Elementary derivations of some results of linear optimization

Darij Grinberg

version 1.0, 7 August 2017

Contents

16. Appendix: Proofs omitted from the early sections 80

17. Appendix: Old (and ugly) proofs of Theorems 2.1c and 2.5c 122

Remark (2017):

This note has been mostly written in 2012, when I was learning combinatorial optimization from Schrijver's online notes [Schrij17] and felt a need for elementary and constructive proofs of the main theorems of linear optimization (as opposed to the standard short proofs using compactness and the Hahn-Banach theorem). This note provides such proofs, although I cannot vouch for their readability[1].

Needless to say, this note breaks no new ground, and probably these proofs (or easier ones) appear often enough in the literature (in fact, I suspect that almost all of the material of this note is covered by each of the two textbooks [Laurit13] and [Schrij98]). Writing them up was a learning experience which, I fear, reading them will not be.

While this note often refers to [Schrij17], it is actually self-contained and can be read separately (except for the few places where it modifies arguments from [Schrij17] to make them constructive; but these are tangential to the note).

The purpose of this note is to give elementary proofs for various results in linear optimization theory. Here, "elementary" means that no analysis is being used, and that the proofs are "morally" constructive.

Let me explain what "morally constructive" means: The results and proofs given below are not valid in constructive logic, but this is solely for the reason that $\mathbb{R}$ is not a discrete field in constructive logic. If we would formulate the results and proofs below for $\mathbb{Q}$ instead of $\mathbb{R}$, then they would become valid in constructive logic. Thus, when I make any claim in constructive logic below, I tacitly want it to be understood with all $\mathbb{R}$'s replaced by $\mathbb{Q}$'s (and all "real numbers" replaced by "rational numbers", and so on).

Note that the following proofs, being free of analysis, generalize to any ordered field instead of $\mathbb{R}$ (for example, to $\mathbb{Q}$ or to $\mathbb{Q}\left(\sqrt{2}\right)$). However, we are going to formulate them for the field $\mathbb{R}$ only (trusting that the reader, if necessary, can generalize them on his own). In particular, whenever we speak of "vector spaces", "vectors" and "matrices" below, we mean vector spaces, vectors and matrices over $\mathbb{R}$.

This note was originally intended as a supplement to Chapter 2 of Schrijver's notes [Schrij17][2]; but it is fully self-contained. It proves separation theorems for convex hulls and cones (of finite sets), the Farkas, Gordan, Stiemke and Motzkin theorems, and two versions of the linear programming duality theorem. Also, it "patches" the proof of Theorem 2.3 in [Schrij17] (the theorem saying that any bounded polyhedron is the convex hull of its vertices) to make it constructive.

---

[1]They are overdetailed in many places, the result of my attempts to ensure their correctness; unfortunately the trees often obscure the forest.

[2]In particular, this explains the strange numbering of results in this note: The numbers have been chosen so as not to conflict with the labeling of [Schrij17].

1. Notations and basics 1: Convex sets

Before we start formulating and proving substantial theorems, let us introduce some notations. The notations that we introduce will usually be identical to the ones introduced in [Schrij17], save for a few exceptions.

In the following, $\mathbb{N}$ will always denote the set $\{0, 1, 2, \ldots\}$; thus, we regard 0 as an element of $\mathbb{N}$ (unlike some other authors).

We will use the words "vector" and "point" as synonyms; both of them will denote elements of a vector space.

> **Definition.** Let E be an $\mathbb{R}$-vector space. Let C be a subset of E. We say that C is *convex* if every two elements $x \in C$ and $y \in C$ and every real number $\lambda \in [0, 1]$ satisfy $\lambda x + (1 - \lambda) y \in C$.

Geometrically, this definition is often put into words as follows: Let E be an $\mathbb{R}$-vector space. Let C be a subset of E. We say that C is *convex* if every point on a segment containing two points of C must also lie in C.

The following property of convex sets is easy to see:

> **Proposition 2.0a.** Let E be an $\mathbb{R}$-vector space. Let C be a convex subset of E. Let I be any set, and let $(x_i)_{i \in I}$ be a family of elements of C indexed by elements of I. Let $(\lambda_i)_{i \in I}$ be a family of nonnegative reals indexed by elements of I such that all but finitely many $i \in I$ satisfy $\lambda_i = 0$. Assume also that $\sum_{i \in I} \lambda_i = 1$. Then, $\sum_{i \in I} \lambda_i x_i \in C$.

For the sake of completeness, we shall give a proof of Proposition 2.0a in Section 16.

> **Definition.** Let E be an $\mathbb{R}$-vector space. Let I be any set, and let $(x_i)_{i \in I}$ be a family of elements of E indexed by elements of I. If $(\lambda_i)_{i \in I}$ is a family of nonnegative reals indexed by elements of I such that all but finitely many $i \in I$ satisfy $\lambda_i = 0$, and such that $\sum_{i \in I} \lambda_i = 1$, then the vector $\sum_{i \in I} \lambda_i x_i$ is said to be a *convex combination* of the vectors x_i for $i \in I$.

Using this definition, Proposition 2.0a rewrites as follows:

> **Proposition 2.0b.** Let E be an $\mathbb{R}$-vector space. Let C be a convex subset of E. Let I be any set, and let $(x_i)_{i \in I}$ be a family of elements of C indexed by elements of I. Then, any convex combination of the vectors x_i for $i \in I$ lies in C.

This result is often stated in words as follows: A convex set is closed under convex combinations.

The notion of convex sets gives rise to another notion: that of a convex hull. The notion of a convex hull can be defined in several ways; here are three:

> **Definition 2.0c.** Let E be an $\mathbb{R}$-vector space. Let S be a subset of E. Then, the *convex hull* of S will denote the intersection of all convex subsets of E which contain S as a subset. We denote the convex hull of S by $\operatorname{conv.hull} S$.

> **Definition 2.0d.** Let E be an $\mathbb{R}$-vector space. Let S be a subset of E. Then, the *convex hull* of S will denote the set of all convex combinations of the vectors s for $s \in S$. (This will often be abbreviated as follows: "The convex hull of S will denote the set of all convex combinations of the elements of S.") We denote the convex hull of S by $\operatorname{conv.hull} S$.

> **Definition 2.0e.** Let E be an $\mathbb{R}$-vector space. Let S be a subset of E. Then, the *convex hull* of S will denote the set
>
> $$\left\{ x \in E \;\middle|\; \left(\begin{array}{c} \text{there exist some } t \in \mathbb{N}, \\ \text{a } t\text{-tuple } (x_1, x_2, \ldots, x_t) \text{ of elements of } S \\ \text{and a } t\text{-tuple } (\lambda_1, \lambda_2, \ldots, \lambda_t) \text{ of nonnegative reals} \\ \text{such that } \sum_{i=1}^{t} \lambda_i = 1 \text{ and } \sum_{i=1}^{t} \lambda_i x_i = x \end{array} \right) \right\}.$$
>
> We denote the convex hull of S by $\operatorname{conv.hull} S$.

Definitions 2.0c, 2.0d and 2.0e are equivalent (according to Proposition 2.0f **(a)** below).

The following result gathers some fundamental properties of convex hulls, most of which are standardly used without explicitly calling them out:

> **Proposition 2.0f. (a)** Definitions 2.0c, 2.0d and 2.0e are equivalent.
>
> Let E be an $\mathbb{R}$-vector space.
>
> **(b)** We have $\operatorname{conv.hull} \varnothing = \varnothing$.
>
> Let now S be a subset of E.
>
> **(c)** We have $S \subseteq \operatorname{conv.hull} S$.
>
> **(d)** The convex hull $\operatorname{conv.hull} S$ is a convex set.
>
> **(e)** If T is a subset of S, then $\operatorname{conv.hull} T \subseteq \operatorname{conv.hull} S$.
>
> **(f)** Every convex subset of E which contains S as a subset also contains $\operatorname{conv.hull} S$ as a subset.
>
> **(g)** If T is a subset of $\operatorname{conv.hull} S$, then $\operatorname{conv.hull} T \subseteq \operatorname{conv.hull} S$.

Again, we refer to Section 16 for a proof of Proposition 2.0f.

> **Remark.** Let E be an $\mathbb{R}$-vector space. Let S be a subset of E.
> Then, the convex hull $\mathrm{conv.\,hull}\, S$ is a convex set (by Proposition 2.0f
> **(d)**) and contains S as a subset (by Proposition 2.0f **(c)**). Moreover,
> every convex subset of E which contains S as a subset also contains
> $\mathrm{conv.\,hull}\, S$ as a subset (by Proposition 2.0f **(f)**). This result is often
> put into words as follows: "The convex hull $\mathrm{conv.\,hull}\, S$ is the smallest
> convex set containing S as a subset."

The following proposition (which is, again, fundamental and will be used without explicit mention) is a simple consequence of Definition 2.0e:

> **Proposition 2.0g.** Let E be an $\mathbb{R}$-vector space.
> **(a)** If S is any subset of E, then
>
> $\mathrm{conv.\,hull}\, S$
>
> $$= \left\{ \sum_{i=1}^{t} \lambda_i x_i \;\middle|\; \left(\begin{array}{c} t \text{ is an element of } \mathbb{N}, \\ \text{and } (x_1, x_2, \ldots, x_t) \text{ is a } t\text{-tuple of elements of } S, \\ \text{and } (\lambda_1, \lambda_2, \ldots, \lambda_t) \text{ is a } t\text{-tuple of nonnegative reals} \\ \text{such that } \sum_{i=1}^{t} \lambda_i = 1 \end{array} \right) \right\}.$$
>
> **(b)** Let F be a vector subspace of E. Let S be a subset of F. Then,
> the convex hull $\mathrm{conv.\,hull}\, S$ does not depend on whether we consider
> S as a subset of F or as a subset of E.

Again, we refer to Section 16 for a proof of Proposition 2.0g.

Here is yet another proposition that will be used tacitly:

> **Proposition 2.0h.** Let E be an $\mathbb{R}$-vector space. Let x_1, x_2, ..., x_n
> be finitely many vectors in E.
> **(a)** Then,
>
> $\mathrm{conv.\,hull}\, \{x_1, x_2, \ldots, x_n\}$
> $= (\text{the set of all convex combinations of the vectors } x_1, x_2, \ldots, x_n).$
>
> **(b)** Let $x \in E$. Then, we have $x \in \mathrm{conv.\,hull}\, \{x_1, x_2, \ldots, x_n\}$ if and
> only if x is a convex combination of the vectors x_1, x_2, ..., x_n.

Again, we refer to Section 16 for a proof of Proposition 2.0h.

Now, we can easily define polytopes:

Definition. Let E be an $\mathbb{R}$-vector space. Let P be a subset of E. Then, P is said to be a *polytope* if P is the convex hull of a finite subset of E. In other words, P is said to be a *polytope* if there are finitely many vectors x_1, x_2, …, x_n in E such that $P = \text{conv . hull}\{x_1, x_2, \ldots, x_n\}$.

We are not going to define the notion of a polyhedron (nor will we use this notion); we refer to [Schrij17] for that.

2. Notations and basics 2: Cones

Convex cones are a notion similar to that of convex sets.

Definition. Let E be an $\mathbb{R}$-vector space. Let C be a subset of E. We say that C is a *convex cone in E* if it satisfies the following two conditions:

- We have $0 \in C$.
- Every two elements $x \in C$ and $y \in C$ and every nonnegative reals λ and μ satisfy $\lambda x + \mu y \in C$.

We will abbreviate "convex cone in E" as "convex cone" when the value of E is clear from the context.

Note that this definition is in slight conflict with the definition of a convex cone in [Schrij17]. In fact, the definition of a convex cone in [Schrij17] does not contain the condition that $0 \in C$. This is irrelevant for nonempty subsets C of E, because if a nonempty subset C of E is a convex cone in the sense of [Schrij17], then it must automatically contain 0 and therefore is also a convex cone in the sense of our definition. So the only difference between our definition of a convex cone and the definition given in [Schrij17] is that the empty set $\varnothing$ is a convex cone in the sense of [Schrij17], but not a convex cone in the sense of our definition. This is not a particularly significant difference.[3]

The name "convex cone" is somewhat presumptuous: It seems to imply that any convex cone is a convex set, although we have not proved this. Fortunately, this is true: Any convex cone is indeed a convex set[4].

---

[3]One advantage of our definition is that it satisfies the identity $\text{cone}(A \cup B) = \text{cone } A + \text{cone } B$ for any two subsets A and B of E (whereas with Schrijver's definition, this holds only when A and B are both nonempty or both empty). We will not use this identity, however.

[4]*Proof.* Let E be an $\mathbb{R}$-vector space. Let C be a convex cone in E. We shall prove that C is a convex set.

Recall that C is a convex cone. Hence, every two elements $x \in C$ and $y \in C$ and every

Remark. Let E be an $\mathbb{R}$-vector space. Let C be a convex cone in E. It is easy to see that every $x \in C$ and every nonnegative real λ satisfy $\lambda x \in C$. In words, this is often stated as follows: "Convex cones are closed under multiplication by a nonnegative scalar".

Similarly to Proposition 2.0a, we have:

Proposition 2.0i. Let E be an $\mathbb{R}$-vector space. Let C be a convex cone in E. Let I be any set, and let $(x_i)_{i \in I}$ be a family of elements of C indexed by elements of I. Let $(\lambda_i)_{i \in I}$ be a family of nonnegative reals indexed by elements of I such that all but finitely many $i \in I$ satisfy $\lambda_i = 0$. Then, $\sum\limits_{i \in I} \lambda_i x_i \in C$.

In words, Proposition 2.0i is often stated as follows: "Any linear combination of finitely many elements of a convex cone with nonnegative coefficients must lie in this cone."

The proof of Proposition 2.0i is similar to (and somewhat easier than) that of Proposition 2.0a; for the sake of completeness, we shall give it in Section 16.

Just as the notion of convex sets gave rise to the notion of convex hulls, the notion of convex cones will give rise to the notion of convex conic hulls. Here are three definitions for this notion:

Definition 2.0j. Let E be an $\mathbb{R}$-vector space. Let S be a subset of E. Then, the *convex conic hull* of S will denote the intersection of all convex cones in E which contain S as a subset. We denote the convex conic hull of S by $\operatorname{cone} S$.

Definition 2.0k. Let E be an $\mathbb{R}$-vector space. Let S be a subset of E. Then, the *convex conic hull* of S will denote the set of all linear combinations of the vectors s for $s \in S$ with nonnegative coefficients. (This will often be abbreviated as follows: "The convex conic hull of S will denote the set of all linear combinations of the elements of S with nonnegative coefficients.") We denote the convex conic hull of S by $\operatorname{cone} S$.

---

nonnegative reals λ and μ satisfy

$$\lambda x + \mu y \in C \tag{1}$$

(by the definition of a convex cone).

Now, let $x \in C$, $y \in C$ and $\lambda \in [0, 1]$. Then, both λ and $1 - \lambda$ are nonnegative reals (since $\lambda \in [0, 1]$). Hence, $\lambda x + (1 - \lambda) y \in C$ (by (1), applied to $\mu = 1 - \lambda$).

Let us now forget that we fixed x, y and λ. We thus have proven that every two elements $x \in C$ and $y \in C$ and every real number $\lambda \in [0, 1]$ satisfy $\lambda x + (1 - \lambda) y \in C$. In other words, C is a convex set (by the definition of a "convex set").

Now, let us forget that we fixed C. We thus have shown that if C is a convex cone in E, then C is a convex set. In other words, any convex cone is a convex set. Qed.

Definition 2.0l. Let E be an $\mathbb{R}$-vector space. Let S be a subset of E. Then, the *convex conic hull* of S will denote the set

$$\left\{ x \in E \;\middle|\; \left(\begin{array}{c} \text{there exist some } t \in \mathbb{N}, \\ \text{a } t\text{-tuple } (x_1, x_2, \ldots, x_t) \text{ of elements of } S \\ \text{and a } t\text{-tuple } (\lambda_1, \lambda_2, \ldots, \lambda_t) \text{ of nonnegative reals} \\ \text{such that } \sum_{i=1}^{t} \lambda_i x_i = x \end{array} \right) \right\}.$$

We denote the convex conic hull of S by $\operatorname{cone} S$.

Definitions 2.0j, 2.0k and 2.0l are equivalent (according to Proposition 2.0m **(a)** below).

The following result is analogous to Proposition 2.0f (except for its part **(h)**, which is easy)[5]:

Proposition 2.0m. (a) Definitions 2.0j, 2.0k and 2.0l are equivalent.

Let E be an $\mathbb{R}$-vector space.

(b) We have[6] $\operatorname{cone} \varnothing = 0$ (not $\varnothing$).

Let now S be a subset of E.

(c) We have $S \subseteq \operatorname{cone} S$.

(d) The convex conic hull $\operatorname{cone} S$ is a convex cone.

(e) If T is a subset of S, then $\operatorname{cone} T \subseteq \operatorname{cone} S$.

(f) Every convex cone in E which contains S as a subset also contains $\operatorname{cone} S$ as a subset.

(g) If T is a subset of $\operatorname{cone} S$, then $\operatorname{cone} T \subseteq \operatorname{cone} S$.

(h) We have $\operatorname{conv.hull} S \subseteq \operatorname{cone} S$.

We refer (again) to Section 16 for a proof of Proposition 2.0m.

Remark. Let E be an $\mathbb{R}$-vector space. Let S be a subset of E. Then, the convex conic hull $\operatorname{cone} S$ is a convex cone (by Proposition 2.0m **(d)**) and contains S as a subset (by Proposition 2.0m **(c)**). Moreover, every convex cone in E which contains S as a subset also contains $\operatorname{cone} S$ as a subset (by Proposition 2.0m **(f)**). This result is often put into words as follows: "The convex conic hull $\operatorname{cone} S$ is the smallest convex cone containing S as a subset."

---

[5]And its proof (apart from part **(h)**) is analogous to the proof of Proposition 2.0f. For example, Proposition 2.0m **(f)** follows from Definition 2.0k in the same way as Proposition 2.0f **(f)** follows from Definition 2.0d.

[6]Here and in the following, we use the symbol "0" not just for the number zero and the zero vector in a vector space, but also for the zero subspace of any vector subspace. In other words, if E is any $\mathbb{R}$-vector space, then the subspace $\{0\}$ of E is also denoted by 0. Hopefully, this abuse of notation will sow no confusion.

The following proposition (which is, again, fundamental and will be used without explicit mention) is a simple consequence of Definition 2.0l:

Proposition 2.0n. Let E be an $\mathbb{R}$-vector space.

(a) If S is any subset of E, then

$$\operatorname{cone} S$$

$$= \left\{ \sum_{i=1}^{t} \lambda_i x_i \ \middle| \ \left(\begin{array}{c} t \text{ is an element of } \mathbb{N}, \\ \text{and } (x_1, x_2, \ldots, x_t) \text{ is a } t\text{-tuple of elements of } S, \\ \text{and } (\lambda_1, \lambda_2, \ldots, \lambda_t) \text{ is a } t\text{-tuple of nonnegative reals} \end{array} \right) \right\}.$$

(b) Let F be a vector subspace of E. Let S be a subset of F. Then, the convex conic hull $\operatorname{cone} S$ does not depend on whether we consider S as a subset of F or as a subset of E.

See Section 16 for a proof of Proposition 2.0n.

Here is yet another proposition (an analogue of Proposition 2.0h) that will be used tacitly:

Proposition 2.0o. Let E be an $\mathbb{R}$-vector space. Let x_1, x_2, ..., x_n be finitely many vectors in E.

(a) Then,

$$\operatorname{cone} \{x_1, x_2, \ldots, x_n\}$$
$$= \text{(the set of all linear combinations of the}$$
$$\text{vectors } x_1, x_2, \ldots, x_n \text{ with nonnegative coefficients}).$$

(b) Let $x \in E$. Then, we have $x \in \operatorname{cone} \{x_1, x_2, \ldots, x_n\}$ if and only if x is a linear combination of the vectors x_1, x_2, ..., x_n with nonnegative coefficients.

A proof of Proposition 2.0o can be found in Section 16.

The following proposition is nearly trivial; we state it merely for convenience:

Proposition 2.0p. Let E be an $\mathbb{R}$-vector space. Let S be a finite subset of E.

(a) Then,

$$\operatorname{cone} S = \left\{ \sum_{s \in S} \nu_s s \ \middle| \ (\nu_s)_{s \in S} \text{ is a family of nonnegative reals} \right\}.$$

In particular:

(b) If $(\nu_s)_{s \in S}$ is a family of nonnegative reals indexed by elements of S, then $\sum\limits_{s \in S} \nu_s s \in \operatorname{cone} S$.

(c) Conversely, if p is an element of $\operatorname{cone} S$, then there exists a family $(\nu_s)_{s \in S}$ of nonnegative reals indexed by elements of S such that $p = \sum\limits_{s \in S} \nu_s s$.

For the sake of completeness, the proof of Proposition 2.0p will be given in Section 16.

3. Notations and basics 3: On the dual space

Here are two further notations that we will be using:

- If E is any $\mathbb{R}$-vector space, then E^* will mean the $\mathbb{R}$-vector space $\operatorname{Hom}_{\mathbb{R}}(E, \mathbb{R})$ of all $\mathbb{R}$-linear maps from E to $\mathbb{R}$. This E^* is called the *dual space* of E.

- If $\ell \in \mathbb{N}$, then we regard the vectors in $\mathbb{R}^\ell$ as column vectors of length ℓ, and we regard the vectors in $\left(\mathbb{R}^\ell\right)^*$ as row vectors of length ℓ. More precisely, we identify every row vector $(a_1, a_2, \ldots, a_\ell) \in \mathbb{R}^\ell$ with the $\mathbb{R}$-linear map $f \in \left(\mathbb{R}^\ell\right)^* = \operatorname{Hom}_{\mathbb{R}}\left(\mathbb{R}^\ell, \mathbb{R}\right)$ which sends every column vector $\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_\ell \end{pmatrix} \in \mathbb{R}^\ell$ to the scalar $(a_1, a_2, \ldots, a_\ell) \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_\ell \end{pmatrix} = \sum\limits_{i=1}^{\ell} a_i b_i \in \mathbb{R}$. Thus, for every $f \in \left(\mathbb{R}^\ell\right)^*$ and $b \in \mathbb{R}^\ell$, we have $f(b) = fb$ (where f is regarded as an element of $\left(\mathbb{R}^\ell\right)^* = \operatorname{Hom}_{\mathbb{R}}\left(\mathbb{R}^\ell, \mathbb{R}\right)$ on the left hand side, and regarded as a row vector on the right hand side).

Let us state a simple property of convex hulls that we will be using several times:

Proposition 2.0r. Let E be an $\mathbb{R}$-vector space. Let $t \in \mathbb{N}$. Let x_1, x_2, ..., x_t be t vectors in E. Let $f \in E^*$ and $\delta \in \mathbb{R}$. Assume that

$$\text{every } i \in \{1, 2, \ldots, t\} \text{ satisfies } f(x_i) < \delta. \qquad (2)$$

Let $C = \operatorname{conv.hull}\{x_1, x_2, \ldots, x_t\}$. Then, every $x \in C$ satisfies $f(x) < \delta$.

The (rather easy) proof of Proposition 2.0r can be found in Section 16 below. Here is an analogous property of convex conic hulls:

Proposition 2.0s. Let E be an $\mathbb{R}$-vector space. Let S be a subset of E. Let $f \in E^*$. Assume that

$$\text{every } s \in S \text{ satisfies } f(s) \leq 0. \tag{3}$$

Then, every $x \in \operatorname{cone} S$ satisfies $f(x) \leq 0$.

Again, we refer to Section 16 for a proof of Proposition 2.0s.

Next, let us define the notion of a hyperplane:

Definition. Let E be an $\mathbb{R}$-vector space. If f is a nonzero element of E^*, and δ is an element of $\mathbb{R}$, then the subset $f^{-1}(\delta)$ of E will be called a *hyperplane* in E.

Note that this definition of a hyperplane is equivalent to the one given in [Schrij17] when $E = \mathbb{R}^n$ for some $n \in \mathbb{N}$. In fact, when $E = \mathbb{R}^n$, then any element $f \in E^*$ has the form

$$E \to \mathbb{R},$$
$$x \mapsto c^T x$$

for some $c \in \mathbb{R}^n$, and this c is nonzero if and only if f is nonzero.

Definition. Let E be an $\mathbb{R}$-vector space. Let H be a hyperplane in E. Let $z \in E$ and $C \subseteq E$. We say that the hyperplane H *separates* z and C if there exist a nonzero $f \in E^*$ and a $\delta \in \mathbb{R}$ such that $H = f^{-1}(\delta)$, $f(z) > \delta$ and (every $x \in C$ satisfies $f(x) < \delta$).

Note that this definition of "separate" is equivalent to the one given by Schrijver in [Schrij17] when $E = \mathbb{R}^n$ for some $n \in \mathbb{N}$. Schrijver's definition, however, has the disadvantage of using a topological notion (that of a "component"), which makes it difficult to generalize to other ordered fields instead of $\mathbb{R}$. It should be remarked that Schrijver never really uses his definition of "separate" in [Schrij17]; instead, he more or less uses my definition.

