maps } \mathbf{inl}_{T(V),M}, \ \mathbf{inr}_{T(V),M} \circ \varphi_1 \text{ and } \iota_{T,V} \text{ are } \mathbf{k}\text{-linear} \end{array} \right) \\
&= \mathbf{inl}_{T(V),M} \circ \iota_{T,V} + \mathbf{inr}_{T(V),M} \circ \underbrace{\varphi_1 \circ \iota_{T,V}}_{=f} \\
&= \mathbf{inl}_{T(V),M} \circ \iota_{T,V} + \mathbf{inr}_{T(V),M} \circ f = g \qquad \text{(by (40))}.
\end{aligned}$$

The same argument (but applied to φ_2 and Φ_2 instead of φ_1 and Φ_1) shows that $\Phi_2 \circ \iota_{T,V} = g$.

Let us recall that there exists a unique $\mathbf{k}$-algebra homomorphism $F : T(V) \to \mathcal{R}_{T(V)}(M)$ such that $F \circ \iota_{T,V} = g$. In particular, there exists **at most one** such homomorphism. In other words, if F_1 and F_2 are two $\mathbf{k}$-algebra homomorphisms $F : T(V) \to \mathcal{R}_{T(V)}(M)$ such that $F \circ \iota_{T,V} = g$, then

$$F_1 = F_2. \tag{43}$$

Now, both Φ_1 and Φ_2 are $\mathbf{k}$-algebra homomorphisms $F : T(V) \to \mathcal{R}_{T(V)}(M)$ such that $F \circ \iota_{T,V} = g$ (since $\Phi_1 \circ \iota_{T,V} = g$ and $\Phi_2 \circ \iota_{T,V} = g$). Thus, we can apply (43) to $F_1 = \Phi_1$ and $F_2 = \Phi_2$. As a result, we obtain $\Phi_1 = \Phi_2$.

Recall that the maps $\operatorname{dth}_{T(V),M}$ and $\operatorname{htd}_{T(V),M}$ are mutually inverse. Thus, the map $\operatorname{dth}_{T(V),M}$ is invertible, i.e., bijective, and therefore injective.

But $\operatorname{dth}_{T(V),M}(\varphi_1) = \Phi_1 = \Phi_2 = \operatorname{dth}_{T(V),M}(\varphi_2)$. This shows that $\varphi_1 = \varphi_2$ (since the map $\operatorname{dth}_{T(V),M}$ is injective).

Let us now forget that we fixed φ_1 and φ_2. We thus have shown that if φ_1 and φ_2 are two derivations $F : T(V) \to M$ such that $F \circ \iota_{T,V} = f$, then $\varphi_1 = \varphi_2$. In other words, there exists **at most one** derivation $F : T(V) \to M$ such that $F \circ \iota_{T,V} = f$. This proves Statement 2.

Now, both Statement 1 and Statement 2 have been proven. Combining these two statements, we obtain Proposition 2.3.

□

Remark 2.4. Our proof of Proposition 2.3 above is not the only one possible; it is probably not even the shortest one, and it is certainly not the most enlightening one. It does, however, illustrate how a result can be proven using universal properties and constructions such as $\mathcal{R}_A(M)$ without "getting one's hands dirty" (e.g., computing with elements).

Let me sketch an alternative way to prove Proposition 2.3: We subdivide the claim into a Statement 1 and a Statement 2 as in our above proof. To prove Statement 1, we can construct a derivation $F : T(V) \to M$ such that $F \circ \iota_{T,V} = f$ explicitly, by setting

$$F(v_1 \otimes v_2 \otimes \cdots \otimes v_n) = \sum_{i=1}^{n} (v_1 \otimes v_2 \otimes \cdots \otimes v_{i-1}) \cdot f(v_i) \cdot (v_{i+1} \otimes v_{i+2} \otimes \cdots \otimes v_n)$$
$$\text{for every } n \in \mathbb{N} \text{ and } v_1, v_2, \ldots, v_n \in V .$$

It is not hard to prove that this is well-defined (i.e., there exists a unique $\mathbf{k}$-linear map $F : T(V) \to M$ satisfying this equality), and indeed defines a derivation $F : T(V) \to M$ satisfying $F \circ \iota_{T,V} = f$. To prove Statement 2, we could use Proposition 1.29 (applied to $A = T(V)$ and $S = \iota_{T,V}(V)$), since $\iota_{T,V}(V)$ generates the $\mathbf{k}$-algebra $T(V)$.

See [EtiGri12, Theorem 4.6.12] for the details of this proof of Proposition 2.3. (The actual statement of [EtiGri12, Theorem 4.6.12] is the particular case of Proposition 2.3 for $\mathbf{k} = \mathbb{C}$; but the proof applies to the general case without changes.)