Furthermore, we introduce the relations $\geq$ and $\leq$ on vectors:

Definition. Let $\ell \in \mathbb{N}$. Let u and v be two column vectors in $\mathbb{R}^\ell$. Then, we write $u \leq v$ if and only if every $i \in \{1, 2, \ldots, \ell\}$ satisfies

$$(\text{the } i\text{-th coordinate of } u) \leq (\text{the } i\text{-th coordinate of } v).$$

Also, we write $u \geq v$ if and only if every $i \in \{1, 2, \ldots, \ell\}$ satisfies

$$(\text{the } i\text{-th coordinate of } u) \geq (\text{the } i\text{-th coordinate of } v).$$

The same notations apply if u and v are row vectors in $\left(\mathbb{R}^\ell\right)^*$ rather than column vectors in $\mathbb{R}^\ell$.

Of course, it is true that if u and v are two column vectors in $\mathbb{R}^\ell$ (or two row vectors in $\left(\mathbb{R}^\ell\right)^*$), then $u \geq v$ holds if and only if $v \leq u$. The relation $\leq$ is the smaller-or-equal relation of a partial order on $\mathbb{R}^\ell$ (namely, the so-called *componentwise partial order*), and the relation $\geq$ is the greater-or-equal relation of this order. However, this order is not a total order (unless $\ell \leq 1$): for example, if $u = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ and $v = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, then we have neither $u \geq v$ nor $v \geq u$.

Clearly, the relation $\leq$ respects the addition of vectors: If $u \in \mathbb{R}^\ell$, $v \in \mathbb{R}^\ell$ and $w \in \mathbb{R}^\ell$ are such that $u \leq v$, then $u + w \leq v + w$. Moreover, if $u \in \mathbb{R}^\ell$, $v \in \mathbb{R}^\ell$, $x \in \mathbb{R}^\ell$ and $y \in \mathbb{R}^\ell$ are such that $u \leq v$ and $x \leq y$, then $u + x \leq v + y$. As a consequence, two vectors $u \in \mathbb{R}^\ell$ and $v \in \mathbb{R}^\ell$ satisfy $u \leq v$ if and only if $u - v \leq 0$. Furthermore, if $x \in \mathbb{R}^\ell$ and $y \in \mathbb{R}^\ell$ are such that $x \leq y$, and if λ is a nonnegative real, then $\lambda x \leq \lambda y$. Similar rules hold for the relation $\geq$.

> **Definition.** Let $\ell \in \mathbb{N}$. A vector v lying either in $\mathbb{R}^\ell$ or in $\left(\mathbb{R}^\ell\right)^*$ is said to be *nonnegative* if it satisfies $v \geq 0$.

It is clear that a sum of nonnegative vectors in $\mathbb{R}^\ell$ is again nonnegative. More generally, any linear combination of nonnegative vectors in $\mathbb{R}^\ell$ with nonnegative coefficients is again nonnegative.

Let us state a trivial but important fact:

> **Lemma 2.0t.** Let $n \in \mathbb{N}$.
>
> **(a)** If a column vector $v \in \mathbb{R}^n$ satisfies $v \geq 0$, then all coordinates of the column vector v are nonnegative.
>
> **(b)** If a row vector $v \in (\mathbb{R}^n)^*$ satisfies $v \geq 0$, then all coordinates of the row vector v are nonnegative.

A proof of Lemma 2.0t can be found in Section 16. Of course, the converse of Lemma 2.0t also holds; we just won't use it often enough to have a reason to state it.

Here is a simple property of nonnegative vectors, which will not be directly used in these notes but which provides some context for them:

> **Lemma 2.0u.** Let $n \in \mathbb{N}$. Let $x \in \mathbb{R}^n$. Then, there exist two vectors y and z in $\mathbb{R}^n$ such that $y \geq 0$, $z \geq 0$ and $x = y - z$.

A proof of Lemma 2.0u can be found in Section 16.
Let us record another simple fact, which we will use many times:

> **Lemma 2.0v.** Let $n \in \mathbb{N}$. Let $x \in (\mathbb{R}^n)^*$ be a row vector such that $x \geq 0$. Let $y \in \mathbb{R}^n$ be a column vector such that $y \geq 0$. Then, $xy \geq 0$.

Again, we refer to Section 16 for the proof of this fact.

4. Closed intervals and [Schrij17, Theorem 2.3]

In this brief section, we shall state three simple lemmas about closed intervals and use them to "patch" the proof of [Schrij17, Theorem 2.3] to make this latter proof constructive. This section can be freely skipped, since none of what is done here will be used afterwards.

In the following, a *closed interval* will mean a set which has the form $\{x \in \mathbb{R} \mid a \leq x \leq b\}$ for some elements a and b of $\mathbb{R} \cup \{-\infty, \infty\}$ (where, of course, $-\infty$ is supposed to be smaller than each element of $\mathbb{R} \cup \{\infty\}$, and ∞ is supposed to be larger than each element of $\mathbb{R} \cup \{-\infty\}$). In particular, $\varnothing$ is a closed interval (since $\varnothing = \{x \in \mathbb{R} \mid 1 \leq x \leq 0\}$), and $\mathbb{R}$ is a closed interval (since $\mathbb{R} = \{x \in \mathbb{R} \mid -\infty \leq x \leq \infty\}$).

We will now state three simple lemmas, the proofs of which can all be found in Section 16.

> **Lemma 2.0x.** The intersection of finitely many closed intervals always is a closed interval.
>
> **Lemma 2.0y.** Let α and β be two reals. Then, the set $\{x \in \mathbb{R} \mid \alpha x \leq \beta\}$ is a closed interval.
>
> **Lemma 2.0z.** Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$. Let $b \in \mathbb{R}^m$. Let $A \in \mathbb{R}^{m \times n}$ be an $m \times n$-matrix. Let $P = \{x \in \mathbb{R}^n \mid Ax \leq b\}$.
>
> Let $z \in \mathbb{R}^n$ and $c \in \mathbb{R}^n$. Then, the set $\{\mu \in \mathbb{R} \mid z + \mu c \in P\}$ is a closed interval.

With the help of Lemma 2.0z, we can modify the proof of Theorem 2.3 in [Schrij17] in such a way that it no longer uses analysis:

Modifications to the proof of Theorem 2.3 in [Schrij17]. In the proof of Theorem 2.3 in [Schrij17], it is claimed that[7] the numbers

$$\mu_0 := \max \{\mu \mid z + \mu c \in P\} \qquad \text{and} \qquad \nu_0 := \max \{\nu \mid z - \nu c \in P\}$$

"exist since P is compact". We want to avoid this use of compactness. Instead, we will prove the existence of these numbers μ_0 and ν_0 as follows:

Alternative proof of the existence of the numbers $\mu_0 := \max \{\mu \mid z + \mu c \in P\}$ *and* $\nu_0 := \max \{\nu \mid z - \nu c \in P\}$ *in the proof of Theorem 2.3 in [Schrij17]:*

Lemma 2.0z shows that $\{\mu \in \mathbb{R} \mid z + \mu c \in P\}$ is a closed interval.

We also notice that $z + \underbrace{0c}_{=0} = z \in P$, so that $0 \in \{\mu \in \mathbb{R} \mid z + \mu c \in P\}$. Hence, the set $\{\mu \in \mathbb{R} \mid z + \mu c \in P\}$ is nonempty.

---

[7]We are using the notations of [Schrij17] here.

Finally, we notice that the set $\{\mu \in \mathbb{R} \mid z + \mu c \in P\}$ is bounded (from both sides)[8].

Now, a nonempty bounded closed interval always has a maximum. Applied to the closed interval $\{\mu \in \mathbb{R} \mid z + \mu c \in P\}$, this yields that $\{\mu \in \mathbb{R} \mid z + \mu c \in P\}$ has a maximum (since we know that $\{\mu \in \mathbb{R} \mid z + \mu c \in P\}$ is a nonempty bounded closed interval). In other words, $\max \{\mu \in \mathbb{R} \mid z + \mu c \in P\}$ exists. Since we defined μ_0 as $\max \{\mu \in \mathbb{R} \mid z + \mu c \in P\}$, this means that μ_0 exists.

We have thus shown that μ_0 exists. Similarly, ν_0 exists.

Also note that 0 is an element of the set $\{\mu \in \mathbb{R} \mid z + \mu c \in P\}$, whereas μ_0 is the maximum of this set (since $\mu_0 = \max \{\mu \in \mathbb{R} \mid z + \mu c \in P\}$). Since any element of a set is $\leq$ to the maximum of this set (if the maximum exists), this yields $0 \leq \mu_0$. Similarly, $0 \geq \nu_0$. This shows that z is a convex combination of the points $x := z + \mu_0 c$ and $y := z - \nu_0 c$. (This is used further below in the proof of Theorem 2.3.)

This completes the modifications necessary to make the proof of Theorem 2.3 in [Schrij17] independent of analysis. $\qquad\square$

5. The separation theorem for polytopes

We now come to the topic of separation theorems. In this section, we are only going to state them; for their proofs, we refer to Section 7 below.

Theorem 2.1 in [Schrij17] is a non-elementary fact; its proof cannot be ridden of analysis. This is not surprising: Theorem 2.1 is too general (it speaks of arbitrary convex sets). However, the following weaker version of Theorem 2.1 (which will still be enough for most of what we want) can be proven elementarily:

> **Theorem 2.1a.** Let C be a polytope in $\mathbb{R}^n$, and let $z \in \mathbb{R}^n$ be such that $z \notin C$. Then, there exists a hyperplane separating z and C.

Let us first give a basis-free version of this theorem:

---

[8]*Proof.* Since P is bounded, there exists an $M \in \mathbb{R}$ such that every $w \in P$ satisfies $|w| \leq M$. Consider this M.

Now, let ξ be an element of $\{\mu \in \mathbb{R} \mid z + \mu c \in P\}$. Then, $\xi \in \mathbb{R}$ and $z + \xi c \in P$. Recall that every $w \in P$ satisfies $|w| \leq M$. Applied to $w = z + \xi c$, this yields $|z + \xi c| \leq M$. But the triangle inequality yields $|z + \xi c| \geq \underbrace{|\xi c|}_{=|\xi||c|} - |z| = |\xi|\,|c| - |z|$. Thus, $M \geq |z + \xi c| \geq |\xi|\,|c| - |z|$, so that $|\xi|\,|c| \leq M + z$. Since $|c| > 0$ (because $c \neq 0$), we can divide this inequality by $|c|$ and obtain $|\xi| \leq \dfrac{M + z}{|c|}$.

Now, forget that we fixed ξ. We thus have proven that every $\xi \in \{\mu \in \mathbb{R} \mid z + \mu c \in P\}$ satisfies $|\xi| \leq \dfrac{M + z}{|c|}$. This yields that the set $\{\mu \in \mathbb{R} \mid z + \mu c \in P\}$ is bounded (from both sides), qed.

Theorem 2.1b. Let E be a finite-dimensional $\mathbb{R}$-vector space. Let C be a polytope in E, and let $z \in E$ be such that $z \notin C$. Then, there exist an $f \in E^*$ and a $\delta \in \mathbb{R}$ such that

$$(f(z) > \delta) \qquad \text{and} \qquad (\text{every } x \in C \text{ satisfies } f(x) < \delta).$$

Here is another way to rewrite this:

Theorem 2.1c. Let E be a finite-dimensional $\mathbb{R}$-vector space. Let C be a polytope in E. Then, exactly one of the following two assertions holds:

Assertion C1: We have $0 \in C$.

Assertion C2: There exists an $f \in E^*$ such that every $x \in C$ satisfies $f(x) < 0$.

We shall prove Theorem 2.1c, Theorem 2.1b and Theorem 2.1a (in this order) in Section 7.

Theorem 2.1a does not have the full strength of Theorem 2.1, but it is enough to replace many applications of Theorem 2.1. For example, in the proof of Theorem 2.4 in [Schrij17], Schrijver writes: "Suppose $x \notin P$. Then there exists a hyperplane separating x and P." This is (tacitly) being derived from Theorem 2.1, but it also follows from Theorem 2.1a (applied to $C = P$ and $z = x$).

6. The separation theorem for finitely generated cones

Next, we will show a separation theorem for finitely generated cones (more precisely, the "cone version" of Theorem 2.1c):

Theorem 2.5c. Let E be a finite-dimensional $\mathbb{R}$-vector space. Let S be a finite subset of E. Let $b \in E$. Then, exactly one of the following two assertions holds:

Assertion D1: We have $b \in \operatorname{cone} S$.

Assertion D2: There exists an $f \in E^*$ such that $f(b) > 0$ and (every $x \in \operatorname{cone} S$ satisfies $f(x) \leq 0$).

Again, we refer to Section 7 for the proof of this theorem.

7. Proofs of the separation theorems

We owe the reader four proofs now: the proofs of Theorems 2.1a, 2.1b, 2.1c and 2.5c. Now is the time to pay back this debt. We begin with a proof of Theorem 2.5c; the following neat proof I have learnt from [Bartl12] (which shows a more general result):[9]

First proof of Theorem 2.5c. We shall prove Theorem 2.5c by strong induction over $|S|$:

Induction step: Let n be a nonnegative integer. Assume that Theorem 2.5c holds whenever $|S| < n$. We will now prove that Theorem 2.5c holds whenever $|S| = n$.

So, let E be a finite-dimensional $\mathbb{R}$-vector space. Let S be a finite subset of E such that $|S| = n$. Let $b \in E$. Then, we must prove that exactly one of the following two assertions holds:

> *Assertion $D_1 1$:* We have $b \in \operatorname{cone} S$.
>
> *Assertion $D_1 2$:* There exists an $f \in E^*$ such that $f(b) > 0$ and (every $x \in \operatorname{cone} S$ satisfies $f(x) \leq 0$).

Since $\operatorname{cone} S$ is a convex cone, it is clear that $\operatorname{cone} S$ is closed under multiplication by a nonnegative scalar, and that any linear combination of finitely many elements of $\operatorname{cone} S$ with nonnegative coefficients must lie in $\operatorname{cone} S$. (These facts follow from general properties of convex cones.)

The Assertions $D_1 1$ and $D_1 2$ cannot hold at the same time[10]. Therefore, **at most** one of the two Assertions $D_1 1$ and $D_1 2$ holds. We will now show that at least one of these assertions holds.

If $n = 0$, then this is easy to prove[11]. Hence, for the rest of this proof, we WLOG assume that $n \neq 0$.

---

[9]A different (and much uglier) proof can be found in Section 17 below.

[10]*Proof.* Assume the opposite. Then, the Assertions $D_1 1$ and $D_1 2$ hold at the same time. Since Assertion $D_1 2$ holds, there exists an $f \in E^*$ such that $f(b) > 0$ and (every $x \in \operatorname{cone} S$ satisfies $f(x) \leq 0$). Consider this f. We know that every $x \in \operatorname{cone} S$ satisfies $f(x) \leq 0$. Since $b \in \operatorname{cone} S$ (because Assertion $D_1 1$ holds), we can apply this to $x = b$, and thus obtain $f(b) \leq 0$. But this contradicts $f(b) > 0$. This contradiction shows that our assumption was wrong, qed.

[11]*Proof.* Assume that $n = 0$. We must show that at least one of the two Assertions $D_1 1$ and $D_1 2$ holds.

We have $|S| = n = 0$, and thus $S = \varnothing$. Hence, $\operatorname{cone} S = \operatorname{cone} \varnothing = 0$. Therefore, $0 \in 0 = \operatorname{cone} S$.

If $b = 0$, then Assertion $D_1 1$ holds (since $b = 0 \in \operatorname{cone} S$). Hence, if $b = 0$, then at least one of the two Assertions $D_1 1$ and $D_1 2$ holds. Thus, for the rest of this proof, we WLOG assume that we don't have $b = 0$. Thus, $b \neq 0$.

Recall the following well-known fact from linear algebra: If v is a vector in a finite-dimensional $\mathbb{R}$-vector space V, and if $v \neq 0$, then there exists some $g \in V^*$ such that $g(v) \neq 0$.

Applying this fact to $V = E$ and $v = b$, we conclude that there exists a $g \in E^*$ such

---

The set S is nonempty (since $|S| = n \neq 0$). In other words, there exists some $t \in S$. Consider such a t.

From $t \in S$, we obtain $|S \setminus \{t\}| = \underbrace{|S|}_{=n} - 1 = n - 1 < n$. Hence, we can apply Theorem 2.5c to E, $S \setminus \{t\}$ and b instead of E, S and b (since we assumed that Theorem 2.5c holds whenever $|S| < n$), and conclude that exactly one of the following two assertions holds:

> *Assertion $D_2 1$:* We have $b \in \operatorname{cone}(S \setminus \{t\})$.
>
> *Assertion $D_2 2$:* There exists an $f \in E^*$ such that $f(b) > 0$ and (every $x \in \operatorname{cone}(S \setminus \{t\})$ satisfies $f(x) \leq 0$).

Thus, we must be in one of the following two cases:

Case 1: Assertion $D_2 1$ holds.

Case 2: Assertion $D_2 2$ holds.

First, let us consider Case 1. In this case, Assertion $D_2 1$ holds. In other words, we have $b \in \operatorname{cone}(S \setminus \{t\})$.

But $S \setminus \{t\} \subseteq S$ and therefore $\operatorname{cone}(S \setminus \{t\}) \subseteq \operatorname{cone} S$ (by Proposition 2.0m **(e)**, applied to $S \setminus \{t\}$ instead of T). Hence, $b \in \operatorname{cone}(S \setminus \{t\}) \subseteq \operatorname{cone} S$. In other words, Assertion $D_1 1$ holds. Hence, at least one of the two Assertions $D_1 1$ and $D_1 2$ holds. We have thus proven that at least one of Assertions $D_1 1$ and $D_1 2$ holds in Case 1.

Let us now consider Case 2. In this case, Assertion $D_2 2$ holds. In other words, there exists an $f \in E^*$ such that $f(b) > 0$ and (every $x \in \operatorname{cone}(S \setminus \{t\})$ satisfies $f(x) \leq 0$). Consider this f, and denote it by g. Thus, g is an element of E^* satisfying $g(b) > 0$ and

$$(\text{every } x \in \operatorname{cone}(S \setminus \{t\}) \text{ satisfies } g(x) \leq 0) \tag{4}$$

We must be in one of the following two subcases:

Subcase 2.1: We have $g(t) \leq 0$.

Subcase 2.2: We have $g(t) > 0$.

Let us first consider Subcase 2.1. In this subcase, we have $g(t) \leq 0$. Thus, every $x \in S$ satisfies $g(x) \leq 0$ [12]. Hence, every $x \in \operatorname{cone} S$ satisfies $g(x) \leq 0$

---

that $g(b) \neq 0$ (since $b \neq 0$). Consider this g. Define an $h \in E^*$ by $h = g(b) \cdot g$. Then, $h(b) = (g(b) \cdot g)(b) = g(b) g(b) = (g(b))^2 > 0$ (since $g(b) \neq 0$). On the other hand, every $x \in \operatorname{cone} S$ satisfies $h(x) \leq 0$ (since $x \in \operatorname{cone} S = 0$, so that $x = 0$, so that $h(x) = h(0) = 0$ (since h is linear)). Altogether, we have thus shown that $h(b) > 0$ and (every $x \in \operatorname{cone} S$ satisfies $h(x) \leq 0$).

Thus, we have proven that there exists an $f \in E^*$ such that $f(b) > 0$ and (every $x \in \operatorname{cone} S$ satisfies $f(x) \leq 0$) (namely, $f = h$). In other words, Assertion $D_1 2$ holds.

Thus, at least one of the two Assertions $D_1 1$ and $D_1 2$ holds. Qed.

[12] *Proof.* Let $x \in S$. We must show that $g(x) \leq 0$.

If $x = t$, then this follows immediately from $g\left(\underbrace{x}_{=t}\right) = g(t) \leq 0$. Hence, for the rest of

---

(by Proposition 2.0s, applied to $f = g$). Now, we know that g is an element of E^* satisfying $g(b) > 0$ and (every $x \in \text{cone } S$ satisfies $g(x) \leq 0$). Hence, there exists an $f \in E^*$ such that $f(b) > 0$ and (every $x \in \text{cone } S$ satisfies $f(x) \leq 0$) (namely, $f = g$). In other words, Assertion $D_1 2$ holds. Hence, at least one of the two Assertions $D_1 1$ and $D_1 2$ holds. We have thus proven that at least one of Assertions $D_1 1$ and $D_1 2$ holds in Subcase 2.1.

Let us now consider Subcase 2.2. In this subcase, we have $g(t) > 0$. Thus, the real number $\dfrac{1}{g(t)}$ is well-defined and satisfies $\dfrac{1}{g(t)} > 0$. In particular, $\dfrac{1}{g(t)}$ is a nonnegative real.

Define a vector $q \in E$ by $q = \dfrac{1}{g(t)} t$. Then, $g(q) = 1$ [13]. Also, $q \in \text{cone } S$ [14].

Define a subset S' of E by

$$S' = \{ s - g(s) q \ \mid \ s \in S \setminus \{t\} \}.$$

Thus, $|S'| \leq |S \setminus \{t\}| < n$. Therefore, we can apply Theorem 2.5c to E, S' and $b - g(b) q$ instead of E, S and b (since we assumed that Theorem 2.5c holds whenever $|S| < n$), and conclude that exactly one of the following two assertions holds:

> *Assertion $D_3 1$:* We have $b - g(b) q \in \text{cone}(S')$.
>
> *Assertion $D_3 2$:* There exists an $f \in E^*$ such that $f(b - g(b) q) > 0$ and (every $x \in \text{cone}(S')$ satisfies $f(x) \leq 0$).

Thus, we must be in one of the following two subsubcases:
Subsubcase 2.2.1: Assertion $D_3 1$ holds.

---

this proof, we can WLOG assume that we don't have $x = t$. Assume this.

We have $x \neq t$ (since we don't have $x = t$). Combining $x \in S$ with $x \neq t$, we find $x \in S \setminus \{t\} \subseteq \text{cone}(S \setminus \{t\})$ (by Proposition 2.0m **(c)**, applied to $S \setminus \{t\}$ instead of S). Hence, (4) shows that $g(x) \leq 0$. Qed.

[13] *Proof.* Applying the map g to the equality $q = \dfrac{1}{g(t)} t$, we obtain

$$g(q) = g\left(\frac{1}{g(t)} t \right) = \frac{1}{g(t)} g(t) \qquad \text{(since the map } g \text{ is } \mathbb{R}\text{-linear)}$$
$$= 1.$$

[14] *Proof.* We have $t \in S \subseteq \text{cone } S$ (by Proposition 2.0m **(c)**).

Recall that $\text{cone } S$ is closed under multiplication by a nonnegative scalar. In other words, every nonnegative real ω and every $y \in \text{cone } S$ satisfy $\omega y \in \text{cone } S$. Applying this to $\omega = \dfrac{1}{g(t)}$ and $y = t$, we obtain $\dfrac{1}{g(t)} t \in \text{cone } S$ (since $\dfrac{1}{g(t)}$ is a nonnegative real). Thus, $q = \dfrac{1}{g(t)} t \in \text{cone } S$.

Subsubcase 2.2.2: Assertion $D_3 2$ holds.

First, let us consider Subsubcase 2.2.1. In this subsubcase, Assertion $D_3 1$ holds. In other words, we have $b - g(b) q \in \text{cone}(S')$.

But $S' \subseteq \text{cone } S$ [15]. Hence, Proposition 2.0m **(g)** (applied to $T = S'$) yields $\text{cone}(S') \subseteq \text{cone } S$. Hence, $b - g(b) q \in \text{cone}(S') \subseteq \text{cone } S$.

Define $c \in E$ by $c = b - g(b) q$. Thus, $c = b - g(b) q \in \text{cone } S$.

Now, $g(b)$ is a nonnegative real (since $g(b) > 0$). Hence, 1 and $g(b)$ are nonnegative reals. Also, c and q are elements of $\text{cone } S$ (since $c \in \text{cone } S$ and $q \in \text{cone } S$). Therefore, $1c + g(b) q$ is a linear combination of finitely many elements of $\text{cone } S$ with nonnegative coefficients. Hence, $1c + g(b) q$ must lie in $\text{cone } S$ (since any linear combination of finitely many elements of $\text{cone } S$ with nonnegative coefficients must lie in $\text{cone } S$). In other words, $1c + g(b) q \in \text{cone } S$. Since $\underbrace{1c}_{=c=b-g(b)q} + g(b) q = (b - g(b) q) + g(b) q = b$, this rewrites as $b \in \text{cone } S$.

In other words, Assertion $D_1 1$ holds. Hence, at least one of the two Assertions $D_1 1$ and $D_1 2$ holds. We have thus proven that at least one of Assertions $D_1 1$ and $D_1 2$ holds in Subsubcase 2.2.1.

Next, let us consider Subsubcase 2.2.2. In this subsubcase, Assertion $D_3 2$ holds. In other words, there exists an $f \in E^*$ such that $f(b - g(b) q) > 0$ and (every $x \in \text{cone}(S')$ satisfies $f(x) \leq 0$). Consider this f, and denote it by h. Thus, h is an element of E^* satisfying $h(b - g(b) q) > 0$ and

$$\text{(every } x \in \text{cone}(S') \text{ satisfies } h(x) \leq 0). \tag{5}$$

Now, define a $k \in E^*$ by $k = h - h(q) \cdot g$. Then, each $x \in E$ satisfies

$$\underbrace{k}_{=h-h(q)\cdot g}(x) = (h - h(q) \cdot g)(x) = h(x) - h(q) \cdot g(x)$$

$$= h(x) - g(x) \cdot h(q) = h(x - g(x) q) \tag{6}$$

(since the map h is $\mathbb{R}$-linear). Applying this to $x = b$, we obtain $k(b) = h(b - g(b) q) > 0$. Moreover, each $x \in S$ satisfies $k(x) \leq 0$ [16]. Hence, each $x \in \text{cone } S$ satisfies $k(x) \leq 0$ (by Proposition 2.0s, applied to $f = k$).