2.2. The symmetric algebra

We now recall the definition of the symmetric algebra of a $\mathbf{k}$-module V:

Definition 2.5. Let V be a $\mathbf{k}$-module. Let J_V denote the two-sided ideal of the $\mathbf{k}$-algebra $T(V)$ generated by the tensors of the form $v \otimes w - w \otimes v$ with $v \in V$ and $w \in V$. We define $\operatorname{Sym} V$ to be the quotient algebra $T(V)/J_V$. We let $\pi_{\operatorname{Sym},V}$ be the canonical projection from $T(V)$ to its quotient $\mathbf{k}$-algebra $\operatorname{Sym} V$. The $\mathbf{k}$-algebra $\operatorname{Sym} V$ is known as the *symmetric algebra* of V. It is well-known that the $\mathbf{k}$-algebra $\operatorname{Sym} V$ is commutative.

For every $n \in \mathbb{N}$, we write $\operatorname{Sym}^n V$ for the $\mathbf{k}$-submodule $\pi_{\operatorname{Sym},V}(V^{\otimes n})$ of $\operatorname{Sym} V$. It is well-known that $\operatorname{Sym} V = \bigoplus_{n \geq 0} \operatorname{Sym}^n V$.

We let $\iota_{\operatorname{Sym},V}$ denote the composition $\pi_{\operatorname{Sym},V} \circ \iota_{T,V}$. This composition $\iota_{\operatorname{Sym},V}$ is a $\mathbf{k}$-linear map $V \to \operatorname{Sym} V$. It is well-known that this map $\iota_{\operatorname{Sym},V}$ is injective and satisfies $\iota_{\operatorname{Sym},V}(V) = \operatorname{Sym}^1 V$.

The following fact is well-known as the universal property of the symmetric algebra:

Proposition 2.6. Let V be a $\mathbf{k}$-module. Let A be a commutative $\mathbf{k}$-algebra. Let $f : V \to A$ be a $\mathbf{k}$-linear map. Then, there exists a unique $\mathbf{k}$-algebra homomorphism $F : \operatorname{Sym} V \to A$ such that $F \circ \iota_{\operatorname{Sym},V} = f$.

This is clearly an analogue of Proposition 2.2. We can similarly state an analogue of Proposition 2.3:

Proposition 2.7. Let V be a $\mathbf{k}$-module. Let M be a symmetric $(\operatorname{Sym} V, \operatorname{Sym} V)$-bimodule. Let $f : V \to M$ be a $\mathbf{k}$-linear map. Then, there exists a unique derivation $F : \operatorname{Sym} V \to M$ such that $F \circ \iota_{\operatorname{Sym},V} = f$.

Proof of Proposition 2.7. To obtain a proof of Proposition 2.7, it suffices to make the following changes to our above proof of Proposition 2.3:

- Add the following argument at the beginning of the proof: "The $\mathbf{k}$-algebra $\mathcal{R}_{\operatorname{Sym} V}(M)$ is commutative (by Proposition 1.41, applied to $A = \operatorname{Sym} V$).".
- Replace each appearance of "$T(V)$" by "$\operatorname{Sym} V$".
- Replace each appearance of "$\iota_{T,V}$" by "$\iota_{\operatorname{Sym},V}$".
- Replace each reference to Proposition 2.2 by a reference to Proposition 2.6. (Proposition 2.6 can indeed be applied since the two $\mathbf{k}$-algebras $\operatorname{Sym} V$ and $\mathcal{R}_{\operatorname{Sym} V}(M)$ are commutative.)

□

Remark 2.8. It is possible to prove Proposition 2.7 in more down-to-earth manners, just as the same is possible for Proposition 2.3 (see Remark 2.4 above).

References

[EtiGri12] Pavel Etingof, *18.747: Infinite-dimensional Lie algebras (Spring term 2012 at MIT)*, scribed by Darij Grinberg, version 0.44.
`https://sites.google.com/site/darijgrinberg/lie`

[Becker14] Hanno Becker, *Extension of R-linear derivation to localization*, math.stackexchange post #1052192 (answer to question #1052153),
`http://math.stackexchange.com/q/1052192`

[CSA14] CSA and others, *Contrasting definitions of bimodules? An illusion?*, math.stackexchange post #889130 and answers,
`http://math.stackexchange.com/q/889130`

[Grinbe16] Darij Grinberg, *Fleck's binomial congruence using circulant matrices*, version 14 July 2019.
`http://www.cip.ifi.lmu.de/~grinberg/fleck.pdf`