---

[15]*Proof.* Let $s \in S \setminus \{t\}$. We shall show that $s - g(s) q \in \text{cone } S$.

Indeed, $s \in S \setminus \{t\} \subseteq \text{cone}(S \setminus \{t\})$ (by Proposition 2.0m **(c)**, applied to $S \setminus \{t\}$ instead of S). Hence, (4) (applied to $x = s$) yields $g(s) \leq 0$. Thus, $-g(s) \geq 0$. In other words, $-g(s)$ is nonnegative.

We have $s \in S \setminus \{t\} \subseteq S \subseteq \text{cone } S$ (by Proposition 2.0m **(c)**) and $q \in \text{cone } S$. Hence, $1s + (-g(s)) q$ is a linear combination of finitely many elements of $\text{cone } S$ with nonnegative coefficients (since $s \in \text{cone } S$ and $q \in \text{cone } S$, and since 1 and $-g(s)$ are nonnegative). Therefore, $1s + (-g(s)) q$ lies in $\text{cone } S$ (since any linear combination of finitely many elements of $\text{cone } S$ with nonnegative coefficients must lie in $\text{cone } S$). In other words, $1s + (-g(s)) q \in \text{cone } S$. Since $1s + (-g(s)) q = s - g(s) q$, this rewrites as $s - g(s) q \in \text{cone } S$.

Now, forget that we fixed s. We thus have shown that $s - g(s) q \in \text{cone } S$ for each $s \in S \setminus \{t\}$. In other words, $\{s - g(s) q \mid s \in S \setminus \{t\}\} \subseteq \text{cone } S$. Thus, $S' = \{s - g(s) q \mid s \in S \setminus \{t\}\} \subseteq \text{cone } S$.

[16]*Proof.* Let $x \in S$. We must prove that $k(x) \leq 0$.

Thus, there exists an $f \in E^*$ such that $f(b) > 0$ and (every $x \in \operatorname{cone} S$ satisfies $f(x) \leq 0$) (namely, $f = k$). In other words, Assertion $D_1 2$ holds. Hence, at least one of Assertions $D_1 1$ and $D_1 2$ holds.

We have thus proven that at least one of Assertions $D_1 1$ and $D_1 2$ holds in Subsubcase 2.2.2.

Recall that our goal was to prove that at least one of Assertions $D_1 1$ and $D_1 2$ holds. We have now proven this in Case 1, in Subcase 2.1, and in Subsubcases 2.2.1 and 2.2.2. Thus, we have proven this in all possible situations. So we know that at least one of Assertions $D_1 1$ and $D_1 2$ holds. Thus, exactly one of Assertions $D_1 1$ and $D_1 2$ holds (since we know that **at most** one of the two Assertions $D_1 1$ and $D_1 2$ holds).

Now, forget that we fixed E, S and b. We have thus proven that if E is a finite-dimensional $\mathbb{R}$-vector space, S is a finite subset of E such that $|S| = n$, and b is an element of E, then exactly one of the following two assertions holds:

> *Assertion $D_1 1$:* We have $b \in \operatorname{cone} S$.
>
> *Assertion $D_1 2$:* There exists an $f \in E^*$ such that $f(b) > 0$ and (every $x \in \operatorname{cone} S$ satisfies $f(x) \leq 0$).

In other words, we have proven that Theorem 2.5c holds in the case when $|S| = n$. This completes the induction step, and thus the induction proof of Theorem 2.5c is complete. $\qquad\square$

Before we step to the proof of Theorem 2.1c, let us state a basic result that connects convex conic hulls with convex hulls:

From $q = \dfrac{1}{g(t)} t$, we obtain $t = g(t) q$. But applying (6) to $x = t$, we obtain

$$k(t) = h\left(\underbrace{t}_{=g(t)q} - g(t) q \right) = h\left(\underbrace{g(t) q - g(t) q}_{=0} \right) = h(0) = 0$$

(since the map h is $\mathbb{R}$-linear). Therefore, $k(t) \leq 0$. Thus, if $x = t$, then $k\left(\underbrace{x}_{=t} \right) = k(t) \leq 0$. Hence, for the rest of our proof of $k(x) \leq 0$, we can WLOG assume that we don't have $x = t$. Assume this.

We have $x \neq t$ (since we don't have $x = t$). Combining $x \in S$ with $x \neq t$, we obtain $x \in S \setminus \{t\}$. Thus, $x - g(x) q \in \{s - g(s) q \mid s \in S \setminus \{t\}\}$ (since $x - g(x) q = s - g(s) q$ for some $s \in S \setminus \{t\}$ (namely, for $s = x$)). Hence,

$$x - g(x) q \in \{s - g(s) q \mid s \in S \setminus \{t\}\} = S' \subseteq \operatorname{cone}(S')$$

(by Proposition 2.0m **(c)**, applied to S' instead of S). Thus, (5) (applied to $x - g(x) q$ instead of x) yields $h(x - g(x) q) \leq 0$.

Now, (6) yields $k(x) = h(x - g(x) q) \leq 0$. Qed.

Proposition 2.4a. Let E be a finite-dimensional $\mathbb{R}$-vector space. Let T be a finite subset of E. Let $x \in E$.

Consider the direct sum $E \oplus \mathbb{R}$ (which is also a finite-dimensional $\mathbb{R}$-vector space). Consider the vector $(x, 1) \in E \oplus \mathbb{R}$. Let S be the subset $\{(y, 1) \mid y \in T\}$ of $E \oplus \mathbb{R}$.

Then, $x \in \operatorname{conv.hull} T$ holds if and only if $(x, 1) \in \operatorname{cone} S$.

Proof of Proposition 2.4a. The definition of S shows that $S = \{(y, 1) \mid y \in T\}$. Thus, the set S is finite (since the set T is finite).

Let us first show that

$$\text{if } x \in \operatorname{conv.hull} T, \text{ then } (x, 1) \in \operatorname{cone} S. \tag{7}$$

[*Proof of (7):* Assume that $x \in \operatorname{conv.hull} T$. We want to show that $(x, 1) \in \operatorname{cone} S$.

We have

$$x \in \operatorname{conv.hull} T$$
$$= \left\{ \sum_{i=1}^{t} \lambda_i x_i \;\middle|\; \left(\begin{array}{c} t \text{ is an element of } \mathbb{N}, \\ \text{and } (x_1, x_2, \ldots, x_t) \text{ is a } t\text{-tuple of elements of } T, \\ \text{and } (\lambda_1, \lambda_2, \ldots, \lambda_t) \text{ is a } t\text{-tuple of nonnegative reals} \\ \text{such that } \sum_{i=1}^{t} \lambda_i = 1 \end{array} \right) \right\}$$

(by Proposition 2.0g **(a)**, applied to T instead of S). In other words, there exist some element t of $\mathbb{N}$, some t-tuple $(x_1, x_2, \ldots, x_t)$ of elements of T, and some t-tuple $(\lambda_1, \lambda_2, \ldots, \lambda_t)$ of nonnegative reals such that $\sum_{i=1}^{t} \lambda_i = 1$ and $x = \sum_{i=1}^{t} \lambda_i x_i$. Consider this t, this $(x_1, x_2, \ldots, x_t)$ and this $(\lambda_1, \lambda_2, \ldots, \lambda_t)$. Clearly, the family $(\lambda_i)_{i \in \{1, 2, \ldots, t\}}$ is a family of nonnegative reals (since $(\lambda_1, \lambda_2, \ldots, \lambda_t)$ is a t-tuple of nonnegative reals) and has the property that all but finitely many $i \in \{1, 2, \ldots, t\}$ satisfy $\lambda_i = 0$ (since there are only finitely many $i \in \{1, 2, \ldots, t\}$).

Each $i \in \{1, 2, \ldots, t\}$ satisfies $(x_i, 1) \in S$ [17] and therefore $(x_i, 1) \in \operatorname{cone} S$ (since $(x_i, 1) \in S \subseteq \operatorname{cone} S$ (by Proposition 2.0m **(c)**, applied to $E \oplus \mathbb{R}$ instead of E)). Thus, $((x_i, 1))_{i \in \{1, 2, \ldots, t\}}$ is a family of elements of $\operatorname{cone} S$. Also, Proposition 2.0m **(d)** (applied to $E \oplus \mathbb{R}$ instead of E) shows that the convex conic hull $\operatorname{cone} S$ is a convex cone. Hence, Proposition 2.0i (applied to $E \oplus \mathbb{R}$, $\operatorname{cone} S$, $\{1, 2, \ldots, t\}$ and $(x_i, 1)$ instead of E, C, I and x_i) shows that $\sum_{i \in \{1, 2, \ldots, t\}} \lambda_i (x_i, 1) \in \operatorname{cone} S$.

---

[17] *Proof.* Let $i \in \{1, 2, \ldots, t\}$. Then, $x_i \in T$ (since $(x_1, x_2, \ldots, x_t)$ is a t-tuple of elements of T). Thus, $(x_i, 1) \in \{(y, 1) \mid y \in T\}$ (since $(x_i, 1) = (y, 1)$ for some $y \in T$ (namely, for $y = x_i$)). Hence, $(x_i, 1) \in \{(y, 1) \mid y \in T\} = S$. Qed.

Since

$$\underbrace{\sum_{i\in\{1,2,\ldots,t\}}}_{=\sum\limits_{i=1}^{t}} \lambda_i\,(x_i,1) = \sum_{i=1}^{t}\lambda_i\,(x_i,1) = \left(\sum_{i=1}^{t}\lambda_i x_i,\ \sum_{i=1}^{t}\underbrace{\lambda_i 1}_{=\lambda_i}\right)$$

$$= \left(\underbrace{\sum_{i=1}^{t}\lambda_i x_i}_{=x},\ \underbrace{\sum_{i=1}^{t}\lambda_i}_{=1}\right) = (x,1),$$

this rewrites as $(x,1)\in\operatorname{cone} S$. Thus, (7) is proven.]

Now, let us show that

$$\text{if } (x,1)\in\operatorname{cone} S, \text{ then } x\in\operatorname{conv.hull} T. \tag{8}$$

[*Proof of (8):* Assume that $(x,1)\in\operatorname{cone} S$. We want to show that $x\in\operatorname{conv.hull} T$.

Write the set T in the form $T = \{t_1,t_2,\ldots,t_n\}$ for some $n\in\mathbb{N}$. (This is possible, since the set T is finite.)

We have

$$S = \left\{(y,1) \ \middle| \ y\in\underbrace{T}_{=\{t_1,t_2,\ldots,t_n\}}\right\} = \{(y,1) \ | \ y\in\{t_1,t_2,\ldots,t_n\}\}$$
$$= \{(t_1,1),(t_2,1),\ldots,(t_n,1)\}.$$

We have

$$(x,1)\in\operatorname{cone}\underbrace{S}_{=\{(t_1,1),(t_2,1),\ldots,(t_n,1)\}} = \operatorname{cone}\{(t_1,1),(t_2,1),\ldots,(t_n,1)\}.$$

But Proposition 2.0o **(b)** (applied to $E\oplus\mathbb{R}$, $(t_i,1)$ and $(x,1)$ instead of E, x_i and x) shows that we have $(x,1)\in\operatorname{cone}\{(t_1,1),(t_2,1),\ldots,(t_n,1)\}$ if and only if $(x,1)$ is a linear combination of the vectors $(t_1,1),(t_2,1),\ldots,(t_n,1)$ with nonnegative coefficients. Therefore, $(x,1)$ is a linear combination of the vectors $(t_1,1),(t_2,1),\ldots,(t_n,1)$ with nonnegative coefficients (since $(x,1)\in\operatorname{cone}\{(t_1,1),(t_2,1),\ldots,(t_n,1)\}$). In other words, there exists an n-tuple $(\lambda_1,\lambda_2,\ldots,\lambda_n)$ of nonnegative reals such that

$$(x,1) = \sum_{i=1}^{n}\lambda_i\,(t_i,1).$$

Consider this $(\lambda_1,\lambda_2,\ldots,\lambda_n)$.

We have

$$(x, 1) = \sum_{i=1}^{n} \lambda_i (t_i, 1) = \left(\sum_{i=1}^{n} \lambda_i t_i, \sum_{i=1}^{n} \lambda_i 1 \right).$$

In other words,

$$x = \sum_{i=1}^{n} \lambda_i t_i \qquad \text{and} \qquad 1 = \sum_{i=1}^{n} \lambda_i 1.$$

Now, $1 = \underbrace{\sum_{i=1}^{n}}_{=\sum_{i \in \{1,2,\ldots,n\}}} \underbrace{\lambda_i 1}_{=\lambda_i} = \sum_{i \in \{1,2,\ldots,n\}} \lambda_i$, so that $\sum_{i \in \{1,2,\ldots,n\}} \lambda_i = 1$.

Clearly, the family $(\lambda_i)_{i \in \{1,2,\ldots,n\}}$ is a family of nonnegative reals (since $(\lambda_1, \lambda_2, \ldots, \lambda_n)$ is a n-tuple of nonnegative reals) and has the property that all but finitely many $i \in \{1, 2, \ldots, n\}$ satisfy $\lambda_i = 0$ (since there are only finitely many $i \in \{1, 2, \ldots, n\}$).

Each $i \in \{1, 2, \ldots, n\}$ satisfies $t_i \in T$ (since $t_i \in \{t_1, t_2, \ldots, t_n\} = T$) and therefore $t_i \in \text{conv.hull}\, T$ (since $t_i \in T \subseteq \text{conv.hull}\, T$ (by Proposition 2.0f **(c)**, applied to T instead of S)). Thus, $(t_i)_{i \in \{1,2,\ldots,n\}}$ is a family of elements of $\text{conv.hull}\, T$. Also, Proposition 2.0f **(d)** (applied to T instead of S) shows that the convex hull $\text{conv.hull}\, T$ is a convex set. Hence, Proposition 2.0a (applied to $\text{conv.hull}\, T$, $\{1, 2, \ldots, n\}$ and t_i instead of C, I and x_i) shows that $\sum_{i \in \{1,2,\ldots,n\}} \lambda_i t_i \in \text{conv.hull}\, T$. Since

$$\underbrace{\sum_{i \in \{1,2,\ldots,n\}}}_{=\sum_{i=1}^{n}} \lambda_i t_i = \sum_{i=1}^{n} \lambda_i t_i = x,$$

this rewrites as $x \in \text{conv.hull}\, T$. Thus, (8) is proven.]

Now, combining the two logical implications (7) and (8), we conclude that $x \in \text{conv.hull}\, T$ holds if and only if $(x, 1) \in \text{cone}\, S$. This proves Proposition 2.4a. $\qquad \square$

Next, we can prove Theorem 2.1c:[18]

First proof of Theorem 2.1c. The Assertions C1 and C2 cannot hold at the same time[19]. We will now show that at least one of these assertions holds.

---

[18]A different (and much uglier) proof can be found in Section 17 below.

[19]*Proof.* Assume the contrary. Thus, the Assertions C1 and C2 hold at the same time. Since Assertion C2 holds, there exists an $f \in E^*$ such that every $x \in C$ satisfies $f(x) < 0$. Consider this f. We know that every $x \in C$ satisfies $f(x) < 0$. Since $0 \in C$ (because Assertion C1 holds), we can apply this to $x = 0$, and thus obtain $f(0) < 0$. But this contradicts $f(0) = 0$ (which is because f is linear). This contradiction shows that our assumption was wrong, qed.

The set C is a polytope in E. In other words, C is the convex hull of a finite subset of E (by the definition of a "polytope"). In other words, there exists a finite subset T of E such that $C = \mathrm{conv\,.\,hull}\,T$. Consider this T.

Let $\iota : E \to E \oplus \mathbb{R}$ be the canonical inclusion of the $\mathbb{R}$-vector space E into the direct sum $E \oplus \mathbb{R}$. This inclusion ι maps each $x \in E$ to $(x, 0) \in E \oplus \mathbb{R}$; it is an $\mathbb{R}$-linear map.

Let S be the subset $\{(y, 1) \mid y \in T\}$ of $E \oplus \mathbb{R}$. Then, S is a finite subset of $E \oplus \mathbb{R}$ (since T is a finite set).

Let b be the vector $(0, 1) \in E \oplus \mathbb{R}$. Theorem 2.5c (applied to $E \oplus \mathbb{R}$ instead of E) then shows that exactly one of the following two assertions holds:

Assertion $C_1 1$: We have $b \in \mathrm{cone}\,S$.

Assertion $C_1 2$: There exists an $f \in (E \oplus \mathbb{R})^*$ such that $f(b) > 0$ and (every $x \in \mathrm{cone}\,S$ satisfies $f(x) \le 0$).

Thus, we are in one of the following two cases:
Case 1: Assertion $C_1 1$ holds.
Case 2: Assertion $C_1 2$ holds.

Let us first consider Case 1. In this case, Assertion $C_1 1$ holds. In other words, we have $b \in \mathrm{cone}\,S$. Thus, $(0, 1) = b \in \mathrm{cone}\,S$. Proposition 2.4a (applied to $x = 0$) shows that $0 \in \mathrm{conv\,.\,hull}\,T$ holds if and only if $(0, 1) \in \mathrm{cone}\,S$. Thus, $0 \in \mathrm{conv\,.\,hull}\,T$ holds (since $(0, 1) \in \mathrm{cone}\,S$). Hence, $0 \in \mathrm{conv\,.\,hull}\,T = C$. In other words, Assertion C1 holds. Thus, at least one of the two assertions C1 and C2 holds.

We thus have shown that in Case 1, at least one of the two assertions C1 and C2 holds.

Now, let us consider Case 2. In this case, Assertion $C_1 2$ holds. In other words, there exists an $f \in (E \oplus \mathbb{R})^*$ such that $f(b) > 0$ and (every $x \in \mathrm{cone}\,S$ satisfies $f(x) \le 0$). Consider such an f, and denote it by h. Thus, h is an element of $(E \oplus \mathbb{R})^*$ satisfying $h(b) > 0$ and

$$(\text{every } x \in \mathrm{cone}\,S \text{ satisfies } h(x) \le 0). \tag{9}$$

The map $h \circ \iota$ is a well-defined $\mathbb{R}$-linear map $E \to \mathbb{R}$ (since h is an $\mathbb{R}$-linear map $E \oplus \mathbb{R} \to \mathbb{R}$, and since ι is an $\mathbb{R}$-linear map $E \to E \oplus \mathbb{R}$). In other words, $h \circ \iota \in E^*$.

Now, let $x \in C$ be arbitrary. Thus, $x \in C = \mathrm{conv\,.\,hull}\,T$. But Proposition 2.4a shows that $x \in \mathrm{conv\,.\,hull}\,T$ holds if and only if $(x, 1) \in \mathrm{cone}\,S$. Hence, $(x, 1) \in \mathrm{cone}\,S$ (since $x \in \mathrm{conv\,.\,hull}\,T$). Therefore, (9) (applied to $(x, 1)$ instead of x) shows that $h((x, 1)) \le 0$.

The definition of ι yields $\iota(x) = (x, 0)$. Hence, $\underbrace{\iota(x)}_{=(x,0)} + \underbrace{b}_{=(0,1)} = (x, 0) + (0, 1) =$

$$\left(\underbrace{x + 0}_{=x}, \underbrace{0 + 1}_{=1} \right) = (x, 1).$$ Applying the map h to both sides of this equality, we find $h\left(\iota\left(x\right) + b\right) = h\left(\left(x, 1\right)\right) \leq 0.$

But the map h is $\mathbb{R}$-linear. Thus,

$$h\left(\iota\left(x\right) + b\right) = h\left(\iota\left(x\right)\right) + \underbrace{h\left(b\right)}_{>0} > h\left(\iota\left(x\right)\right) = \left(h \circ \iota\right)\left(x\right).$$

Hence, $\left(h \circ \iota\right)\left(x\right) < h\left(\iota\left(x\right) + b\right) \leq 0.$

Now, forget that we fixed x. We thus have shown that every $x \in C$ satisfies $\left(h \circ \iota\right)\left(x\right) < 0$. Hence, there exists an $f \in E^*$ such that every $x \in C$ satisfies $f\left(x\right) < 0$ (namely, $f = h \circ \iota$). In other words, Assertion C2 holds. Thus, at least one of the two assertions C1 and C2 holds.

We thus have shown that in Case 2, at least one of the two assertions C1 and C2 holds.

Thus, in each of the two Cases 1 and 2, we have proven that at least one of the two assertions C1 and C2 holds. Since these two Cases cover all possibilities, we therefore conclude that at least one of the two assertions C1 and C2 holds. Therefore, exactly one of the two assertions C1 and C2 holds (since we already know that the Assertions C1 and C2 cannot hold at the same time). This proves Theorem 2.1c. $\qquad\square$

Now, we are going to prove Theorem 2.1b and Theorem 2.1a:

Proof of Theorem 2.1b. We know that C is a polytope. By the definition of a polytope, this shows that C is the convex hull of a finite set of vectors in E. In other words, there exist some $t \in \mathbb{N}$ and some vectors x_1, x_2, ..., x_t in E such that $C = \operatorname{conv.hull}\left\{x_1, x_2, \ldots, x_t\right\}$. Consider this t and these x_1, x_2, ..., x_t.

Clearly, $\operatorname{conv.hull}\left\{x_1 - z, x_2 - z, \ldots, x_t - z\right\}$ is a polytope. Moreover,

$$0 \notin \operatorname{conv.hull}\left\{x_1 - z, x_2 - z, \ldots, x_t - z\right\}$$

[20].

Now, applying Theorem 2.1c to $\operatorname{conv.hull}\left\{x_1 - z, x_2 - z, \ldots, x_t - z\right\}$ instead of C, we obtain that exactly one of the following two assertions holds:

---

[20] *Proof.* Assume the opposite. Then, $0 \in \operatorname{conv.hull}\left\{x_1 - z, x_2 - z, \ldots, x_t - z\right\}$. But Proposition 2.0h **(b)** (applied to t, $x_i - z$ and 0 instead of n, x_i and x) shows that we have $0 \in \operatorname{conv.hull}\left\{x_1 - z, x_2 - z, \ldots, x_t - z\right\}$ if and only if 0 is a convex combination of the vectors $x_1 - z$, $x_2 - z$, ..., $x_t - z$. Hence, 0 is a convex combination of the vectors $x_1 - z$, $x_2 - z$, ..., $x_t - z$ (since we have $0 \in \operatorname{conv.hull}\left\{x_1 - z, x_2 - z, \ldots, x_t - z\right\}$). In other words, there exist nonnegative elements λ_1, λ_2, ..., λ_t of $\mathbb{R}$ such that $\sum_{i=1}^{t} \lambda_i = 1$ and $\sum_{i=1}^{t} \lambda_i\left(x_i - z\right) = 0$. Consider these λ_1, λ_2, ..., λ_t.

We have

$$0 = \sum_{i=1}^{t} \lambda_i\left(x_i - z\right) = \sum_{i=1}^{t} \lambda_i x_i - \underbrace{\sum_{i=1}^{t} \lambda_i}_{=1} z = \sum_{i=1}^{t} \lambda_i x_i - z,$$

Assertion $C_z 1$: We have $0 \in \text{conv.hull}\left\{x_1 - z, x_2 - z, \ldots, x_t - z\right\}$.

Assertion $C_z 2$: There exists an $f \in E^*$ such that every
$x \in \text{conv.hull}\left\{x_1 - z, x_2 - z, \ldots, x_t - z\right\}$ satisfies $f(x) < 0$.

Since Assertion $C_z 1$ cannot hold (because $0 \notin \text{conv.hull}\left\{x_1 - z, x_2 - z, \ldots, x_t - z\right\}$), this yields that Assertion $C_z 2$ must hold. In other words, there exists an $f \in E^*$ such that every $x \in \text{conv.hull}\left\{x_1 - z, x_2 - z, \ldots, x_t - z\right\}$ satisfies $f(x) < 0$. Consider this f.

Clearly, every $i \in \{1, 2, \ldots, t\}$ satisfies $f(z) - f(x_i) > 0$ [21]. Hence,

$$\left\{f(z) - f(x_1), f(z) - f(x_2), \ldots, f(z) - f(x_t)\right\}$$

is a set of positive reals. Since this set is finite, it must be bounded from below by a positive real (because any finite set of positive reals is bounded from below by a positive real). In other words, there exists a positive real ε such that every $i \in \{1, 2, \ldots, t\}$ satisfies $f(z) - f(x_i) \geq \varepsilon$. Consider this ε.

Since ε is positive, we have $\varepsilon > \dfrac{\varepsilon}{2}$.

Now, let $\delta = f(z) - \dfrac{\varepsilon}{2}$. Then,

$$\text{every } i \in \{1, 2, \ldots, t\} \text{ satisfies } f(x_i) < \delta \tag{10}$$

(since every $i \in \{1, 2, \ldots, t\}$ satisfies $f(z) - f(x_i) \geq \varepsilon$ and thus $f(x_i) \leq f(z) - \underbrace{\varepsilon}_{\substack{> \frac{\varepsilon}{2}}} < f(z) - \dfrac{\varepsilon}{2} = \delta$). Therefore, Proposition 2.0r shows that

$$\text{every } x \in C \text{ satisfies } f(x) < \delta. \tag{11}$$

Altogether, we thus know that $f(z) > \delta$ (since ε is positive, so that $\dfrac{\varepsilon}{2} > 0$, and thus $\delta = f(z) - \underbrace{\dfrac{\varepsilon}{2}}_{>0} < f(z)$) and that every $x \in C$ satisfies $f(x) < \delta$ (by (11)). We thus have shown that there exist an $f \in E^*$ and a $\delta \in \mathbb{R}$ such that

$$(f(z) > \delta) \qquad \text{and} \qquad (\text{every } x \in C \text{ satisfies } f(x) < \delta).$$

---

so that $z = \sum_{i=1}^{t} \lambda_i x_i$. Since $\lambda_1, \lambda_2, \ldots, \lambda_t$ are nonnegative elements of $\mathbb{R}$ such that $\sum_{i=1}^{t} \lambda_i = 1$, this yields that z is a convex combination of the vectors $x_1, x_2, \ldots, x_t$. In other words, $z \in \text{conv.hull}\{x_1, x_2, \ldots, x_t\}$. Hence, $z \in \text{conv.hull}\{x_1, x_2, \ldots, x_t\} = C$, which contradicts $z \notin C$. This contradiction shows that our assumption was wrong, qed.

[21] *Proof.* Let $i \in \{1, 2, \ldots, t\}$. Then, $x_i - z \in \{x_1 - z, x_2 - z, \ldots, x_t - z\} \subseteq \text{conv.hull}\{x_1 - z, x_2 - z, \ldots, x_t - z\}$ (by Proposition 2.0f **(c)**, applied to $\{x_1 - z, x_2 - z, \ldots, x_t - z\}$ instead of S). But we know that every $x \in \text{conv.hull}\{x_1 - z, x_2 - z, \ldots, x_t - z\}$ satisfies $f(x) < 0$. Applying this to $x = x_i - z$, we obtain $f(x_i - z) < 0$ (since $x_i - z \in \text{conv.hull}\{x_1 - z, x_2 - z, \ldots, x_t - z\}$). Since $f(x_i - z) = f(x_i) - f(z)$ (since f is linear), this rewrites as $f(x_i) - f(z) < 0$. In other words, $f(x_i) < f(z)$. Hence, $f(z) - f(x_i) > 0$, qed.

This proves Theorem 2.1b. □

Proof of Theorem 2.1a. Theorem 2.1b (applied to $E = \mathbb{R}^n$) shows that there exist an $f \in (\mathbb{R}^n)^*$ and a $\delta \in \mathbb{R}$ such that

$$(f(z) > \delta) \qquad \text{and} \qquad (\text{every } x \in C \text{ satisfies } f(x) < \delta).$$

Consider this f. Then, the hyperplane $f^{-1}(\delta) \subseteq \mathbb{R}^n$ clearly separates z and C. Thus, there exists a hyperplane separating z and C. This proves Theorem 2.1a. □

8. The Farkas lemma

Next, we are going to give a proof of [Schrij17, Theorem 2.5] without recourse to [Schrij17, Exercise 2.7]. This won't be very rich in substance, since the main work has already been done proving Theorems 2.1c and 2.5c.

First, we reformulate [Schrij17, Theorem 2.5]:

> **Theorem 2.5d.** Let $m \in \mathbb{N}$ and $n \in \mathbb{N}$. Let A be an $m \times n$-matrix and let $b \in \mathbb{R}^m$. Then, exactly one of the following two assertions holds:
>
> *Assertion F1:* The system $Ax = b$ has a nonnegative solution $x \in \mathbb{R}^n$.
>
> *Assertion F2:* There exists a vector $y \in \mathbb{R}^m$ such that $y^T A \geq 0$ and $y^T b < 0$.

In classical logic, Theorem 2.5d is equivalent to [Schrij17, Theorem 2.5], but constructively Theorem 2.5d is stronger.

Notice that Theorem 2.5d appears in [Jacim11, Teorema 1] (in an equivalent form) and in [Bartl12, Lemma 1] (in a generalized version).

Proof of Theorem 2.5d. The Assertions F1 and F2 cannot hold at the same time[22]. We will now show that at least one of these assertions holds.

Let a_1, a_2, ..., a_n be the columns of A. Let $S = \{a_1, a_2, \ldots, a_n\}$. Then, S is a finite subset of $\mathbb{R}^m$. Hence, Theorem 2.5c (applied to $E = \mathbb{R}^m$) yields that exactly one of the following two assertions holds:

---

[22]*Proof.* Assume the opposite. Then, the Assertions F1 and F2 hold at the same time. Since Assertion F2 holds, there exists a vector $y \in \mathbb{R}^m$ such that $y^T A \geq 0$ and $y^T b < 0$. Consider this y. Since Assertion F1 holds, the system $Ax = b$ has a nonnegative solution $x \in \mathbb{R}^n$. Consider this solution x. We have $x \geq 0$ (since x is nonnegative).

Lemma 2.0v (applied to $y^T A$ and x instead of x and y) shows that $y^T Ax \geq 0$ (since $y^T A \geq 0$ and $x \geq 0$). This contradicts $y^T \underbrace{Ax}_{=b} = y^T b < 0$. This contradiction shows that our assumption was wrong, qed.

Assertion $F_1 1$: We have $b \in \operatorname{cone} S$.

Assertion $F_2 2$: There exists an $f \in (\mathbb{R}^m)^*$ such that $f(b) > 0$ and (every $x \in \operatorname{cone} S$ satisfies $f(x) \leq 0$).

We must thus be in one of the following two cases:
Case 1: Assertion $F_1 1$ holds.
Case 2: Assertion $F_1 2$ holds.

Let us first consider Case 1. In this case, Assertion $F_1 1$ holds. In other words, $b \in \operatorname{cone} S$. Thus,

$$b \in \operatorname{cone} S = \left\{ \sum_{s \in S} \nu_s s \ \middle| \ (\nu_s)_{s \in S} \text{ is a family of nonnegative reals} \right\}$$

(by Proposition 2.0p **(a)**). Hence, there exists a family $(\nu_s)_{s \in S}$ of nonnegative reals such that $b = \sum\limits_{s \in S} \nu_s s$. Consider this family $(\nu_s)_{s \in S}$.

Let $(e_1, e_2, \ldots, e_n)$ be the standard basis of $\mathbb{R}^n$. In other words, for every $i \in \{1, 2, \ldots, n\}$, let e_i be the vector in $\mathbb{R}^n$ whose i-th coordinate is 1 and whose other coordinates are all 0. Then, for every $i \in \{1, 2, \ldots, n\}$, we have

$$Ae_i = (\text{the } i\text{-th column of } A) = a_i \tag{12}$$

(since the columns of A are a_1, a_2, ..., a_n). Also, for every $i \in \{1, 2, \ldots, n\}$, we have $e_i \geq 0$ (since every coordinate of e_i is either 1 or 0, and thus nonnegative).

Now, it is pretty clear that

for every $s \in S$, there exists a vector $x_s \in \mathbb{R}^n$ such that $x_s \geq 0$ and $Ax_s = s$. $\tag{13}$

[23] Consider this x_s.

Now, $\sum\limits_{s \in S} \nu_s x_s$ is a linear combination of nonnegative vectors (namely, the vectors x_s) with nonnegative coefficients (namely, the coefficients ν_s). Thus, $\sum\limits_{s \in S} \nu_s x_s$ must itself be a nonnegative vector (since any linear combination of nonnegative vectors with nonnegative coefficients must itself be a nonnegative vector). Since

$$A \left(\sum_{s \in S} \nu_s x_s \right) = \sum_{s \in S} \nu_s \underbrace{Ax_s}_{=s} = \sum_{s \in S} \nu_s s = b,$$

this yields that the system $Ax = b$ has a nonnegative solution $x \in \mathbb{R}^n$ (namely, $x = \sum\limits_{s \in S} \nu_s x_s$). In other words, Assertion F1 holds. Hence, at least one of Assertions F1 and F2 holds.

---

[23] *Proof of (13):* Let $s \in S$. Then, $s \in S = \{a_1, a_2, \ldots, a_n\}$. Hence, there exists some $i \in \{1, 2, \ldots, n\}$ such that $s = a_i$. Consider this i. Then, $s = a_i = Ae_i$ (by (12)). Thus, $e_i \geq 0$ and $Ae_i = s$. Thus, there exists a vector $x_s \in \mathbb{R}^n$ such that $x_s \geq 0$ and $Ax_s = s$ (namely, $x_s = e_i$). This proves (13).

We thus have proven that at least one of Assertions F1 and F2 holds in Case 1.

Now, let us consider Case 2. In this case, Assertion $F_1 2$ holds. In other words, there exists an $f \in (\mathbb{R}^m)^*$ such that $f(b) > 0$ and (every $x \in \text{cone } S$ satisfies $f(x) \leq 0$). Consider this f.

Let T denote the map

$$\mathbb{R}^m \to (\mathbb{R}^m)^*,$$
$$v \mapsto v^T.$$

Then, it is known that T is an isomorphism of $\mathbb{R}$-vector spaces. Hence, T^{-1} is well-defined. Let $z = T^{-1}(f)$. Then, $f = T(z) = z^T$ (by the definition of T). Thus, $f(b) = z^T b$. Hence, $z^T b = f(b) > 0$, so that $(-z)^T b = - \underbrace{z^T b}_{>0} < -0 = 0$.

Also, every $i \in \{1, 2, \ldots, n\}$ satisfies $(-z)^T a_i \geq 0$ [24]. Now, for every $i \in \{1, 2, \ldots, n\}$, the rule for multiplying a row vector by a matrix yields that

$$\left(\text{the } i\text{-th coordinate of } (-z)^T A \right)$$
$$= (-z)^T \cdot \underbrace{(\text{the } i\text{-th column of } A)}_{\substack{=a_i \\ \text{(since the columns of } A \text{ are } a_1, a_2, \ldots, a_n)}} = (-z)^T a_i$$
$$\geq 0 = (\text{the } i\text{-th coordinate of } 0).$$

In other words, $(-z)^T A \geq 0$.

We thus know that $(-z)^T A \geq 0$ and $(-z)^T b < 0$. Hence, there exists a vector $y \in \mathbb{R}^m$ such that $y^T A \geq 0$ and $y^T b < 0$ (namely, $y = -z$). In other words, Assertion F2 holds. Hence, at least one of Assertions F1 and F2 holds.

We thus have proven that at least one of Assertions F1 and F2 holds in Case 2.

Hence, in each of the Cases 1 and 2, at least one of Assertions F1 and F2 holds. Since these Cases 1 and 2 cover all possibilities, this yields that, in every situation, at least one of Assertions F1 and F2 holds. Since we know that the Assertions F1 and F2 cannot hold at the same time, this yields that exactly one of Assertions F1 and F2 holds. This proves Theorem 2.5d. $\square$

Thus, of course, [Schrij17, Theorem 2.5] is also proven.

---

[24] *Proof.* Let $i \in \{1, 2, \ldots, n\}$. Then, $a_i \in \{a_1, a_2, \ldots, a_n\} = S \subseteq \text{cone } S$ (by Proposition 2.0m **(c)**). But we know that every $x \in \text{cone } S$ satisfies $f(x) \leq 0$. Applying this to $x = a_i$, we obtain $f(a_i) \leq 0$ (since $a_i \in \text{cone } S$). Since $f = z^T$, this rewrites as $z^T a_i \leq 0$, so that $(-z)^T a_i = - \underbrace{z^T a_i}_{\leq 0} \geq -0 = 0$, qed.

9. The $<$ and $>$ relations for vectors

Let us now take a break from proving theorems, and introduce another piece of notation:

> **Definition.** Let $\ell \in \mathbb{N}$. Let u and v be two column vectors in $\mathbb{R}^\ell$. Then, we write $u < v$ if and only if every $i \in \{1, 2, \ldots, \ell\}$ satisfies
>
> $$(\text{the } i\text{-th coordinate of } u) < (\text{the } i\text{-th coordinate of } v).$$
>
> Also, we write $u > v$ if and only if every $i \in \{1, 2, \ldots, \ell\}$ satisfies
>
> $$(\text{the } i\text{-th coordinate of } u) > (\text{the } i\text{-th coordinate of } v).$$
>
> The same notations apply if u and v are row vectors in $\left(\mathbb{R}^\ell\right)^*$ rather than column vectors in $\mathbb{R}^\ell$.

Note that for a positive integer ℓ and two vectors $u \in \mathbb{R}^\ell$ and $v \in \mathbb{R}^\ell$, the assertion $(u > v)$ is **not** (in general) equivalent to the assertion $(u \geq v$ and $u \neq v)$, but is stronger. Similarly, the assertion $(u < v)$ is stronger than $(u \leq v$ and $u \neq v)$. For example, if $u = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $v = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, then we do have $(u \geq v$ and $u \neq v)$, but we don't have $u > v$.

Here is one further caveat: If 0_0 denotes the zero vector in $\mathbb{R}^0$, then $0_0 > 0_0$ and $0_0 < 0_0$. This is not surprising, since 0_0 has **no coordinates at all**.

The newly introduced notation has several properties which are similar to properties proven before. For example, here is an analogue of Lemma 2.0t:

> **Lemma 2.2a.** Let $n \in \mathbb{N}$.
>
> **(a)** If a column vector $v \in \mathbb{R}^n$ satisfies $v > 0$, then all coordinates of the column vector v are positive.
>
> **(b)** If a row vector $v \in (\mathbb{R}^n)^*$ satisfies $v > 0$, then all coordinates of the row vector v are positive.

A proof of Lemma 2.2a can be found in Section 16. Of course, the converse of Lemma 2.2a also holds; we just won't use it often enough to have a reason to state it.

Let us state three further facts, which also are proven in Section 16 below.

> **Lemma 2.2b.** Let $n \in \mathbb{N}$. Let $x \in (\mathbb{R}^n)^*$ be a nonzero row vector such that $x \geq 0$. Let $y \in \mathbb{R}^n$ be a column vector such that $y > 0$. Then, $xy > 0$.
>
> **Lemma 2.2c.** Let $n \in \mathbb{N}$. Let $x \in (\mathbb{R}^n)^*$ be a row vector such that $x > 0$. Let $y \in \mathbb{R}^n$ be a nonzero column vector such that $y \geq 0$. Then, $xy > 0$.
>
> **Lemma 2.2d.** Let $n \in \mathbb{N}$. Let $x \in (\mathbb{R}^n)^*$ be a row vector such that $x \geq 0$. Let $y \in \mathbb{R}^n$ be a column vector such that $y > 0$. Assume that $xy = 0$. Then, $x = 0$.

10. The Gordan and Stiemke theorems

Next, we will state two further important results.

> **Theorem 2.5e.** Let A be an $m \times n$-matrix. Then, exactly one of the following two assertions holds:
>
> *Assertion G1:* There exists a nonzero vector $x \in \mathbb{R}^n$ such that $x \geq 0$ and $Ax = 0$.
>
> *Assertion G2:* There exists a vector $y \in \mathbb{R}^m$ such that $y^T A > 0$.

Theorem 2.5e is known as *Gordan's theorem* and is equivalent to Exercise 2.17 in [Schrij17] in classical logic. In constructive logic, Theorem 2.5e is somewhat stronger than Exercise 2.17 in [Schrij17].

> **Theorem 2.5f.** Let A be an $m \times n$-matrix. Then, exactly one of the following two assertions holds:
>
> *Assertion S1:* There exists a vector $x \in \mathbb{R}^n$ such that $x > 0$ and $Ax = 0$.
>
> *Assertion S2:* There exists a vector $y \in \mathbb{R}^m$ such that $y^T A \geq 0$ and $y^T A \neq 0$.

Theorem 2.5f is known as *Stiemke's theorem* and is equivalent to Exercise 2.16 in [Schrij17] in classical logic. In constructive logic, Theorem 2.5f is somewhat stronger than Exercise 2.16 in [Schrij17].

Proof of Theorem 2.5e. The Assertions G1 and G2 cannot hold at the same time[25]. We will now show that at least one of these assertions holds.

Let a_1, a_2, ..., a_n be the columns of A. Let $C = \operatorname{conv.hull} \{a_1, a_2, \ldots, a_n\}$. Then, C is a polytope in $\mathbb{R}^m$. Hence, Theorem 2.1c (applied to $E = \mathbb{R}^m$) yields that exactly one of the following two assertions holds:

> *Assertion $G_1 1$:* We have $0 \in C$.
>
> *Assertion $G_1 2$:* There exists an $f \in (\mathbb{R}^m)^*$ such that every $x \in C$ satisfies $f(x) < 0$.

---

[25] *Proof.* Assume the opposite. Then, the Assertions G1 and G2 hold at the same time. Since Assertion G2 holds, there exists a vector $y \in \mathbb{R}^m$ such that $y^T A > 0$. Consider this y. Since Assertion G1 holds, there exists a nonzero vector $x \in \mathbb{R}^n$ such that $x \geq 0$ and $Ax = 0$. Consider this x.

Lemma 2.2c (applied to $y^T A$ and x instead of x and y) yields $y^T Ax > 0$ (since $y^T A > 0$ and $x \geq 0$ and since x is nonzero). This contradicts $y^T Ax = y^T \underbrace{Ax}_{=0} = 0$. This contradiction shows that our assumption was wrong, qed.

Thus, we must be in one of the following two cases:

Case 1: Assertion $G_1 1$ holds.

Case 2: Assertion $G_1 2$ holds.

Let us first consider Case 1. In this case, Assertion $G_1 1$ holds. In other words, $0 \in C$. Thus, $0 \in C = \mathrm{conv.\, hull}\,\{a_1, a_2, \ldots, a_n\}$.

Proposition 2.0h **(b)** (applied to $\mathbb{R}^m$, a_i and 0 instead of E, x_i and x) shows that we have $0 \in \mathrm{conv.\, hull}\,\{a_1, a_2, \ldots, a_n\}$ if and only if 0 is a convex combination of the vectors a_1, a_2, $\ldots$, a_n. Thus, 0 is a convex combination of the vectors a_1, a_2, $\ldots$, a_n (since we have $0 \in \mathrm{conv.\, hull}\,\{a_1, a_2, \ldots, a_n\}$). In other words, there exist n nonnegative elements λ_1, λ_2, $\ldots$, λ_n of $\mathbb{R}$ such that $\sum_{i=1}^{n} \lambda_i = 1$ and $\sum_{i=1}^{n} \lambda_i a_i = 0$. Consider these λ_1, λ_2, $\ldots$, λ_n.

Let u be the vector $\begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix} \in \mathbb{R}^n$. Then, the sum of the coordinates of the vector u is $\sum_{i=1}^{n} \lambda_i = 1 \neq 0$, while the sum of the coordinates of the vector 0 is 0.

Thus, $u \neq 0$. Also, $u \geq 0$ (because $u = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix}$, but every λ_i is nonnegative).

Moreover, by the definition of the product of a matrix with a column vector, we have

$$
Au = \sum_{i=1}^{n} \underbrace{(\text{the } i\text{-th column of } A)}_{\substack{=a_i \\ (\text{since the columns of } A \\ \text{are } a_1,\, a_2,\, \ldots,\, a_n)}} \cdot \underbrace{(\text{the } i\text{-th coordinate of } u)}_{\substack{=\lambda_i \\ (\text{since } u = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix})}} = \sum_{i=1}^{n} a_i \lambda_i
$$

$$
= \sum_{i=1}^{n} \lambda_i a_i = 0.
$$

Hence, u is nonzero, and satisfies $u \geq 0$ and $Au = 0$. Thus, there exists a nonzero vector $x \in \mathbb{R}^n$ such that $x \geq 0$ and $Ax = 0$ (namely, $x = u$). In other words, Assertion G1 holds. Hence, at least one of Assertions G1 and G2 holds.

We thus have proven that at least one of Assertions G1 and G2 holds in Case 1.

Next, let us consider Case 2. In this case, Assertion $G_1 2$ holds. In other words, there exists an $f \in (\mathbb{R}^m)^*$ such that

$$
\text{every } x \in C \text{ satisfies } f(x) < 0. \tag{14}
$$

Consider this f.

Let T denote the map

$$\mathbb{R}^m \to (\mathbb{R}^m)^*,$$
$$v \mapsto v^T.$$

Then, it is known that T is an isomorphism of $\mathbb{R}$-vector spaces. Hence, T^{-1} is well-defined. Let $z = T^{-1}(f)$. Then, $f = T(z) = z^T$ (by the definition of T).

Every $i \in \{1, 2, \ldots, n\}$ satisfies

$$a_i \in \{a_1, a_2, \ldots, a_n\} \subseteq \text{conv . hull}\{a_1, a_2, \ldots, a_n\}$$
$$\text{(by Proposition 2.0f } \textbf{(c)}, \text{ applied to } S = \{a_1, a_2, \ldots, a_n\})$$
$$= C$$

and thus $f(a_i) < 0$ (by (14), applied to $x = a_i$). Now, for every $i \in \{1, 2, \ldots, n\}$, we have

$$\left(\text{the } i\text{-th coordinate of the row vector } (-z)^T A\right)$$
$$= \underbrace{(-z)^T}_{\substack{= -z^T = -f \\ (\text{since } f = z^T)}} \cdot \underbrace{(\text{the } i\text{-th column of } A)}_{\substack{= a_i \\ (\text{since the columns of } A \\ \text{are } a_1, a_2, \ldots, a_n)}}$$
$$\text{(by the definition of the product of a row vector with a matrix)}$$
$$= (-f) \cdot a_i = -\underbrace{f(a_i)}_{<0} > -0 = 0 = (\text{the } i\text{-th coordinate of the row vector } 0).$$

Thus, $(-z)^T A > 0$. Hence, there exists a vector $y \in \mathbb{R}^m$ such that $y^T A > 0$ (namely, $y = -z$). In other words, Assertion G2 holds. Hence, at least one of Assertions G1 and G2 holds.

We thus have proven that at least one of Assertions G1 and G2 holds in Case 2.

Hence, in each of the Cases 1 and 2, at least one of Assertions G1 and G2 holds. Since these Cases 1 and 2 cover all possibilities, this yields that, in every situation, at least one of Assertions G1 and G2 holds. Since we know that the Assertions G1 and G2 cannot hold at the same time, this yields that exactly one of Assertions G1 and G2 holds. This proves Theorem 2.5e. $\square$

As we have seen, Theorem 2.5e is no more than a simple corollary of Theorem 2.1c. Theorem 2.5f is more interesting.

Proof of Theorem 2.5f. The Assertions S1 and S2 cannot hold at the same time[26]. We will now show that at least one of these assertions holds.

---

[26]*Proof.* Assume the opposite. Then, the Assertions S1 and S2 hold at the same time. Since Assertion S2 holds, there exists a vector $y \in \mathbb{R}^m$ such that $y^T A \geq 0$ and $y^T A \neq 0$. Consider

Let p be the vector $\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \in \mathbb{R}^n$. Clearly, $p > 0$.

Applying Theorem 2.5d to $b = -Ap$, we see that exactly one of the following two assertions holds:

> *Assertion $S_1$1:* The system $Ax = -Ap$ has a nonnegative solution $x \in \mathbb{R}^n$.

> *Assertion $S_1$2:* There exists a vector $y \in \mathbb{R}^m$ such that $y^T A \geq 0$ and $y^T (-Ap) < 0$.

Therefore, we must be in one of the following two cases:

Case 1: Assertion $S_1$1 holds.

Case 2: Assertion $S_1$2 holds.

Let us first consider Case 1. In this case, Assertion $S_1$1 holds. In other words, the system $Ax = -Ap$ has a nonnegative solution $x \in \mathbb{R}^n$. Let y be this solution. Then, $y \in \mathbb{R}^n$ is nonnegative and satisfies $Ay = -Ap$. Since $y \geq 0$ (because y is nonnegative) and $p > 0$, we have $\underbrace{y}_{\geq 0} + \underbrace{p}_{>0} > 0$. Also, $A(y + p) = \underbrace{Ay}_{=-Ap} + Ap = -Ap + Ap = 0$.

Hence, we know that $y + p > 0$ and $A(y + p) = 0$. Thus, there exists a vector $x \in \mathbb{R}^n$ such that $x > 0$ and $Ax = 0$ (namely, $x = y + p$). In other words, Assertion S1 holds. Hence, at least one of Assertions S1 and S2 holds.

We have thus proven that at least one of Assertions S1 and S2 holds in Case 1.

Now, let us consider Case 2. In this case, Assertion $S_1$2 holds. In other words, there exists a vector $y \in \mathbb{R}^m$ such that $y^T A \geq 0$ and $y^T (-Ap) < 0$. Denote this vector y by w. Then, $w \in \mathbb{R}^m$ satisfies $w^T A \geq 0$ and $w^T (-Ap) < 0$.

Since $-w^T Ap = w^T (-Ap) < 0$, we have $-w^T Ap \neq 0$, so that $w^T A \neq 0$.

We thus know that $w^T A \geq 0$ and $w^T A \neq 0$. Hence, there exists a vector $y \in \mathbb{R}^m$ such that $y^T A \geq 0$ and $y^T A \neq 0$ (namely, $y = w$). In other words, Assertion S2 holds. Hence, at least one of Assertions S1 and S2 holds.

We have thus proven that at least one of Assertions S1 and S2 holds in Case 2.

Hence, in each of the Cases 1 and 2, at least one of Assertions S1 and S2 holds. Since these Cases 1 and 2 cover all possibilities, this yields that, in every situation, at least one of Assertions S1 and S2 holds. Since we know that the

---

this y. Since Assertion S1 holds, there exists a vector $x \in \mathbb{R}^n$ such that $x > 0$ and $Ax = 0$. Consider this x.

The row vector $y^T A$ is nonzero (since $y^T A \neq 0$) and satisfies $y^T A \geq 0$. Hence, Lemma 2.2b (applied to $y^T A$ and x instead of x and y) shows that $y^T A x > 0$ (since $x > 0$). This contradicts $y^T A x = y^T \underbrace{Ax}_{=0} = 0$. This contradiction shows that our assumption was wrong, qed.

Assertions S1 and S2 cannot hold at the same time, this yields that exactly one of Assertions S1 and S2 holds. This proves Theorem 2.5f. □

11. Block matrices

Let us recall the following convention:

Convention 2.5g. **(a)** Let $u \in \mathbb{N}$ and $v \in \mathbb{N}$. Let n_1, n_2, …, n_v be nonnegative integers, and let m_1, m_2, …, m_u be nonnegative integers. For every $(i, j) \in \{1, 2, \ldots, u\} \times \{1, 2, \ldots, v\}$, let $A_{i,j}$ be an $m_i \times n_j$-matrix. Then, $\begin{pmatrix} A_{1,1} & A_{1,2} & \ldots & A_{1,v} \\ A_{2,1} & A_{2,2} & \ldots & A_{2,v} \\ \vdots & \vdots & \ddots & \vdots \\ A_{u,1} & A_{u,2} & \ldots & A_{u,v} \end{pmatrix}$ will not mean a $u \times v$-matrix whose entries themselves are matrices, but rather a **block matrix** of size $(m_1 + m_2 + \cdots + m_u) \times (n_1 + n_2 + \cdots + n_v)$ whose blocks are the matrices $A_{i,j}$.

(b) Let $n \in \mathbb{N}$. Column vectors in $\mathbb{R}^n$ are considered as $n \times 1$-matrices, and row vectors in $(\mathbb{R}^n)^*$ are considered as $1 \times n$-matrices. In particular, Convention 2.5g **(a)** also applies when some of the $A_{i,j}$ are column vectors or row vectors (as long as their sizes "fit"). For example, if x is the vector $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ and y is the vector $\begin{pmatrix} 3 \\ 7 \\ 5 \end{pmatrix}$, then $\begin{pmatrix} x \\ y \end{pmatrix}$ denotes the vector $\begin{pmatrix} 2 \\ 1 \\ 3 \\ 7 \\ 5 \end{pmatrix}$.

Let us state yet another trivial property of nonnegative vectors using Convention 2.5g:

Lemma 2.5h. Let $N \in \mathbb{N}$ and $M \in \mathbb{N}$. Let $x \in \mathbb{R}^N$ and $y \in \mathbb{R}^M$. Consider the block matrix $\begin{pmatrix} x \\ y \end{pmatrix}$; this is an $(N + M) \times 1$-matrix, i. e., a vector in $\mathbb{R}^{N+M}$ (since we identify matrices having only one column with column vectors).

(a) If $x \geq 0$ and $y \geq 0$, then $\begin{pmatrix} x \\ y \end{pmatrix} \geq 0$.

(b) If $\begin{pmatrix} x \\ y \end{pmatrix} \geq 0$, then $x \geq 0$ and $y \geq 0$.

See Section 16 for the proof of Lemma 2.5h.

The following simple lemma further illustrates the use of Convention 2.5g:

> **Lemma 2.5i.** Let $n \in \mathbb{N}$. Let $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^n$ be such that $\begin{pmatrix} x \\ -x \end{pmatrix} \geq \begin{pmatrix} y \\ -y \end{pmatrix}$ (where $\begin{pmatrix} x \\ -x \end{pmatrix}$ and $\begin{pmatrix} y \\ -y \end{pmatrix}$ are to be understood according to Convention 2.5g). Then, $x = y$.

We shall not use Lemma 2.5i, but a proof nevertheless is given in Section 16.

12. The Motzkin theorem

We will now prove another of the basic theorems of linear optimization theory:

> **Theorem 2.5k.** Let $n \in \mathbb{N}$, $m \in \mathbb{N}$ and $m' \in \mathbb{N}$. Let A be an $m \times n$-matrix. Let $b \in \mathbb{R}^m$. Let A' be an $m' \times n$-matrix. Let $b' \in \mathbb{R}^{m'}$. Then, exactly one of the following two assertions holds:
>
> *Assertion M1:* There exists a vector $x \in \mathbb{R}^n$ such that $Ax < b$ and $A'x \leq b'$.
>
> *Assertion M2:* There exist two vectors $y \in \mathbb{R}^m$ and $y' \in \mathbb{R}^{m'}$ such that $y \geq 0$, $y' \geq 0$, $y^T A + y'^T A' = 0$ and
>
> $$\left(\left(y^T b + y'^T b' < 0 \right) \text{ or } \left(y \neq 0 \text{ and } y^T b + y'^T b' \leq 0 \right) \right).$$

Theorem 2.5k is known as *Motzkin's theorem* and is equivalent to Exercise 2.19 in [Schrij17] in classical logic. In constructive logic, Theorem 2.5k is somewhat stronger than Exercise 2.19 in [Schrij17].

To prove Theorem 2.5k, we will first show:

> **Theorem 2.5l.** Let $n \in \mathbb{N}$, $n' \in \mathbb{N}$ and $m \in \mathbb{N}$. Let A be an $m \times n$-matrix. Let A' be an $m \times n'$-matrix. Then, exactly one of the following two assertions holds:
>
> *Assertion L1:* There exist two vectors $x \in \mathbb{R}^n$ and $x' \in \mathbb{R}^{n'}$ such that $x > 0$, $x' \geq 0$ and $Ax + A'x' = 0$.
>
> *Assertion L2:* There exists a vector $y \in \mathbb{R}^m$ such that $y^T A \geq 0$, $y^T A' \geq 0$ and $y^T A \neq 0$.

Proof of Theorem 2.5l. The Assertions L1 and L2 cannot hold at the same time[27]. We will now show that at least one of these assertions holds.

Consider the block matrix $\begin{pmatrix} A & A' \end{pmatrix}$; this is an $m \times (n + n')$-matrix.

Let p be the vector $\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \in \mathbb{R}^n$. Clearly, $p > 0$.

Applying Theorem 2.5d to $n + n'$, $\begin{pmatrix} A & A' \end{pmatrix}$ and $-Ap$ instead of n, A and b, we conclude that exactly one of the following two assertions holds:

> *Assertion $L_1 1$:* The system $\begin{pmatrix} A & A' \end{pmatrix} x = -Ap$ has a nonnegative solution $x \in \mathbb{R}^{n+n'}$.

> *Assertion $L_1 2$:* There exists a vector $y \in \mathbb{R}^m$ such that $y^T \begin{pmatrix} A & A' \end{pmatrix} \geq 0$ and $y^T (-Ap) < 0$.

Therefore, we must be in one of the following two cases:

Case 1: Assertion $L_1 1$ holds.

Case 2: Assertion $L_1 2$ holds.

Let us first consider Case 1. In this case, Assertion $L_1 1$ holds. In other words, the system $\begin{pmatrix} A & A' \end{pmatrix} x = -Ap$ has a nonnegative solution $x \in \mathbb{R}^{n+n'}$. Let ξ be this solution. Then, $\xi \in \mathbb{R}^{n+n'}$ is a nonnegative vector satisfying $\begin{pmatrix} A & A' \end{pmatrix} \xi = -Ap$.

Let us write the vector ξ in the form $\begin{pmatrix} u \\ v \end{pmatrix}$, where $u \in \mathbb{R}^n$ and $v \in \mathbb{R}^{n'}$.

We have $\begin{pmatrix} u \\ v \end{pmatrix} = \xi \geq 0$ (since ξ is a nonnegative vector). Hence, Lemma 2.5h **(b)** (applied to n, n', u and v instead of N, M, x and y) shows that $u \geq 0$ and $v \geq 0$.

---

[27] *Proof.* Assume the opposite. Then, the Assertions L1 and L2 hold at the same time. Since Assertion L2 holds, there exists a vector $y \in \mathbb{R}^m$ such that $y^T A \geq 0$, $y^T A' \geq 0$ and $y^T A \neq 0$. Consider this y.

Since Assertion L1 holds, there exist two vectors $x \in \mathbb{R}^n$ and $x' \in \mathbb{R}^{n'}$ such that $x > 0$, $x' \geq 0$ and $Ax + A'x' = 0$. Consider these two vectors x and x'.

From $x > 0$, we conclude that $x \geq 0$.

Applying Lemma 2.0v to $y^T A$ and x instead of x and y, we obtain $y^T Ax \geq 0$. Applying Lemma 2.0v to n', $y^T A'$ and x' instead of n, x and y, we obtain $y^T A'x' \geq 0$.

If we had $y^T Ax = 0$, then Lemma 2.2d (applied to $y^T A$ and x instead of x and y) would yield $y^T A = 0$, contradicting $y^T A \neq 0$. Thus, we cannot have $y^T Ax = 0$. Hence, $y^T Ax \neq 0$, so that $y^T Ax > 0$ (since $y^T Ax \geq 0$).

We have $y^T \underbrace{(Ax + A'x')}_{=0} = 0$, thus

$$0 = y^T (Ax + A'x') = \underbrace{y^T Ax}_{>0} + \underbrace{y^T A'x'}_{\geq 0} > 0.$$

This is absurd. This contradiction shows that our assumption was wrong, qed.

Since $\xi = \begin{pmatrix} u \\ v \end{pmatrix}$, we have

$$\begin{pmatrix} A & A' \end{pmatrix} \xi = \begin{pmatrix} A & A' \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = Au + A'v$$

$$\text{(by the multiplication rule for block matrices)}.$$

Compared with $\begin{pmatrix} A & A' \end{pmatrix} \xi = -Ap$, this yields $-Ap = Au + A'v$. Thus, $0 = \underbrace{Ap + Au}_{=A(p+u)} + A'v = A(p+u) + A'v$. Also, $\underbrace{p}_{>0} + \underbrace{u}_{\geq 0} > 0$.

Altogether, we know that $p + u > 0$, $v \geq 0$ and $A(p+u) + A'v = 0$. Thus, there exist two vectors $x \in \mathbb{R}^n$ and $x' \in \mathbb{R}^{n'}$ such that $x > 0$, $x' \geq 0$ and $Ax + A'x' = 0$ (namely, $x = p + u$ and $x' = v$). In other words, Assertion L1 holds. Hence, at least one of Assertions L1 and L2 holds.

We have thus proven that at least one of Assertions L1 and L2 holds in Case 1.

Now, let us consider Case 2. In this case, Assertion $L_1 2$ holds. In other words, there exists a vector $y \in \mathbb{R}^m$ such that $y^T \begin{pmatrix} A & A' \end{pmatrix} \geq 0$ and $y^T (-Ap) < 0$. Denote this y by w. Then, $w \in \mathbb{R}^m$ satisfies $w^T \begin{pmatrix} A & A' \end{pmatrix} \geq 0$ and $w^T (-Ap) < 0$.

Since $w^T (-Ap) < 0$, we have $w^T (-Ap) \neq 0$ and thus $w^T A (-p) = w^T (-Ap) \neq 0$, hence $w^T A \neq 0$.

We have $w^T \begin{pmatrix} A & A' \end{pmatrix} \geq 0$. Since $w^T \begin{pmatrix} A & A' \end{pmatrix} = \begin{pmatrix} w^T A & w^T A' \end{pmatrix}$ (by the multiplication rule for block matrices), this rewrites as $\begin{pmatrix} w^T A & w^T A' \end{pmatrix} \geq 0$. Thus, $w^T A \geq 0$ and $w^T A' \geq 0$ (by an analogue of Lemma 2.5h **(b)** for block matrices of the form $\begin{pmatrix} x & y \end{pmatrix}$ instead of $\begin{pmatrix} x \\ y \end{pmatrix}$).

So we have proven that $w^T A \geq 0$, $w^T A' \geq 0$ and $w^T A \neq 0$. Thus, there exists a vector $y \in \mathbb{R}^m$ such that $y^T A \geq 0$, $y^T A' \geq 0$ and $y^T A \neq 0$ (namely, $y = w$). In other words, Assertion L2 holds. Hence, at least one of Assertions L1 and L2 holds.

We have thus proven that at least one of Assertions L1 and L2 holds in Case 2.

Hence, in each of the Cases 1 and 2, at least one of Assertions L1 and L2 holds. Since these Cases 1 and 2 cover all possibilities, this yields that, in every situation, at least one of Assertions L1 and L2 holds. Since we know that the Assertions L1 and L2 cannot hold at the same time, this yields that exactly one of Assertions L1 and L2 holds. This proves Theorem 2.5l. $\qquad \square$

Proof of Theorem 2.5k. The Assertions M1 and M2 cannot hold at the same time[28]. We will now show that at least one of these assertions holds.

For any nonnegative integers α and β, let $0_{\alpha,\beta}$ denote the $\alpha \times \beta$ zero matrix. For any nonnegative integer γ, let I_γ denote the $\gamma \times \gamma$ identity matrix.

---

[28]*Proof.* Assume the opposite. Then, the Assertions M1 and M2 hold at the same time. Since Assertion M2 holds, there exist two vectors $y \in \mathbb{R}^m$ and $y' \in \mathbb{R}^{m'}$ such that $y \geq 0$, $y' \geq 0$,

Consider the block matrix $\begin{pmatrix} I_m & -b \\ 0_{m',m} & -b' \end{pmatrix}$; this is a $(m+m') \times (m+1)$-matrix.

Also, consider the block matrix $\begin{pmatrix} 0_{m,m'} & A & -A \\ I_{m'} & A' & -A' \end{pmatrix}$; this is a $(m+m') \times (m'+2n)$-matrix.

Applying Theorem 2.5l to $m+m'$, $m+1$, $m'+2n$, $\begin{pmatrix} I_m & -b \\ 0_{m',m} & -b' \end{pmatrix}$ and $\begin{pmatrix} 0_{m,m'} & A & -A \\ I_{m'} & A' & -A' \end{pmatrix}$ instead of m, n, n', A and A', we conclude that exactly one of the following two assertions holds:

$y^T A + y'^T A' = 0$ and

$$\left((y^T b + y'^T b' < 0) \text{ or } (y \neq 0 \text{ and } y^T b + y'^T b' \leq 0) \right).$$

Consider these two vectors y and y'.

Since Assertion M1 holds, there exists a vector $x \in \mathbb{R}^n$ such that $Ax < b$ and $A'x \leq b'$. Consider this x.

From $Ax < b$, we obtain $b > Ax$, so that $b - Ax > 0$. Thus, $b - Ax \geq 0$. Also, $y \geq 0$, so that $y^T \geq 0$ (since the transpose of any nonnegative vector is nonnegative). Now, applying Lemma 2.0v to m, y^T and $b - Ax$ instead of n, x and y, we obtain $y^T (b - Ax) \geq 0$. Thus, $y^T b - y^T Ax = y^T (b - Ax) \geq 0$, so that $y^T b \geq y^T Ax$.

From $A'x \leq b'$, we obtain $b' \geq A'x$, so that $b' - A'x \geq 0$. Also, $y' \geq 0$, so that $y'^T \geq 0$ (since the transpose of any nonnegative vector is nonnegative). Now, applying Lemma 2.0v to m', y'^T and $b' - A'x$ instead of n, x and y, we obtain $y'^T (b' - A'x) \geq 0$. Thus, $y'^T b' - y'^T A'x = y'^T (b' - A'x) \geq 0$, so that $y'^T b' \geq y'^T A'x$.

Now, recall that we have $\left(y^T b + y'^T b' < 0 \right)$ or $\left(y \neq 0 \text{ and } y^T b + y'^T b' \leq 0 \right)$. Hence, we must be in one of the following two cases:

Case 1: We have $y^T b + y'^T b' < 0$.

Case 2: We have $y \neq 0$ and $y^T b + y'^T b' \leq 0$.

Let us first consider Case 1. In this case, $y^T b + y'^T b' < 0$. Thus,

$$0 > \underbrace{y^T b}_{\geq y^T Ax} + \underbrace{y'^T b'}_{\geq y'^T A'x} \geq y^T Ax + y'^T A'x = \underbrace{\left(y^T A + y'^T A' \right)}_{=0} x = 0.$$

This is absurd. Thus, we have obtained a contradiction in Case 1.

Let us now consider Case 2. In this case, $y \neq 0$ and $y^T b + y'^T b' \leq 0$. From $y \neq 0$, we obtain $y^T \neq 0$. If we had $y^T (b - Ax) = 0$, then Lemma 2.2d (applied to m, y^T and $b - Ax$ instead of n, x and y) would yield $y^T = 0$ (since $y^T \geq 0$, $y^T (b - Ax) = 0$ and $b - Ax > 0$), contradicting $y^T \neq 0$. Hence, we cannot have $y^T (b - Ax) = 0$. In other words, $y^T (b - Ax) \neq 0$. Combined with $y^T (b - Ax) \geq 0$, this yields $y^T (b - Ax) > 0$. Hence, $y^T b - y^T Ax = y^T (b - Ax) > 0$, so that $y^T b > y^T Ax$.

Now, from $y^T b + y'^T b' \leq 0$, we obtain

$$0 \geq \underbrace{y^T b}_{> y^T Ax} + \underbrace{y'^T b'}_{\geq y'^T A'x} > y^T Ax + y'^T A'x = \underbrace{\left(y^T A + y'^T A' \right)}_{=0} x = 0.$$

This is absurd. Thus, we have obtained a contradiction in Case 2.

Hence, in each of the cases 1 and 2, we have obtained a contradiction. Since cases 1 and 2 cover all possibilities, this yields that we have a contradiction in any situation. Thus, our assumption was wrong, qed.

Assertion $M_1 1$: There exist two vectors $x \in \mathbb{R}^{m+1}$ and $x' \in \mathbb{R}^{m'+2n}$ such that $x > 0$, $x' \geq 0$ and $\begin{pmatrix} I_m & -b \\ 0_{m',m} & -b' \end{pmatrix} x + \begin{pmatrix} 0_{m,m'} & A & -A \\ I_{m'} & A' & -A' \end{pmatrix} x' = 0$.

Assertion $M_1 2$: There exists a vector $y \in \mathbb{R}^{m+m'}$ such that $y^T \begin{pmatrix} I_m & -b \\ 0_{m',m} & -b' \end{pmatrix} \geq 0$, $y^T \begin{pmatrix} 0_{m,m'} & A & -A \\ I_{m'} & A' & -A' \end{pmatrix} \geq 0$ and $y^T \begin{pmatrix} I_m & -b \\ 0_{m',m} & -b' \end{pmatrix} \neq 0$.

Therefore, we must be in one of the following two cases:

Case 1: Assertion $M_1 1$ holds.

Case 2: Assertion $M_1 2$ holds.

Let us first consider Case 1. In this case, Assertion $M_1 1$ holds. In other words, there exist two vectors $x \in \mathbb{R}^{m+1}$ and $x' \in \mathbb{R}^{m'+2n}$ such that $x > 0$, $x' \geq 0$ and $\begin{pmatrix} I_m & -b \\ 0_{m',m} & -b' \end{pmatrix} x + \begin{pmatrix} 0_{m,m'} & A & -A \\ I_{m'} & A' & -A' \end{pmatrix} x' = 0$. Denote these vectors x and x' by ξ and ξ', respectively. Then, $\xi \in \mathbb{R}^{m+1}$ and $\xi' \in \mathbb{R}^{m'+2n}$ satisfy $\xi > 0$, $\xi' \geq 0$ and $\begin{pmatrix} I_m & -b \\ 0_{m',m} & -b' \end{pmatrix} \xi + \begin{pmatrix} 0_{m,m'} & A & -A \\ I_{m'} & A' & -A' \end{pmatrix} \xi' = 0$.

Denote the $(m+1)$-th coordinate of the vector ξ by t. Then, $t > 0$ (since $\xi > 0$). Hence, $t \neq 0$, so that t is invertible, and $\frac{1}{t} > 0$ (since $t > 0$).

Since ξ is a vector in $\mathbb{R}^{m+1}$ whose $(m+1)$-th coordinate is t, we can write ξ in the form $\xi = \begin{pmatrix} u \\ t \end{pmatrix}$ for some $u \in \mathbb{R}^m$. Consider this u. Then, $u > 0$ (since $\xi > 0$).

Write the vector $\xi' \in \mathbb{R}^{m'+2n}$ in the form $\begin{pmatrix} v \\ w \\ z \end{pmatrix}$, where $v \in \mathbb{R}^{m'}$, $w \in \mathbb{R}^n$ and $z \in \mathbb{R}^n$. Then, from $\xi' \geq 0$, we obtain $v \geq 0$, $w \geq 0$ and $z \geq 0$.

But

$$0 = \begin{pmatrix} I_m & -b \\ 0_{m',m} & -b' \end{pmatrix} \underbrace{\xi}_{=\begin{pmatrix} u \\ t \end{pmatrix}} + \begin{pmatrix} 0_{m,m'} & A & -A \\ I_{m'} & A' & -A' \end{pmatrix} \underbrace{\xi'}_{=\begin{pmatrix} v \\ w \\ z \end{pmatrix}}$$

$$= \underbrace{\begin{pmatrix} I_m & -b \\ 0_{m',m} & -b' \end{pmatrix}\begin{pmatrix} u \\ t \end{pmatrix}}_{=\begin{pmatrix} I_m u + (-b)t \\ 0_{m',m}u + (-b')t \end{pmatrix} \atop \text{(by the multiplication rule for block matrices)}} + \underbrace{\begin{pmatrix} 0_{m,m'} & A & -A \\ I_{m'} & A' & -A' \end{pmatrix}\begin{pmatrix} v \\ w \\ z \end{pmatrix}}_{=\begin{pmatrix} 0_{m,m'}v + Aw + (-A)z \\ I_{m'}v + A'w + (-A')z \end{pmatrix} \atop \text{(by the multiplication rule for block matrices)}}$$

$$= \begin{pmatrix} I_m u + (-b)t \\ 0_{m',m}u + (-b')t \end{pmatrix} + \begin{pmatrix} 0_{m,m'}v + Aw + (-A)z \\ I_{m'}v + A'w + (-A')z \end{pmatrix}$$

$$= \begin{pmatrix} I_m u + (-b)t + 0_{m,m'}v + Aw + (-A)z \\ 0_{m',m}u + (-b')t + I_{m'}v + A'w + (-A')z \end{pmatrix}.$$

Hence, $\begin{pmatrix} I_m u + (-b)t + 0_{m,m'}v + Aw + (-A)z \\ 0_{m',m}u + (-b')t + I_{m'}v + A'w + (-A')z \end{pmatrix} = 0$, so that $I_m u + (-b)t + 0_{m,m'}v + Aw + (-A)z = 0$ and $0_{m',m}u + (-b')t + I_{m'}v + A'w + (-A')z = 0$.

Thus,

$$0 = \underbrace{I_m u}_{=u>0} + \underbrace{(-b)t}_{=-bt} + \underbrace{0_{m,m'}v}_{=0} + Aw + \underbrace{(-A)z}_{=-Az} > -bt + \underbrace{Aw - Az}_{=A(w-z)} = -bt + A(w-z).$$

Since $t > 0$, we can divide this inequality by t, and obtain $0 > -b + \frac{1}{t}A(w-z) = -b + A \cdot \frac{1}{t}(w-z)$. Thus, $b > A \cdot \frac{1}{t}(w-z)$, so that $A \cdot \frac{1}{t}(w-z) < b$.

On the other hand,

$$0 = \underbrace{0_{m',m}u}_{=0} + \underbrace{(-b')t}_{=-b't} + \underbrace{I_{m'}v}_{=v\geq 0} + A'w + \underbrace{(-A')z}_{=-A'z} \geq -b't + \underbrace{A'w - A'z}_{=A'(w-z)} = -b't + A'(w-z).$$

Since $t > 0$, we can divide this inequality by t, and obtain $0 \geq -b' + \frac{1}{t}A'(w-z) = -b' + A' \cdot \frac{1}{t}(w-z)$. Thus, $b' \geq A' \cdot \frac{1}{t}(w-z)$, so that $A' \cdot \frac{1}{t}(w-z) \leq b'$.

So we know that $A \cdot \frac{1}{t}(w-z) < b$ and $A' \cdot \frac{1}{t}(w-z) \leq b'$. Hence, there exists a vector $x \in \mathbb{R}^n$ such that $Ax < b$ and $A'x \leq b'$ (namely, $x = \frac{1}{t}(w-z)$). In other words, Assertion M1 holds. Hence, at least one of Assertions M1 and M2 holds.

We have thus proven that at least one of Assertions M1 and M2 holds in Case 1.

Next, let us consider Case 2. In this case, Assertion $M_1 2$ holds. In other words, there exists a vector $y \in \mathbb{R}^{m+m'}$ such that $y^T \begin{pmatrix} I_m & -b \\ 0_{m',m} & -b' \end{pmatrix} \geq 0$,

$y^T \begin{pmatrix} 0_{m,m'} & A & -A \\ I_{m'} & A' & -A' \end{pmatrix} \geq 0$ and $y^T \begin{pmatrix} I_m & -b \\ 0_{m',m} & -b' \end{pmatrix} \neq 0$. Denote this vector y by

η. Then, $\eta \in \mathbb{R}^{m+m'}$ satisfies $\eta^T \begin{pmatrix} I_m & -b \\ 0_{m',m} & -b' \end{pmatrix} \geq 0$, $\eta^T \begin{pmatrix} 0_{m,m'} & A & -A \\ I_{m'} & A' & -A' \end{pmatrix} \geq$

0 and $\eta^T \begin{pmatrix} I_m & -b \\ 0_{m',m} & -b' \end{pmatrix} \neq 0$.

Let us write the vector η in the form $\begin{pmatrix} v \\ w \end{pmatrix}$, where $v \in \mathbb{R}^m$ and $w \in \mathbb{R}^{m'}$.

Then, $\eta^T = \begin{pmatrix} v \\ w \end{pmatrix}^T = \begin{pmatrix} v^T & w^T \end{pmatrix}$. Hence,

$\eta^T \begin{pmatrix} I_m & -b \\ 0_{m',m} & -b' \end{pmatrix}$

$= \begin{pmatrix} v^T & w^T \end{pmatrix} \begin{pmatrix} I_m & -b \\ 0_{m',m} & -b' \end{pmatrix} = \begin{pmatrix} v^T I_m + w^T 0_{m',m} & v^T (-b) + w^T (-b') \end{pmatrix}$

(by the multiplication rule for block matrices)

$= \begin{pmatrix} v^T & -(v^T b + w^T b') \end{pmatrix}$

$\left(\text{since } \underbrace{v^T I_m}_{=v^T} + \underbrace{w^T 0_{m',m}}_{=0} = v^T \text{ and } v^T (-b) + w^T (-b') = -(v^T b + w^T b') \right)$.

Thus, $\begin{pmatrix} v^T & -(v^T b + w^T b') \end{pmatrix} = \eta^T \begin{pmatrix} I_m & -b \\ 0_{m',m} & -b' \end{pmatrix} \geq 0$. Therefore, $v^T \geq 0$

and $-(v^T b + w^T b') \geq 0$ (by an analogue of Lemma 2.5h **(b)** for block matrices

of the form $\begin{pmatrix} x & y \end{pmatrix}$ instead of $\begin{pmatrix} x \\ y \end{pmatrix}$). From $-(v^T b + w^T b') \geq 0$, we obtain

$v^T b + w^T b' \leq 0$. From $v^T \geq 0$, we obtain $v \geq 0$ (since a column vector is nonnegative if and only if its transpose is nonnegative).

Since $\eta^T = \begin{pmatrix} v^T & w^T \end{pmatrix}$, we have

$$\eta^T \begin{pmatrix} 0_{m,m'} & A & -A \\ I_{m'} & A' & -A' \end{pmatrix}$$

$$= \begin{pmatrix} v^T & w^T \end{pmatrix} \begin{pmatrix} 0_{m,m'} & A & -A \\ I_{m'} & A' & -A' \end{pmatrix}$$

$$= \begin{pmatrix} v^T 0_{m,m'} + w^T I_{m'} & v^T A + w^T A' & v^T(-A) + w^T(-A') \end{pmatrix}$$

(by the multiplication rule for block matrices)

$$= \begin{pmatrix} w^T & v^T A + w^T A' & -\left(v^T A + w^T A'\right) \end{pmatrix}$$

$$\left(\text{since } \underbrace{v^T 0_{m,m'}}_{=0} + \underbrace{w^T I_{m'}}_{=w^T} = w^T \text{ and } v^T(-A) + w^T(-A') = -\left(v^T A + w^T A'\right) \right).$$

Thus, $\begin{pmatrix} w^T & v^T A + w^T A' & -\left(v^T A + w^T A'\right) \end{pmatrix} = \eta^T \begin{pmatrix} 0_{m,m'} & A & -A \\ I_{m'} & A' & -A' \end{pmatrix} \geq 0.$
Hence, we have $w^T \geq 0$, $v^T A + w^T A' \geq 0$ and $-\left(v^T A + w^T A'\right) \geq 0$ (by an analogue of Lemma 2.5h **(b)** for block matrices of the form $\begin{pmatrix} x & y & z \end{pmatrix}$ instead of $\begin{pmatrix} x \\ y \end{pmatrix}$). From $-\left(v^T A + w^T A'\right) \geq 0$, we obtain $v^T A + w^T A' \leq 0$. Combining this with $v^T A + w^T A' \geq 0$, we obtain $v^T A + w^T A' = 0$. From $w^T \geq 0$, we obtain $w \geq 0$ (since a column vector is nonnegative if and only if its transpose is nonnegative).

Altogether, we now know that $v \geq 0$, $w \geq 0$, $v^T A + w^T A' = 0$, $v^T b + w^T b' \leq 0$ and $\begin{pmatrix} v^T & -\left(v^T b + w^T b'\right) \end{pmatrix} = \eta^T \begin{pmatrix} I_m & -b \\ 0_{m',m} & -b' \end{pmatrix} \neq 0.$

Since $\begin{pmatrix} v^T & -\left(v^T b + w^T b'\right) \end{pmatrix} \neq 0$, at least one of the relations $v^T \neq 0$ and $-\left(v^T b + w^T b'\right) \neq 0$ must hold. Thus, we must be in one of the following two subcases:

Subcase 2.1: We have $v^T \neq 0$.
Subcase 2.2: We have $-\left(v^T b + w^T b'\right) \neq 0$.

Let us consider Subcase 2.1 first. In this subcase, $v^T \neq 0$. Thus, $v \neq 0$. Combined with $v^T b + w^T b' \leq 0$, this yields $\left(v \neq 0 \text{ and } v^T b + w^T b' \leq 0\right)$, so that

$$\left(\left(v^T b + w^T b' < 0\right) \text{ or } \left(v \neq 0 \text{ and } v^T b + w^T b' \leq 0\right) \right).$$

Thus, we know that $v \geq 0$, $w \geq 0$, $v^T A + w^T A' = 0$ and

$$\left(\left(v^T b + w^T b' < 0\right) \text{ or } \left(v \neq 0 \text{ and } v^T b + w^T b' \leq 0\right) \right).$$

Hence, there exist two vectors $y \in \mathbb{R}^m$ and $y' \in \mathbb{R}^{m'}$ such that $y \geq 0$, $y' \geq 0$, $y^T A + y'^T A' = 0$ and

$$\left(\left(y^T b + y'^T b' < 0\right) \text{ or } \left(y \neq 0 \text{ and } y^T b + y'^T b' \leq 0\right) \right)$$

(namely, $y = v$ and $y' = w$). In other words, Assertion M2 holds. Thus, Assertion M2 holds in Subcase 2.1.

Let us now consider Subcase 2.2. In this subcase, $-\left(v^T b + w^T b'\right) \neq 0$, so that $v^T b + w^T b' \neq 0$. Combined with $v^T b + w^T b' \leq 0$, this yields $v^T b + w^T b' < 0$. Hence,

$$\left(\left(v^T b + w^T b' < 0\right) \text{ or } \left(v \neq 0 \text{ and } v^T b + w^T b' \leq 0\right)\right).$$

Thus, we know that $v \geq 0$, $w \geq 0$, $v^T A + w^T A' = 0$ and

$$\left(\left(v^T b + w^T b' < 0\right) \text{ or } \left(v \neq 0 \text{ and } v^T b + w^T b' \leq 0\right)\right).$$

Hence, there exist two vectors $y \in \mathbb{R}^m$ and $y' \in \mathbb{R}^{m'}$ such that $y \geq 0$, $y' \geq 0$, $y^T A + y'^T A' = 0$ and

$$\left(\left(y^T b + y'^T b' < 0\right) \text{ or } \left(y \neq 0 \text{ and } y^T b + y'^T b' \leq 0\right)\right)$$

(namely, $y = v$ and $y' = w$). In other words, Assertion M2 holds. Thus, Assertion M2 holds in Subcase 2.2.

We have thus proven that Assertion M2 holds in each of Subcases 2.1 and 2.2. Thus, in Case 2, Assertion M2 always holds (because Subcases 2.1 and 2.2 cover all of Case 2). Hence, in Case 2, at least one of Assertions M1 and M2 holds.

Hence, in each of the Cases 1 and 2, at least one of Assertions M1 and M2 holds. Since these Cases 1 and 2 cover all possibilities, this yields that, in every situation, at least one of Assertions M1 and M2 holds. Since we know that the Assertions M1 and M2 cannot hold at the same time, this yields that exactly one of Assertions M1 and M2 holds. This proves Theorem 2.5k. $\qquad\square$

Let us mention two direct corollaries of Theorem 2.5k:

Corollary 2.5n. Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$. Let A be an $m \times n$-matrix. Let $b \in \mathbb{R}^m$. Then, exactly one of the following two assertions holds:

Assertion N1: There exists a vector $x \in \mathbb{R}^n$ such that $Ax < b$.

Assertion N2: There exists a nonzero vector $y \in \mathbb{R}^m$ such that $y \geq 0$, $y^T A = 0$ and $y^T b \leq 0$.

Corollary 2.5n is classically equivalent to Exercise 2.18 in [Schrij17], but constructively stronger.

Here is the second corollary:

Corollary 2.5o. Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$. Let A be an $m \times n$-matrix. Let $b \in \mathbb{R}^m$. Then, exactly one of the following two assertions holds:

Assertion O1: There exists a vector $x \in \mathbb{R}^n$ such that $Ax \leq b$.

Assertion O2: There exists a vector $y \in \mathbb{R}^m$ such that $y \geq 0$, $y^T A = 0$ and $y^T b < 0$.

Corollary 2.5o is classically equivalent to [Schrij17, Corollary 2.5a], but constructively stronger.

Proof of Corollary 2.5n. The Assertions N1 and N2 cannot hold at the same time[29]. We will now show that at least one of these assertions holds.

Let $0_{0,n}$ denote the zero $0 \times n$-matrix, and let 0_0 denote the zero vector in $\mathbb{R}^0$. (Both $0_{0,n}$ and 0_0 are matrices with no entries at all.) Applying Theorem 2.5k to $m' = 0$, $A' = 0_{0,n}$ and $b' = 0_0$, we conclude that exactly one of the following two assertions holds:

Assertion $N_1 1$: There exists a vector $x \in \mathbb{R}^n$ such that $Ax < b$ and $0_{0,n} x \leq 0_0$.

Assertion $N_1 2$: There exist two vectors $y \in \mathbb{R}^m$ and $y' \in \mathbb{R}^0$ such that $y \geq 0$, $y' \geq 0$, $y^T A + y'^T 0_{0,n} = 0$ and

$$\left(\left(y^T b + y'^T 0_0 < 0 \right) \text{ or } \left(y \neq 0 \text{ and } y^T b + y'^T 0_0 \leq 0 \right) \right).$$

Hence, we must be in one of the following two cases:
Case 1: Assertion $N_1 1$ holds.
Case 2: Assertion $N_1 2$ holds.

Let us first consider Case 1. In this case, Assertion $N_1 1$ holds. In other words, there exists a vector $x \in \mathbb{R}^n$ such that $Ax < b$ and $0_{0,n} x \leq 0_0$. Thus, Assertion N1 holds. Hence, at least one of Assertions N1 and N2 holds.

We have thus proven that at least one of Assertions N1 and N2 holds in Case 1.

Next, let us consider Case 2. In this case, Assertion $N_1 2$ holds. In other words, there exist two vectors $y \in \mathbb{R}^m$ and $y' \in \mathbb{R}^0$ such that $y \geq 0$, $y' \geq 0$, $y^T A + y'^T 0_{0,n} = 0$ and

$$\left(\left(y^T b + y'^T 0_0 < 0 \right) \text{ or } \left(y \neq 0 \text{ and } y^T b + y'^T 0_0 \leq 0 \right) \right).$$

Denote these vectors y and y' by η and η'. Then, $\eta \in \mathbb{R}^m$ and $\eta' \in \mathbb{R}^0$ satisfy $\eta \geq 0$, $\eta' \geq 0$, $\eta^T A + \eta'^T 0_{0,n} = 0$ and

$$\left(\left(\eta^T b + \eta'^T 0_0 < 0 \right) \text{ or } \left(\eta \neq 0 \text{ and } \eta^T b + \eta'^T 0_0 \leq 0 \right) \right). \tag{15}$$

---

[29]*Proof.* Assume the opposite. Then, the Assertions N1 and N2 hold at the same time. Since Assertion N2 holds, there exists a nonzero vector $y \in \mathbb{R}^m$ such that $y \geq 0$, $y^T A = 0$ and $y^T b \leq 0$. Consider this y.

Since Assertion N1 holds, there exists a vector $x \in \mathbb{R}^n$ such that $Ax < b$. Consider this x.

From $Ax < b$, we obtain $b > Ax$, hence $b - Ax > 0$. Thus, $b - Ax \geq 0$. Also, $y \geq 0$, so that $y^T \geq 0$ (since the transpose of any nonnegative vector is nonnegative). Now, applying Lemma 2.0v to m, y^T and $b - Ax$ instead of n, x and y, we obtain $y^T (b - Ax) \geq 0$. Combined with $y^T (b - Ax) = y^T b - \underbrace{y^T A}_{=0} x = y^T b \leq 0$, this yields $y^T (b - Ax) = 0$. Hence, applying Lemma 2.2d to m, y^T and $b - Ax$ instead of n, x and y, we obtain $y^T = 0$, thus $y = 0$, contradicting the fact that y be nonzero. Thus, our assumption was wrong, qed.

Since $\eta'^T 0_0 = 0$, the relation (15) simplifies to $\left(\left(\eta^T b < 0 \right) \ \text{or} \ \left(\eta \neq 0 \ \text{and} \ \eta^T b \leq 0 \right) \right)$. Thus, $\eta^T b \leq 0$ (because otherwise, neither $\eta^T b < 0$ nor $\left(\eta \neq 0 \ \text{and} \ \eta^T b \leq 0 \right)$ would be possible) and $\eta \neq 0$ (for the same reason). Also, $\eta^T A + \eta'^T 0_{0,n} = 0$ simplifies to $\eta^T A = 0$.

Altogether, we know that η is nonzero and satisfies $\eta \geq 0$, $\eta^T A = 0$ and $\eta^T b \leq 0$. Thus, there exists a nonzero vector $y \in \mathbb{R}^m$ such that $y \geq 0$, $y^T A = 0$ and $y^T b \leq 0$ (namely, $y = \eta$). In other words, Assertion N2 holds. Hence, at least one of Assertions N1 and N2 holds.

We have thus proven that at least one of Assertions N1 and N2 holds in Case 2.

Hence, in each of the Cases 1 and 2, at least one of Assertions N1 and N2 holds. Since these Cases 1 and 2 cover all possibilities, this yields that, in every situation, at least one of Assertions N1 and N2 holds. Since we know that the Assertions N1 and N2 cannot hold at the same time, this yields that exactly one of Assertions N1 and N2 holds. This proves Corollary 2.5n. $\qquad\square$

Proof of Corollary 2.5o. The Assertions O1 and O2 cannot hold at the same time[30]. We will now show that at least one of these assertions holds.

Let $0_{0,n}$ denote the zero $0 \times n$-matrix, and let 0_0 denote the zero vector in $\mathbb{R}^0$. (Both $0_{0,n}$ and 0_0 are matrices with no entries at all.) Applying Theorem 2.5k to 0, m, $0_{0,n}$, 0_0, A and b instead of m, m', A, b, A' and b', we conclude that exactly one of the following two assertions holds:

> *Assertion $O_1 1$:* There exists a vector $x \in \mathbb{R}^n$ such that $0_{0,n} x < 0_0$ and $Ax \leq b$.
>
> *Assertion $O_1 2$:* There exist two vectors $y \in \mathbb{R}^0$ and $y' \in \mathbb{R}^m$ such that $y \geq 0$, $y' \geq 0$, $y^T 0_{0,n} + y'^T A = 0$ and
> $$\left(\left(y^T 0_0 + y'^T b < 0 \right) \ \text{or} \ \left(y \neq 0 \ \text{and} \ y^T 0_0 + y'^T b \leq 0 \right) \right).$$

Hence, we must be in one of the following two cases:
Case 1: Assertion $O_1 1$ holds.
Case 2: Assertion $O_1 2$ holds.

---

[30]*Proof.* Assume the opposite. Then, the Assertions O1 and O2 hold at the same time. Since Assertion O2 holds, there exists a vector $y \in \mathbb{R}^m$ such that $y \geq 0$, $y^T A = 0$ and $y^T b < 0$. Consider this y.

Since Assertion O1 holds, there exists a vector $x \in \mathbb{R}^n$ such that $Ax \leq b$. Consider this x.

From $Ax \leq b$, we obtain $b \geq Ax$, thus $b - Ax \geq 0$. Also, $y \geq 0$, so that $y^T \geq 0$ (since the transpose of any nonnegative vector is nonnegative). Now, applying Lemma 2.0v to m, y^T and $b - Ax$ instead of n, x and y, we obtain $y^T (b - Ax) \geq 0$. This contradicts $y^T (b - Ax) = y^T b - \underbrace{y^T A}_{=0} x = y^T b < 0$. This contradiction shows that our assumption was wrong, qed.

Let us first consider Case 1. In this case, Assertion $O_1 1$ holds. In other words, there exists a vector $x \in \mathbb{R}^n$ such that $0_{0,n} x < 0_0$ and $Ax \leq b$. Hence, Assertion O1 holds. Hence, at least one of Assertions O1 and O2 holds.

We have thus proven that at least one of Assertions O1 and O2 holds in Case 1.

Next, let us consider Case 2. In this case, Assertion $O_1 2$ holds. In other words, there exist two vectors $y \in \mathbb{R}^0$ and $y' \in \mathbb{R}^m$ such that $y \geq 0$, $y' \geq 0$, $y^T 0_{0,n} + y'^T A = 0$ and

$$\left(\left(y^T 0_0 + y'^T b < 0 \right) \text{ or } \left(y \neq 0 \text{ and } y^T 0_0 + y'^T b \leq 0 \right) \right).$$

Denote these vectors y and y' by η and η'. Then, $\eta \in \mathbb{R}^0$ and $\eta' \in \mathbb{R}^m$ satisfy $\eta \geq 0$, $\eta' \geq 0$, $\eta^T 0_{0,n} + \eta'^T A = 0$ and

$$\left(\left(\eta^T 0_0 + \eta'^T b < 0 \right) \text{ or } \left(\eta \neq 0 \text{ and } \eta^T 0_0 + \eta'^T b \leq 0 \right) \right). \tag{16}$$

Since $\eta \in \mathbb{R}^0$, we have $\eta = 0$ (since the only element of $\mathbb{R}^0$ is 0). Hence, the assertion $\eta \neq 0$ is false. Thus, the assertion $\left(\eta \neq 0 \text{ and } \eta^T 0_0 + \eta'^T b \leq 0 \right)$ is also false. Hence, from (16), we conclude that we must have $\eta^T 0_0 + \eta'^T b < 0$. Since $\underbrace{\eta^T 0_0}_{=0} + \eta'^T b = \eta'^T b$, this simplifies to $\eta'^T b < 0$.

Furthermore, comparing $\underbrace{\eta^T 0_{0,n}}_{=0} + \eta'^T A = \eta'^T A$ with $\eta^T 0_{0,n} + \eta'^T A = 0$, we obtain $\eta'^T A = 0$.

Altogether, we now know that $\eta' \geq 0$, $\eta'^T A = 0$ and $\eta'^T b < 0$. Hence, there exists a vector $y \in \mathbb{R}^m$ such that $y \geq 0$, $y^T A = 0$ and $y^T b < 0$ (namely, $y = \eta'$). In other words, Assertion O2 holds. Hence, at least one of Assertions O1 and O2 holds.

We have thus proven that at least one of Assertions O1 and O2 holds in Case 2.

Hence, in each of the Cases 1 and 2, at least one of Assertions O1 and O2 holds. Since these Cases 1 and 2 cover all possibilities, this yields that, in every situation, at least one of Assertions O1 and O2 holds. Since we know that the Assertions O1 and O2 cannot hold at the same time, this yields that exactly one of Assertions O1 and O2 holds. This proves Corollary 2.5o. $\square$

13. The weak Duality theorem

The following theorem is classically equivalent to, but constructively stronger than [Schrij17, Corollary 2.5b]:

Theorem 2.5p. Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$. Let A be an $m \times n$-matrix. Let $b \in \mathbb{R}^m$. Let $c \in \mathbb{R}^n$ and $\delta \in \mathbb{R}$. Assume that the set

$\{x \in \mathbb{R}^n \mid Ax \le b\}$ has at least one element. Then, exactly one of the following two assertions holds:

Assertion P1: There exists a vector $x \in \mathbb{R}^n$ such that $Ax \le b$ and $c^T x > \delta$.

Assertion P2: There exists a vector $y \in \mathbb{R}^m$ such that $y \ge 0$, $y^T A = c^T$ and $y^T b \le \delta$.

The following proof of Theorem 2.5p is being given only for the sake of completeness. It is more or less a copy of that of [Schrij17, Corollary 2.5b], with the only difference that the reductio to absurdum part has been replaced by a constructive case distinction:

Proof of Theorem 2.5p. The Assertions P1 and P2 cannot hold at the same time[31]. We will now show that at least one of these assertions holds.

Let 0_n denote the zero vector in $\mathbb{R}^n$.

Consider the block matrix $\begin{pmatrix} A^T & 0_n \\ b^T & 1 \end{pmatrix}$ (where the scalar 1 is considered as a vector in $\mathbb{R}^1$, thus as an 1×1-matrix); this is an $(n+1) \times (m+1)$-matrix.

Also, consider the block matrix $\begin{pmatrix} c \\ \delta \end{pmatrix}$. This block matrix is an $(n+1) \times 1$-matrix, i. e., a vector in $\mathbb{R}^{n+1}$.

Applying Theorem 2.5d to $n+1$, $m+1$, $\begin{pmatrix} A^T & 0_n \\ b^T & 1 \end{pmatrix}$ and $\begin{pmatrix} c \\ \delta \end{pmatrix}$ instead of m, n, A and b, we conclude that exactly one of the following two assertions holds:

Assertion $P_1$1: The system $\begin{pmatrix} A^T & 0_n \\ b^T & 1 \end{pmatrix} x = \begin{pmatrix} c \\ \delta \end{pmatrix}$ has a nonnegative solution $x \in \mathbb{R}^{m+1}$.

Assertion $P_1$2: There exists a vector $y \in \mathbb{R}^{n+1}$ such that $y^T \begin{pmatrix} A^T & 0_n \\ b^T & 1 \end{pmatrix} \ge 0$ and $y^T \begin{pmatrix} c \\ \delta \end{pmatrix} < 0$.

---

[31]*Proof.* Assume the opposite. Then, the Assertions P1 and P2 hold at the same time. Since Assertion P2 holds, there exists a vector $y \in \mathbb{R}^m$ such that $y \ge 0$, $y^T A = c^T$ and $y^T b \le \delta$. Consider this y.

Since Assertion P1 holds, there exists a vector $x \in \mathbb{R}^n$ such that $Ax \le b$ and $c^T x > \delta$. Consider this x.

From $Ax \le b$, we obtain $b \ge Ax$, so that $b - Ax \ge 0$. Also, $y \ge 0$, so that $y^T \ge 0$ (since the transpose of any nonnegative vector is nonnegative). Now, applying Lemma 2.0v to m, y^T and $b - Ax$ instead of n, x and y, we obtain $y^T (b - Ax) \ge 0$. This contradicts $y^T (b - Ax) = \underbrace{y^T b}_{\le \delta} - \underbrace{y^T A}_{=c^T} x \le \delta - \underbrace{c^T x}_{>\delta} < \delta - \delta = 0$. This contradiction shows that our assumption was wrong, qed.

Hence, we must be in one of the following two cases:

Case 1: Assertion P$_1$1 holds.

Case 2: Assertion P$_1$2 holds.

Let us first consider Case 1. In this case, Assertion P$_1$1 holds. In other words, the system $\begin{pmatrix} A^T & 0_n \\ b^T & 1 \end{pmatrix} x = \begin{pmatrix} c \\ \delta \end{pmatrix}$ has a nonnegative solution $x \in \mathbb{R}^{m+1}$. Let ξ be this solution. Then, $\xi \in \mathbb{R}^{m+1}$ is a nonnegative vector such that $\begin{pmatrix} A^T & 0_n \\ b^T & 1 \end{pmatrix} \xi = \begin{pmatrix} c \\ \delta \end{pmatrix}$.

Let us write the vector ξ in the form $\begin{pmatrix} v \\ w \end{pmatrix}$, where $v \in \mathbb{R}^m$ and $w \in \mathbb{R}^1$. Consider $w \in \mathbb{R}^1$ as a scalar (because elements of $\mathbb{R}^1$ can be identified with scalars). Then,

$$\begin{pmatrix} c \\ \delta \end{pmatrix} = \begin{pmatrix} A^T & 0_n \\ b^T & 1 \end{pmatrix} \underbrace{\xi}_{= \begin{pmatrix} v \\ w \end{pmatrix}} = \begin{pmatrix} A^T & 0_n \\ b^T & 1 \end{pmatrix} \begin{pmatrix} v \\ w \end{pmatrix} = \begin{pmatrix} A^T v + 0_n w \\ b^T v + 1w \end{pmatrix}$$

(by the rules for multiplying block matrices). Thus, $c = A^T v + 0_n w$ and $\delta = b^T v + 1w$. Hence, $c = A^T v + \underbrace{0_n w}_{=0} = A^T v$ and $\delta = b^T v + \underbrace{1w}_{=w} = b^T v + w$.

Now, $\begin{pmatrix} v \\ w \end{pmatrix} = \xi \geq 0$ (since ξ is nonnegative). Thus, $v \geq 0$ and $w \geq 0$ (by Lemma 2.5h **(b)**).

Thus,

$$\delta = b^T v + \underbrace{w}_{\geq 0} \geq b^T v = \left(b^T v\right)^T \qquad \text{(since } b^T v \text{ is a scalar, i. e., a } 1 \times 1\text{-matrix)}$$
$$= v^T \underbrace{\left(b^T\right)^T}_{=b} = v^T b.$$

In other words, $v^T b \leq \delta$.

Also, from $c = A^T v$, we obtain $c^T = \left(A^T v\right)^T = v^T \underbrace{\left(A^T\right)^T}_{=A} = v^T A$.

We thus know that $v \geq 0$, $v^T A = c^T$ and $v^T b \leq \delta$. Hence, there exists a vector $y \in \mathbb{R}^m$ such that $y \geq 0$, $y^T A = c^T$ and $y^T b \leq \delta$ (namely, $y = v$). In other words, Assertion P2 holds. Hence, at least one of Assertions P1 and P2 holds.

We have thus proven that at least one of Assertions P1 and P2 holds in Case 1.

Let us now consider Case 2. In this case, Assertion P$_1$2 holds. In other words, there exists a vector $y \in \mathbb{R}^{n+1}$ such that $y^T \begin{pmatrix} A^T & 0_n \\ b^T & 1 \end{pmatrix} \geq 0$ and $y^T \begin{pmatrix} c \\ \delta \end{pmatrix} <$

0. Denote this y by w. Then, $w \in \mathbb{R}^{n+1}$ satisfies $w^T \begin{pmatrix} A^T & 0_n \\ b^T & 1 \end{pmatrix} \geq 0$ and $w^T \begin{pmatrix} c \\ \delta \end{pmatrix} < 0$.

Let us write the vector w in the form $\begin{pmatrix} v \\ t \end{pmatrix}$, where $v \in \mathbb{R}^n$ and $t \in \mathbb{R}^1$. Consider $t \in \mathbb{R}^1$ as a scalar (because elements of $\mathbb{R}^1$ can be identified with scalars). Then, $t^T = t$.

We have $\left(c^T v\right)^T = c^T v$ (since $c^T v$ is a scalar). Thus, $c^T v = \left(c^T v\right)^T = v^T \underbrace{\left(c^T\right)^T}_{=c} = v^T c$, therefore $v^T c = c^T v$.

From $w = \begin{pmatrix} v \\ t \end{pmatrix}$, we obtain $w^T = \begin{pmatrix} v \\ t \end{pmatrix}^T = \begin{pmatrix} v^T & t^T \end{pmatrix} = \begin{pmatrix} v^T & t \end{pmatrix}$ (since $t^T = t$). Thus,

$$w^T \begin{pmatrix} c \\ \delta \end{pmatrix} = \begin{pmatrix} v^T & t \end{pmatrix} \begin{pmatrix} c \\ \delta \end{pmatrix} = \underbrace{v^T c}_{=c^T v} + t\delta$$

$$\text{(by the multiplication rule for block matrices)}$$

$$= c^T v + t\delta.$$

Hence, the relation $w^T \begin{pmatrix} c \\ \delta \end{pmatrix} < 0$ (which we know to be true) rewrites as $c^T v + t\delta < 0$.

On the other hand, from $w^T = \begin{pmatrix} v^T & t \end{pmatrix}$, we deduce that

$$w^T \begin{pmatrix} A^T & 0_n \\ b^T & 1 \end{pmatrix} = \begin{pmatrix} v^T & t \end{pmatrix} \begin{pmatrix} A^T & 0_n \\ b^T & 1 \end{pmatrix} = \begin{pmatrix} v^T A^T + t b^T & v^T 0_n + t \cdot 1 \end{pmatrix}$$

(by the multiplication rule for block matrices). Thus, the relation $w^T \begin{pmatrix} A^T & 0_n \\ b^T & 1 \end{pmatrix} \geq 0$ (which we know to be true) rewrites as $\begin{pmatrix} v^T A^T + t b^T & v^T 0_n + t \cdot 1 \end{pmatrix} \geq 0$. Thus, (using the analogue of Lemma 2.5h **(b)** for block matrices of the form $\begin{pmatrix} x & y \end{pmatrix}$ instead of $\begin{pmatrix} x \\ y \end{pmatrix}$) we conclude that $v^T A^T + t b^T \geq 0$ and $v^T 0_n + t \cdot 1 \geq 0$. From $v^T A^T + t b^T \geq 0$, we obtain $\left(v^T A^T + t b^T\right)^T \geq 0$ (since the transpose of a nonnegative vector is always nonnegative). In light of $\left(v^T A^T + t b^T\right)^T = \underbrace{\left(A^T\right)^T}_{=A} \underbrace{\left(v^T\right)^T}_{=v} + \underbrace{\left(b^T\right)^T}_{=b} \underbrace{t^T}_{=t} = Av + bt$, this rewrites as $Av + bt \geq 0$. Since $bt = tb$ (as we can regard t as a scalar), this rewrites as $Av + tb \geq 0$.

Of course, $v^T 0_n + t \cdot 1 \geq 0$ simplifies to $t \geq 0$. Hence, we must be in one of the following two subcases:

Subcase 2.1: We have $t = 0$.

Subcase 2.2: We have $t > 0$.

Let us first consider Subcase 2.1. In this subcase, we have $t = 0$. Thus, $c^T v + \underbrace{t}_{=0} \delta = c^T v + 0\delta = c^T v$. Hence, the relation $c^T v + t\delta < 0$ (which we know to be true) rewrites as $c^T v < 0$. Thus, $c^T v \neq 0$, so that $c^T v$ is invertible.

We know that the set $\{x \in \mathbb{R}^n \mid Ax \leq b\}$ has at least one element. Let x_0 be such an element. Then, $x_0 \in \{x \in \mathbb{R}^n \mid Ax \leq b\}$, so that $x_0 \in \mathbb{R}^n$ and $Ax_0 \leq b$.

Let $\tau = \max\left\{0, \dfrac{c^T x_0 - \delta}{c^T v} + 1\right\}$ [32]. Then, $\tau \geq 0$ and $\tau \geq \dfrac{c^T x_0 - \delta}{c^T v} + 1$.

We have $\tau \geq \dfrac{c^T x_0 - \delta}{c^T v} + 1 > \dfrac{c^T x_0 - \delta}{c^T v}$. Multiplying this inequality with $c^T v$, we obtain

$$\tau c^T v < c^T x_0 - \delta \tag{17}$$

(since $c^T v < 0$).

On the other hand, $Av + \underbrace{t}_{=0} b = Av + 0b = Av$, so that $Av = Av + tb \geq 0$.

Now, using $\tau \geq 0$, we obtain $\tau \underbrace{Av}_{\geq 0} \geq \tau 0 = 0$. Hence, $A(x_0 - \tau v) = Ax_0 - \underbrace{\tau Av}_{\geq 0} \leq Ax_0 \leq b$.

On the other hand,

$$c^T (x_0 - \tau v) = c^T x_0 - \underbrace{\tau c^T v}_{\substack{< c^T x_0 - \delta \\ \text{(by (17))}}} > c^T x_0 - \left(c^T x_0 - \delta\right) = \delta.$$

Altogether, we thus know that $A(x_0 - \tau v) \leq b$ and $c^T(x_0 - \tau v) > \delta$. Hence, there exists a vector $x \in \mathbb{R}^n$ such that $Ax \leq b$ and $c^T x > \delta$ (namely, $x = x_0 - \tau v$). In other words, Assertion P1 holds. We have thus proven that Assertion P1 holds in Subcase 2.1.

Now, let us consider Subcase 2.2. In this subcase, $t > 0$. Thus, $t \neq 0$, so that t is invertible, and $\dfrac{1}{t} > 0$ (since $t > 0$). Hence, from $Av + tb \geq 0$, we obtain $Av \geq -tb$. Multiplying this inequality by $\dfrac{1}{t}$ (which is allowed, since $\dfrac{1}{t} > 0$), we obtain $\dfrac{1}{t} Av \geq -b$. Hence, $A\left(-\dfrac{1}{t}v\right) = -\underbrace{\dfrac{1}{t}Av}_{\geq -b} \leq -(-b) = b$.

On the other hand,

$$\delta - \underbrace{c^T\left(-\dfrac{1}{t}v\right)}_{= -\frac{1}{t}c^T v} = \delta - \left(-\dfrac{1}{t}c^T v\right) = \delta + \dfrac{1}{t}c^T v = \dfrac{1}{t}\left(t\delta + c^T v\right) = \underbrace{\dfrac{1}{t}}_{>0}\underbrace{\left(c^T v + t\delta\right)}_{<0} < 0,$$

---

[32]This is well-defined, since the real number $c^T v$ is invertible.

so that $c^T \left(-\dfrac{1}{t} v \right) > \delta$.

Altogether, we know that $A \left(-\dfrac{1}{t} v \right) \leq b$ and $c^T \left(-\dfrac{1}{t} v \right) > \delta$. Thus, there exists a vector $x \in \mathbb{R}^n$ such that $Ax \leq b$ and $c^T x > \delta$ (namely, $x = -\dfrac{1}{t} v$). In other words, Assertion P1 holds. We have thus proven that Assertion P1 holds in Subcase 2.2.

Altogether, we know that Assertion P1 holds in each of Subcases 2.1 and 2.2. Thus, in Case 2, Assertion P1 always holds (since Subcases 2.1 and 2.2 cover all of Case 2). Thus, in Case 2, at least one of Assertions P1 and P2 holds.

Hence, in each of the Cases 1 and 2, at least one of Assertions P1 and P2 holds. Since these Cases 1 and 2 cover all possibilities, this yields that, in every situation, at least one of Assertions P1 and P2 holds. Since we know that the Assertions P1 and P2 cannot hold at the same time, this yields that exactly one of Assertions P1 and P2 holds. This proves Theorem 2.5p. $\qquad \square$

Theorem 2.5p yields a weak version of linear programming duality:

> **Corollary 2.5q.** Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$. Let A be an $m \times n$-matrix. Let $b \in \mathbb{R}^m$. Let $c \in \mathbb{R}^n$. Assume that the number $\max \left\{ c^T x \mid x \in \mathbb{R}^n; \ Ax \leq b \right\}$ exists. Then, the number $\min \left\{ y^T b \mid y \in \mathbb{R}^m; \ y \geq 0 \text{ and } y^T A = c^T \right\}$ exists and satisfies
>
> $$\max \left\{ c^T x \mid x \in \mathbb{R}^n; \ Ax \leq b \right\} = \min \left\{ y^T b \mid y \in \mathbb{R}^m; \ y \geq 0 \text{ and } y^T A = c^T \right\}.$$

I am calling this version weak because it requires the existence of $\max \left\{ c^T x \mid x \in \mathbb{R}^n; \ Ax \leq b \right\}$, while the same conclusion can be proven under the (easier to verify) assumption that the set $\left\{ c^T x \mid x \in \mathbb{R}^n; \ Ax \leq b \right\}$ is nonempty and bounded from below. (This will follow from Theorem 2.6c below.) Here is yet another weak version of linear programming duality:

> **Corollary 2.5r.** Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$. Let A be an $m \times n$-matrix. Let $b \in \mathbb{R}^m$. Let $c \in \mathbb{R}^n$. Assume that the number $\min \left\{ y^T b \mid y \in \mathbb{R}^m; \ y \geq 0 \text{ and } y^T A = c^T \right\}$ exists. Then, the number $\max \left\{ c^T x \mid x \in \mathbb{R}^n; \ Ax \leq b \right\}$ exists and satisfies
>
> $$\max \left\{ c^T x \mid x \in \mathbb{R}^n; \ Ax \leq b \right\} = \min \left\{ y^T b \mid y \in \mathbb{R}^m; \ y \geq 0 \text{ and } y^T A = c^T \right\}.$$

Both Corollary 2.5q and Corollary 2.5r are easy corollaries of Theorem 2.6c below. We will, however, give an alternative proof of Corollary 2.5q using Theorem 2.5p first. In a similar vein, Corollary 2.5r could be shown using an analogue of Theorem 2.5p.

Proof of Corollary 2.5q. The number $\max\left\{c^T x \mid x \in \mathbb{R}^n; \ Ax \leq b\right\}$ exists (by assumption). Denote this number by δ. Then, every element of $\left\{c^T x \mid x \in \mathbb{R}^n; \ Ax \leq b\right\}$ is $\leq \delta$. In other words,

$$c^T x \leq \delta \text{ for every } x \in \mathbb{R}^n \text{ satisfying } Ax \leq b. \tag{18}$$

Also, the set $\left\{c^T x \mid x \in \mathbb{R}^n; \ Ax \leq b\right\}$ has a maximum (since $\max\left\{c^T x \mid x \in \mathbb{R}^n; \ Ax \leq b\right\}$ exists) and thus has at least one element (since a set that has a maximum must always have at least one element). In other words, there exists some $x \in \mathbb{R}^n$ satisfying $Ax \leq b$. In other words, the set $\left\{x \in \mathbb{R}^n \mid Ax \leq b\right\}$ has at least one element. Hence, Theorem 2.5p yields that exactly one of the following two assertions holds:

> *Assertion $Q_1 1$:* There exists a vector $x \in \mathbb{R}^n$ such that $Ax \leq b$ and $c^T x > \delta$.

> *Assertion $Q_1 2$:* There exists a vector $y \in \mathbb{R}^m$ such that $y \geq 0$, $y^T A = c^T$ and $y^T b \leq \delta$.

Since Assertion $Q_1 1$ cannot hold[33], this yields that Assertion $Q_1 2$ must hold. In other words, there exists a vector $y \in \mathbb{R}^m$ such that $y \geq 0$, $y^T A = c^T$ and $y^T b \leq \delta$. Denote this y by w. Thus, $w \in \mathbb{R}^m$ satisfies $w \geq 0$, $w^T A = c^T$ and $w^T b \leq \delta$. Since $w \in \mathbb{R}^m$ satisfies $w \geq 0$ and $w^T A = c^T$, we have $w \in \left\{y \in \mathbb{R}^m \mid y \geq 0 \text{ and } y^T A = c^T\right\}$ and thus

$$w^T b \in \left\{y^T b \mid y \in \mathbb{R}^m; \ y \geq 0 \text{ and } y^T A = c^T\right\}. \tag{19}$$

Now, we will prove that

$$\lambda \geq \delta \text{ for every } \lambda \in \left\{y^T b \mid y \in \mathbb{R}^m; \ y \geq 0 \text{ and } y^T A = c^T\right\}. \tag{20}$$

[Proof of (20): Let $\lambda \in \left\{y^T b \mid y \in \mathbb{R}^m; \ y \geq 0 \text{ and } y^T A = c^T\right\}$. Thus, there exists a $z \in \mathbb{R}^m$ such that $z \geq 0$, $z^T A = c^T$ and $z^T b = \lambda$. Consider this z.

By the definition of δ, we have

$$\delta = \max\left\{c^T x \mid x \in \mathbb{R}^n; \ Ax \leq b\right\} \in \left\{c^T x \mid x \in \mathbb{R}^n; \ Ax \leq b\right\}.$$

Hence, there exists some $x \in \mathbb{R}^n$ such that $Ax \leq b$ and $c^T x = \delta$. Denote this x by q. Then, $q \in \mathbb{R}^n$ satisfies $Aq \leq b$ and $c^T q = \delta$. From $Aq \leq b$, we obtain $b \geq Aq$, thus $b - Aq \geq 0$. Also, $z \geq 0$, so that $z^T \geq 0$ (since the transpose of a nonnegative vector must always be nonnegative). Thus, applying Lemma 2.0v to m, z^T and $b - Aq$ instead of n, x and y, we obtain $z^T (b - Aq) \geq 0$ (since

---

[33]*Proof.* Assume the contrary. Then, Assertion $Q_1 1$ holds. In other words, there exists a vector $x \in \mathbb{R}^n$ such that $Ax \leq b$ and $c^T x > \delta$. Consider this x. Then, $Ax \leq b$, so that $c^T x \leq \delta$ (by (18)), contradicting $c^T x > \delta$. This contradiction shows that our assumption was wrong, qed.

$b - Aq \geq 0$). Since $z^T (b - Aq) = \underbrace{z^T b}_{=\lambda} - \underbrace{z^T A}_{=c^T} q = \lambda - \underbrace{c^T q}_{=\delta} = \lambda - \delta$, this rewrites as $\lambda - \delta \geq 0$. In other words, $\lambda \geq \delta$. This proves (20).]

From (19), we know that $w^T b \in \{y^T b \mid y \in \mathbb{R}^m; \ y \geq 0 \text{ and } y^T A = c^T\}$. Hence, applying (20) to $\lambda = w^T b$, we obtain $w^T b \geq \delta$. Combined with $w^T b \leq \delta$, this yields $w^T b = \delta$. Hence, (19) rewrites as

$$\delta \in \{y^T b \mid y \in \mathbb{R}^m; \ y \geq 0 \text{ and } y^T A = c^T\}.$$

Thus, δ is an element of the set $\{y^T b \mid y \in \mathbb{R}^m; \ y \geq 0 \text{ and } y^T A = c^T\}$. Combined with the fact that every element of the set $\{y^T b \mid y \in \mathbb{R}^m; \ y \geq 0 \text{ and } y^T A = c^T\}$ is $\geq \delta$ (by (20)), this yields that δ is the minimum of the set $\{y^T b \mid y \in \mathbb{R}^m; \ y \geq 0 \text{ and } y^T A = c^T\}$. In other words, the number $\min\{y^T b \mid y \in \mathbb{R}^m; \ y \geq 0 \text{ and } y^T A = c^T\}$ exists and satisfies

$$\min\{y^T b \mid y \in \mathbb{R}^m; \ y \geq 0 \text{ and } y^T A = c^T\} = \delta.$$

Since $\delta = \max\{c^T x \mid x \in \mathbb{R}^n; \ Ax \leq b\}$, this rewrites as follows: The number $\min\{y^T b \mid y \in \mathbb{R}^m; \ y \geq 0 \text{ and } y^T A = c^T\}$ exists and satisfies

$$\min\{y^T b \mid y \in \mathbb{R}^m; \ y \geq 0 \text{ and } y^T A = c^T\} = \max\{c^T x \mid x \in \mathbb{R}^n; \ Ax \leq b\}.$$

This proves Corollary 2.5q. $\square$

14. The Duality theorem

We are almost ready to formulate the linear programming duality theorem in one of its strongest forms. First, let us define two basic notions:

> **Definition 2.6b. (i)** A subset S of $\mathbb{R}$ is said to be *unbounded from above* if for every $\delta \in \mathbb{R}$, there exists a $t \in S$ such that $t \geq \delta$.
>
> **(ii)** A subset S of $\mathbb{R}$ is said to be *unbounded from below* if for every $\delta \in \mathbb{R}$, there exists a $t \in S$ such that $t \leq \delta$.

In classical logic, a subset S of $\mathbb{R}$ is unbounded from above if and only if it is not bounded from above (i. e., there does not exist any $m \in \mathbb{R}$ such that every $t \in S$ satisfies $t \leq m$). In constructive logic, this is not generally true, and the assertion that S be unbounded from above is stronger than the assertion that S not be bounded from above.

We now state a theorem that is somewhat stronger (both classically and constructively) than [Schrij17, Theorem 2.6]:

> **Theorem 2.6c.** Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$. Let A be an $m \times n$-matrix. Let $b \in \mathbb{R}^m$. Let $c \in \mathbb{R}^n$. Then, exactly one of the following four assertions holds:

Assertion I1: The sets $\{x \in \mathbb{R}^n \mid Ax \le b\}$ and $\{y \in \mathbb{R}^m \mid y \ge 0 \text{ and } y^T A = c^T\}$ are empty.

Assertion I2: The set $\{c^T x \mid x \in \mathbb{R}^n; Ax \le b\}$ is unbounded from above, and the set $\{y \in \mathbb{R}^m \mid y \ge 0 \text{ and } y^T A = c^T\}$ is empty.

Assertion I3: The set $\{y^T b \mid y \in \mathbb{R}^m; y \ge 0 \text{ and } y^T A = c^T\}$ is unbounded from below, and the set $\{x \in \mathbb{R}^n \mid Ax \le b\}$ is empty.

Assertion I4: The numbers $\max\{c^T x \mid x \in \mathbb{R}^n; Ax \le b\}$ and $\min\{y^T b \mid y \in \mathbb{R}^m; y \ge 0 \text{ and } y^T A = c^T\}$ exist and satisfy

$$\max\{c^T x \mid x \in \mathbb{R}^n; Ax \le b\} = \min\{y^T b \mid y \in \mathbb{R}^m; y \ge 0 \text{ and } y^T A = c^T\}.$$

Theorem 2.6c is often referred to as the "asymmetric version of the strong duality theorem of linear programming".

Let me reiterate that Theorem 2.6c still holds if we replace $\mathbb{R}$ by any other ordered field, such as $\mathbb{Q}$. This cannot be said of the proof of Theorem 2.6 in [Schrij17], since it uses the fact that any set of real numbers has a supremum or is unbounded from above (and this fact does not hold for $\mathbb{Q}$). So we are going to give a different proof.

First, we prove a slightly weaker result:

Lemma 2.6d. Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$. Let A be an $m \times n$-matrix. Let $b \in \mathbb{R}^m$. Let $c \in \mathbb{R}^n$. Assume that the set $\{x \in \mathbb{R}^n \mid Ax \le b\}$ has at least one element. Then, exactly one of the following two assertions holds:

Assertion J1: The set $\{c^T x \mid x \in \mathbb{R}^n; Ax \le b\}$ is unbounded from above, and the set $\{y \in \mathbb{R}^m \mid y \ge 0 \text{ and } y^T A = c^T\}$ is empty.

Assertion J2: The numbers $\max\{c^T x \mid x \in \mathbb{R}^n; Ax \le b\}$ and $\min\{y^T b \mid y \in \mathbb{R}^m; y \ge 0 \text{ and } y^T A = c^T\}$ exist and satisfy

$$\max\{c^T x \mid x \in \mathbb{R}^n; Ax \le b\} = \min\{y^T b \mid y \in \mathbb{R}^m; y \ge 0 \text{ and } y^T A = c^T\}.$$

It should be said that Lemma 2.6d is weaker than Theorem 2.6c, but still stronger than [Schrij17, Theorem 2.6].

Proof of Lemma 2.6d. The Assertions J1 and J2 cannot hold at the same time[34]. We will now show that at least one of these assertions holds.

---

[34]*Proof.* Assume the opposite. Then, the Assertions J1 and J2 hold at the same time.

The set $\{c^T x \mid x \in \mathbb{R}^n; Ax \le b\}$ is unbounded from above (by Assertion J1), and thus has no maximum. In other words, the number $\max\{c^T x \mid x \in \mathbb{R}^n; Ax \le b\}$ does not exist. But this contradicts Assertion J2. This contradiction shows that our assumption was wrong, qed.

For any nonnegative integers α and β, let $0_{\alpha,\beta}$ denote the $\alpha \times \beta$ zero matrix. For any nonnegative integer γ, let I_γ denote the $\gamma \times \gamma$ identity matrix.

Consider the block matrix $\begin{pmatrix} 0_{m,n} & -I_m \\ A & 0_{m,m} \\ 0_{n,n} & A^T \\ 0_{n,n} & -A^T \\ -c^T & b^T \end{pmatrix}$; this is an $(2m+2n+1) \times (n+m)$-matrix.

Consider the block matrix $\begin{pmatrix} 0_{m,1} \\ b \\ c \\ -c \\ 0_{1,1} \end{pmatrix}$; this is an $(2m+2n+1) \times 1$-matrix, i. e., a vector in $\mathbb{R}^{2m+2n+1}$. We identify the 1×1-matrix $0_{1,1}$ with the scalar 0; thus,

the block matrix $\begin{pmatrix} 0_{m,1} \\ b \\ c \\ -c \\ 0_{1,1} \end{pmatrix}$ rewrites as $\begin{pmatrix} 0_{m,1} \\ b \\ c \\ -c \\ 0 \end{pmatrix}$.

Applying Corollary 2.5o to $2m+2n+1$, $n+m$, $\begin{pmatrix} 0_{m,n} & -I_m \\ A & 0_{m,m} \\ 0_{n,n} & A^T \\ 0_{n,n} & -A^T \\ -c^T & b^T \end{pmatrix}$ and $\begin{pmatrix} 0_{m,1} \\ b \\ c \\ -c \\ 0_{1,1} \end{pmatrix}$

instead of m, n, A and b, we see that exactly one of the following two assertions holds:

Assertion $J_1 1$: There exists a vector $x \in \mathbb{R}^{n+m}$ such that $\begin{pmatrix} 0_{m,n} & -I_m \\ A & 0_{m,m} \\ 0_{n,n} & A^T \\ 0_{n,n} & -A^T \\ -c^T & b^T \end{pmatrix} x \leq$

$\begin{pmatrix} 0_{m,1} \\ b \\ c \\ -c \\ 0_{1,1} \end{pmatrix}$.

Assertion $J_1 2$: There exists a vector $y \in \mathbb{R}^{2m+2n+1}$ such that $y \geq 0$,

$y^T \begin{pmatrix} 0_{m,n} & -I_m \\ A & 0_{m,m} \\ 0_{n,n} & A^T \\ 0_{n,n} & -A^T \\ -c^T & b^T \end{pmatrix} = 0$ and $y^T \begin{pmatrix} 0_{m,1} \\ b \\ c \\ -c \\ 0_{1,1} \end{pmatrix} < 0$.

Hence, we must be in one of the following two cases:

Case 1: Assertion $J_1 1$ holds.

Case 2: Assertion $J_1 2$ holds.

Let us first consider Case 1. In this case, Assertion $J_1 1$ holds. In other words, there exists a vector $x \in \mathbb{R}^{n+m}$ such that $\begin{pmatrix} 0_{m,n} & -I_m \\ A & 0_{m,m} \\ 0_{n,n} & A^T \\ 0_{n,n} & -A^T \\ -c^T & b^T \end{pmatrix} x \leq \begin{pmatrix} 0_{m,1} \\ b \\ c \\ -c \\ 0_{1,1} \end{pmatrix}$. Denote this vector x by ξ. Then, $\xi \in \mathbb{R}^{n+m}$ satisfies $\begin{pmatrix} 0_{m,n} & -I_m \\ A & 0_{m,m} \\ 0_{n,n} & A^T \\ 0_{n,n} & -A^T \\ -c^T & b^T \end{pmatrix} \xi \leq \begin{pmatrix} 0_{m,1} \\ b \\ c \\ -c \\ 0_{1,1} \end{pmatrix}$.

Let us write the vector ξ in the form $\begin{pmatrix} u \\ v \end{pmatrix}$, where $u \in \mathbb{R}^n$ and $v \in \mathbb{R}^m$. Since $\xi = \begin{pmatrix} u \\ v \end{pmatrix}$, we have

$$\begin{pmatrix} 0_{m,n} & -I_m \\ A & 0_{m,m} \\ 0_{n,n} & A^T \\ 0_{n,n} & -A^T \\ -c^T & b^T \end{pmatrix} \xi$$

$$= \begin{pmatrix} 0_{m,n} & -I_m \\ A & 0_{m,m} \\ 0_{n,n} & A^T \\ 0_{n,n} & -A^T \\ -c^T & b^T \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} 0_{m,n}u + (-I_m)v \\ Au + 0_{m,m}v \\ 0_{n,n}u + A^T v \\ 0_{n,n}u + (-A^T)v \\ -c^T u + b^T v \end{pmatrix}$$

(by the multiplication rule for block matrices)

$$= \begin{pmatrix} -v \\ Au \\ A^T v \\ -A^T v \\ -c^T u + b^T v \end{pmatrix} \begin{pmatrix} \text{since } \underbrace{0_{m,n}u}_{=0} + \underbrace{(-I_m)v}_{=-v} = -v, \ Au + \underbrace{0_{m,m}v}_{=0} = Au, \\ \underbrace{0_{n,n}u}_{=0} + A^T v = A^T v, \text{ and } \underbrace{0_{n,n}u}_{=0} + \underbrace{(-A^T)v}_{=-A^T v} = -A^T v \end{pmatrix}.$$

Hence,

$$\begin{pmatrix} -v \\ Au \\ A^T v \\ -A^T v \\ -c^T u + b^T v \end{pmatrix} = \begin{pmatrix} 0_{m,n} & -I_m \\ A & 0_{m,m} \\ 0_{n,n} & A^T \\ 0_{n,n} & -A^T \\ -c^T & b^T \end{pmatrix} \xi \leq \begin{pmatrix} 0_{m,1} \\ b \\ c \\ -c \\ 0_{1,1} \end{pmatrix}.$$

Hence, we have the five inequalities[35]

$$-v \leq 0_{m,1}, \qquad Au \leq b, \qquad A^T v \leq c,$$
$$-A^T v \leq -c, \qquad \text{and } -c^T u + b^T v \leq 0_{1,1}.$$

The inequality $-v \leq 0_{m,1}$ leads to $v \geq -0_{m,1} = 0_{m,1} = 0$. Thus, $v^T \geq 0$ (since the transpose of any nonnegative vector is nonnegative). The inequality $-A^T v \leq -c$ rewrites as $A^T v \geq c$; combining this with $A^T v \leq c$, we obtain $A^T v = c$. The inequality $-c^T u + b^T v \leq 0_{1,1}$ becomes $-c^T u + b^T v \leq 0_{1,1} = 0$, so that $b^T v \leq c^T u$.

Since $b^T v$ is a scalar, we have $\left(b^T v\right)^T = b^T v$. Thus, $b^T v = \left(b^T v\right)^T = v^T \underbrace{\left(b^T\right)^T}_{=b} = v^T b$. Hence, $v^T b = b^T v \leq c^T u$.

But $u \in \mathbb{R}^n$ satisfies $Au \leq b$. Thus, $u \in \{x \in \mathbb{R}^n \mid Ax \leq b\}$ and therefore $c^T u \in \{c^T x \mid x \in \mathbb{R}^n; Ax \leq b\}$.

On the other hand,

$$\text{every } w \in \{c^T x \mid x \in \mathbb{R}^n; Ax \leq b\} \text{ satisfies } w \leq v^T b. \tag{21}$$

[36] Applying this to $w = c^T u$, we obtain $c^T u \leq v^T b$ (since $c^T u \in \{c^T x \mid x \in \mathbb{R}^n; Ax \leq b\}$). Combined with $v^T b \leq c^T u$, this yields

$$v^T b = c^T u. \tag{22}$$

Thus, $v^T b = c^T u \in \{c^T x \mid x \in \mathbb{R}^n; Ax \leq b\}$. In other words, $v^T b$ is an element of the set $\{c^T x \mid x \in \mathbb{R}^n; Ax \leq b\}$. Combined with the fact that every element of the set $\{c^T x \mid x \in \mathbb{R}^n; Ax \leq b\}$ is $\leq v^T b$ (due to (21)), this yields that $v^T b$ is the maximum of the set $\{c^T x \mid x \in \mathbb{R}^n; Ax \leq b\}$. In other words, the number $\max \{c^T x \mid x \in \mathbb{R}^n; Ax \leq b\}$ exists and satisfies

$$\max \{c^T x \mid x \in \mathbb{R}^n; Ax \leq b\} = v^T b. \tag{23}$$

---

[35]Here, we are using the following simple fact: Let n_1, n_2, n_3, n_4, and n_5 be five nonnegative integers. Let $\alpha_1 \in \mathbb{R}^{n_1}$, $\alpha_2 \in \mathbb{R}^{n_2}$, $\alpha_3 \in \mathbb{R}^{n_3}$, $\alpha_4 \in \mathbb{R}^{n_4}$, and $\alpha_5 \in \mathbb{R}^{n_5}$ be five column vectors. Let $\beta_1 \in \mathbb{R}^{n_1}$, $\beta_2 \in \mathbb{R}^{n_2}$, $\beta_3 \in \mathbb{R}^{n_3}$, $\beta_4 \in \mathbb{R}^{n_4}$, and $\beta_5 \in \mathbb{R}^{n_5}$ be five column vectors.

Assume that $\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \\ \alpha_5 \end{pmatrix} \geq \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_5 \end{pmatrix}$. Then, we have the five inequalities

$$\alpha_1 \geq \beta_1, \qquad \alpha_2 \geq \beta_2, \qquad \alpha_3 \geq \beta_3, \qquad \alpha_4 \geq \beta_4, \qquad \text{and } \alpha_5 \geq \beta_5.$$

[36]*Proof of (21):* Let $w \in \{c^T x \mid x \in \mathbb{R}^n; Ax \leq b\}$. Then, there exists some $x \in \mathbb{R}^n$ such that $Ax \leq b$ and $w = c^T x$. Consider this x.

We have $Ax \leq b$, hence $b \geq Ax$. Thus, $b - Ax \geq 0$. Thus, Lemma 2.0v (applied to m, v^T and $b - Ax$ instead of n, x and y) yields $v^T (b - Ax) \geq 0$ (since $v^T \geq 0$). Thus, $0 \leq v^T (b - Ax) = v^T b - \underbrace{v^T A}_{=c^T} x = v^T b - \underbrace{c^T x}_{=w} = v^T b - w$, so that $w \leq v^T b$. This proves (21).

Now,

$$\text{every } w \in \left\{ y^T b \mid y \in \mathbb{R}^m; \ y \geq 0 \text{ and } y^T A = c^T \right\} \text{ satisfies } w \geq v^T b. \qquad (24)$$

[37] Also, from $v \geq 0$ and $v^T A = c^T$, we conclude that $v \in \left\{ y \in \mathbb{R}^m \mid y \geq 0 \text{ and } y^T A = c^T \right\}$. Hence, $v^T b \in \left\{ y^T b \mid y \in \mathbb{R}^m; \ y \geq 0 \text{ and } y^T A = c^T \right\}$. Combined with the fact that every element of the set $\left\{ y^T b \mid y \in \mathbb{R}^m; \ y \geq 0 \text{ and } y^T A = c^T \right\}$ is $\geq v^T b$ (by (24)), this yields that $v^T b$ is the minimum of the set $\left\{ y^T b \mid y \in \mathbb{R}^m; \ y \geq 0 \text{ and } y^T A = c^T \right\}$. In other words, the number $\min \left\{ y^T b \mid y \in \mathbb{R}^m; \ y \geq 0 \text{ and } y^T A = c^T \right\}$ exists and satisfies

$$\min \left\{ y^T b \mid y \in \mathbb{R}^m; \ y \geq 0 \text{ and } y^T A = c^T \right\} = v^T b.$$

Compared with (23), this yields

$$\max \left\{ c^T x \mid x \in \mathbb{R}^n; \ Ax \leq b \right\} = \min \left\{ y^T b \mid y \in \mathbb{R}^m; \ y \geq 0 \text{ and } y^T A = c^T \right\}.$$

Altogether, we have shown that the numbers $\max \left\{ c^T x \mid x \in \mathbb{R}^n; \ Ax \leq b \right\}$ and $\min \left\{ y^T b \mid y \in \mathbb{R}^m; \ y \geq 0 \text{ and } y^T A = c^T \right\}$ exist and satisfy

$$\max \left\{ c^T x \mid x \in \mathbb{R}^n; \ Ax \leq b \right\} = \min \left\{ y^T b \mid y \in \mathbb{R}^m; \ y \geq 0 \text{ and } y^T A = c^T \right\}.$$

In other words, Assertion J2 holds. Thus, at least one of Assertions J1 and J2 holds.

We have thus proven that in Case 1, at least one of Assertions J1 and J2 holds.

Now, let us consider Case 2. In this case, Assertion $J_1 2$ holds. In other words, there exists a vector $y \in \mathbb{R}^{2m+2n+1}$ such that $y \geq 0$, $y^T \begin{pmatrix} 0_{m,n} & -I_m \\ A & 0_{m,m} \\ 0_{n,n} & A^T \\ 0_{n,n} & -A^T \\ -c^T & b^T \end{pmatrix} = 0$ and

---

[37] *Proof of (24):* Let $w \in \left\{ y^T b \mid y \in \mathbb{R}^m; \ y \geq 0 \text{ and } y^T A = c^T \right\}$. Then, there exists an $y \in \mathbb{R}^m$ such that $y \geq 0$, $y^T A = c^T$ and $w = y^T b$. Consider this y. Since $y \geq 0$, we have $y^T \geq 0$ (since the transpose of any nonnegative vector is nonnegative).

We have $Au \leq b$. Thus, $b \geq Au$, so that $b - Au \geq 0$. Thus, Lemma 2.0v (applied to m, y^T and $b - Au$ instead of n, x and y) yields $y^T (b - Au) \geq 0$ (since $y^T \geq 0$). Thus, $0 \leq y^T (b - Au) = y^T b - \underbrace{y^T A}_{=c^T} u = \underbrace{y^T b}_{=w} - \underbrace{c^T u}_{\substack{=v^T b \\ (\text{by } (22))}} = w - v^T b$, so that $w \geq v^T b$. This proves

(24).

$$y^T \begin{pmatrix} 0_{m,1} \\ b \\ c \\ -c \\ 0_{1,1} \end{pmatrix} < 0. \text{ Denote this vector } y \text{ by } \eta. \text{ Then, } \eta \in \mathbb{R}^{2m+2n+1} \text{ satisfies } \eta \geq 0,$$

$$\eta^T \begin{pmatrix} 0_{m,n} & -I_m \\ A & 0_{m,m} \\ 0_{n,n} & A^T \\ 0_{n,n} & -A^T \\ -c^T & b^T \end{pmatrix} = 0 \tag{25}$$

and

$$\eta^T \begin{pmatrix} 0_{m,1} \\ b \\ c \\ -c \\ 0_{1,1} \end{pmatrix} < 0. \tag{26}$$

Let us write the vector η in the form $\begin{pmatrix} u \\ v \\ q \\ r \\ w \end{pmatrix}$, where $u \in \mathbb{R}^m$, $v \in \mathbb{R}^m$, $q \in \mathbb{R}^n$, $r \in \mathbb{R}^n$ and $w \in \mathbb{R}^1$. The vector $w \in \mathbb{R}^1$ will be regarded as a scalar (since we can identify vectors of length 1 with scalars). Thus, $w^T = w$.

Define a vector $s \in \mathbb{R}^n$ by $s = q - r$.

Since $\eta = \begin{pmatrix} u \\ v \\ q \\ r \\ w \end{pmatrix}$, we have

$$\eta^T = \begin{pmatrix} u \\ v \\ q \\ r \\ w \end{pmatrix}^T = \begin{pmatrix} u^T & v^T & q^T & r^T & w^T \end{pmatrix}$$
$$= \begin{pmatrix} u^T & v^T & q^T & r^T & w \end{pmatrix} \tag{27}$$

(because $w^T = w$).

Now, (25) yields

$$\left(\eta^T \left(\begin{pmatrix} 0_{m,n} & -I_m \\ A & 0_{m,m} \\ 0_{n,n} & A^T \\ 0_{n,n} & -A^T \\ -c^T & b^T \end{pmatrix} \right) \right)^T = 0^T = 0.$$

Hence,

$$0 = \left(\eta^T \left(\begin{pmatrix} 0_{m,n} & -I_m \\ A & 0_{m,m} \\ 0_{n,n} & A^T \\ 0_{n,n} & -A^T \\ -c^T & b^T \end{pmatrix} \right) \right)^T$$

$$= \underbrace{\begin{pmatrix} 0_{m,n} & -I_m \\ A & 0_{m,m} \\ 0_{n,n} & A^T \\ 0_{n,n} & -A^T \\ -c^T & b^T \end{pmatrix}^T} \underbrace{\left(\eta^T \right)^T}_{=\eta}$$

$$= \begin{pmatrix} (0_{m,n})^T & A^T & (0_{n,n})^T & (0_{n,n})^T & (-c^T)^T \\ (-I_m)^T & (0_{m,m})^T & (A^T)^T & (-A^T)^T & (b^T)^T \end{pmatrix}$$

$$= \begin{pmatrix} (0_{m,n})^T & A^T & (0_{n,n})^T & (0_{n,n})^T & (-c^T)^T \\ (-I_m)^T & (0_{m,m})^T & (A^T)^T & (-A^T)^T & (b^T)^T \end{pmatrix} \eta$$

$$= \begin{pmatrix} 0_{n,m} & A^T & 0_{n,n} & 0_{n,n} & -c \\ -I_m & 0_{m,m} & A & -A & b \end{pmatrix} \begin{pmatrix} u \\ v \\ q \\ r \\ w \end{pmatrix}$$

$$\left(\begin{array}{c} \text{since } (0_{m,n})^T = 0_{n,m} \text{ and } (-I_m)^T = - \underbrace{(I_m)^T}_{=I_m} = -I_m \text{ and } (0_{m,m})^T = 0_{m,m} \\[2mm] \text{and } (0_{n,n})^T = 0_{n,n} \text{ and } (A^T)^T = A \text{ and } (-A^T)^T = - \underbrace{(A^T)^T}_{=A} = -A \\[2mm] \text{and } (-c^T)^T = - \underbrace{(c^T)^T}_{=c} = -c \text{ and } (b^T)^T = b \text{ and } \eta = \begin{pmatrix} u \\ v \\ q \\ r \\ w \end{pmatrix} \end{array} \right)$$

$$= \begin{pmatrix} 0_{n,m}u + A^Tv + 0_{n,n}q + 0_{n,n}r + (-c)w \\ (-I_m)u + 0_{m,m}v + Aq + (-A)r + bw \end{pmatrix}$$

(by the multiplication rule for block matrices).

Thus,

$$\begin{pmatrix} 0_{n,m}u + A^Tv + 0_{n,n}q + 0_{n,n}r + (-c)w \\ (-I_m)u + 0_{m,m}v + Aq + (-A)r + bw \end{pmatrix} = 0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Thus,

$$0_{n,m}u + A^Tv + 0_{n,n}q + 0_{n,n}r + (-c)w = 0 \tag{28}$$

and

$$(-I_m)u + 0_{m,m}v + Aq + (-A)r + bw = 0. \tag{29}$$

Now, (28) yields

$$0 = \underbrace{0_{n,m}u}_{=0} + A^Tv + \underbrace{0_{n,n}q}_{=0} + \underbrace{0_{n,n}r}_{=0} + \underbrace{(-c)w}_{=-cw} = A^Tv - \underbrace{cw}_{\substack{=wc \\ \text{(since } w \text{ is a scalar)}}} = A^Tv - wc,$$

so that $A^Tv = wc$. Thus, $wc = A^Tv$, so that $\left(\underbrace{wc}_{=A^Tv} \right)^T = \left(A^Tv \right)^T = v^T \underbrace{\left(A^T \right)^T}_{=A} = v^TA$. Hence,

$$v^TA = (wc)^T = wc^T. \tag{30}$$

Also, (29) yields

$$0 = \underbrace{(-I_m)u}_{=-u} + \underbrace{0_{m,m}v}_{=0} + Aq + \underbrace{(-A)r}_{=-Ar} + bw = -u + \underbrace{Aq - Ar}_{=A(q-r)} + \underbrace{bw}_{\substack{=wb \\ \text{(since } w \text{ is a scalar)}}}$$

$$= -u + A\underbrace{(q-r)}_{=s} + wb = -u + As + wb,$$

so that

$$u = As + wb. \tag{31}$$

On the other hand, from (27), we obtain

$$\eta^T \begin{pmatrix} 0_{m,1} \\ b \\ c \\ -c \\ 0_{1,1} \end{pmatrix} = \begin{pmatrix} u^T & v^T & q^T & r^T & w \end{pmatrix} \begin{pmatrix} 0_{m,1} \\ b \\ c \\ -c \\ 0_{1,1} \end{pmatrix}$$

$$= \underbrace{u^T 0_{m,1}}_{=0} + v^T b + q^T c + \underbrace{r^T(-c)}_{=-r^Tc} + \underbrace{w 0_{1,1}}_{=0}$$

(by the multiplication rule for block matrices)

$$= v^T b + \underbrace{q^T c - r^T c}_{=(q^T - r^T)c} = v^T b + \underbrace{(q^T - r^T)}_{=(q-r)^T} c = v^T b + \left(\underbrace{q-r}_{=s} \right)^T c = v^T b + s^T c.$$

Hence, (26) rewrites as
$$v^T b + s^T c < 0. \tag{32}$$

Since $s^T c$ is a scalar, we have $\left(s^T c\right)^T = s^T c$, so that $s^T c = \left(s^T c\right)^T = c^T \underbrace{\left(s^T\right)^T}_{=s} = c^T s$. Hence, (32) rewrites as

$$v^T b + c^T s < 0. \tag{33}$$

We have $\begin{pmatrix} u \\ v \\ q \\ r \\ w \end{pmatrix} = \eta \geq 0$. Hence, we have the five inequalities[38]

$$u \geq 0, \qquad v \geq 0, \qquad q \geq 0, \qquad r \geq 0, \qquad \text{and } w \geq 0.$$

Since $w \geq 0$, we can multiply the inequality (33) with w, and obtain

$$w \left(v^T b + c^T s\right) \leq 0. \tag{34}$$

But from $v \geq 0$, we obtain $v^T \geq 0$ (since the transpose of any nonnegative vector is nonnegative). Thus, Lemma 2.0v (applied to m, v^T and u instead of n, x and y) yields $v^T u \geq 0$ (because $u \geq 0$). Hence,

$$0 \leq v^T \underbrace{u}_{\substack{=As+wb \\ \text{(by (31))}}} = v^T (As + wb) = \underbrace{v^T A}_{\substack{=wc^T \\ \text{(by (30))}}} s + wv^T b$$

$$= wc^T s + wv^T b = w \left(v^T b + c^T s\right). \tag{35}$$

Combined with (33), this easily yields that $w = 0$ [39]. As a consequence, (30) rewrites as

$$v^T A = 0c^T = 0. \tag{36}$$

---

[38]Here, we are using the following simple fact: Let n_1, n_2, n_3, n_4, and n_5 be five nonnegative integers. Let $\alpha_1 \in \mathbb{R}^{n_1}$, $\alpha_2 \in \mathbb{R}^{n_2}$, $\alpha_3 \in \mathbb{R}^{n_3}$, $\alpha_4 \in \mathbb{R}^{n_4}$, and $\alpha_5 \in \mathbb{R}^{n_5}$ be five column vectors. Let $\beta_1 \in \mathbb{R}^{n_1}$, $\beta_2 \in \mathbb{R}^{n_2}$, $\beta_3 \in \mathbb{R}^{n_3}$, $\beta_4 \in \mathbb{R}^{n_4}$, and $\beta_5 \in \mathbb{R}^{n_5}$ be five column vectors.

Assume that $\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \\ \alpha_5 \end{pmatrix} \geq \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_5 \end{pmatrix}$. Then, we have the five inequalities

$$\alpha_1 \geq \beta_1, \qquad \alpha_2 \geq \beta_2, \qquad \alpha_3 \geq \beta_3, \qquad \alpha_4 \geq \beta_4, \qquad \text{and } \alpha_5 \geq \beta_5.$$

[39]*Proof.* Assume that $w > 0$. Then, we can divide the inequality (35) by w, and obtain $0 \leq v^T b + c^T s$. This contradicts (33). This contradiction shows that our assumption was wrong. Thus, we don't have $w > 0$. Hence, $w \leq 0$. Combined with $w \geq 0$, this yields $w = 0$, qed.

Also, (31) becomes

$$u = As + \underbrace{w}_{=0} b = As. \tag{37}$$

Let us summarize what we have found so far: We have found vectors $u \in \mathbb{R}^m$, $v \in \mathbb{R}^m$ and $s \in \mathbb{R}^n$ satisfying $u \geq 0$ and $v \geq 0$ and the equations (33), (36) and (37). These equations (along with $u \geq 0$ and $v \geq 0$) are all that we are going to need from now on; we can forget about q, r, η and w.

Our goal now is to prove that Assertion J1 holds.

By the assumptions of Lemma 2.6d, the set $\{x \in \mathbb{R}^n \mid Ax \leq b\}$ has at least one element. Let z be this element. Then, $z \in \{x \in \mathbb{R}^n \mid Ax \leq b\}$. In other words, $z \in \mathbb{R}^n$ and $Az \leq b$. From $Az \leq b$, we obtain $b \geq Az$, thus $b - Az \geq 0$.

Also, $v^T \geq 0$ (since $v \geq 0$, and since the transpose of any nonnegative vector is nonnegative). Thus, Lemma 2.0v (applied to m, v^T and $b - Az$ instead of n, x and y) yields $v^T (b - Az) \geq 0$ (since $b - Az \geq 0$). Thus, $v^T b - v^T Az = v^T (b - Az) \geq 0$, so that $v^T b \geq \underbrace{v^T A}_{\substack{=0 \\ \text{(by (36))}}} z = 0$. But from (33), it follows that

$$c^T s < - \underbrace{v^T b}_{\geq 0} \leq -0 = 0. \tag{38}$$

Hence, $c^T s \neq 0$, so that $c^T s$ is invertible.

Now, let $\delta \in \mathbb{R}$. We will show that there exists a $t \in \{c^T x \mid x \in \mathbb{R}^n; \ Ax \leq b\}$ such that $t \geq \delta$.

Let $\lambda = \max \left\{ 0, \dfrac{c^T z - \delta}{c^T s} \right\}$ (this is well-defined since $c^T s$ is invertible). Then, $\lambda \geq 0$ and $\lambda \geq \dfrac{c^T z - \delta}{c^T s}$.

Multiplying the inequality $\lambda \geq \dfrac{c^T z - \delta}{c^T s}$ with $c^T s$, we obtain

$$\lambda \cdot c^T s \leq c^T z - \delta \tag{39}$$

(the sign has flipped since $c^T s < 0$).

From (37), we have $As = u \geq 0$. Hence, $\lambda As \geq 0$ (since $\lambda \geq 0$). The vector $z - \lambda s \in \mathbb{R}^n$ satisfies $A(z - \lambda s) = Az - \underbrace{\lambda As}_{\geq 0} \leq Az \leq b$. Hence, $z - \lambda s \in \{x \in \mathbb{R}^n \mid Ax \leq b\}$, so that $c^T (z - \lambda s) \in \{c^T x \mid x \in \mathbb{R}^n; \ Ax \leq b\}$.

But

$$c^T (z - \lambda s) = c^T z - \underbrace{\lambda \cdot c^T s}_{\substack{\leq c^T z - \delta \\ \text{(by (39))}}} \geq c^T z - (c^T z - \delta) = \delta.$$

Since $c^T (z - \lambda s) \in \{c^T x \mid x \in \mathbb{R}^n; \ Ax \leq b\}$, this yields that there exists a $t \in \{c^T x \mid x \in \mathbb{R}^n; \ Ax \leq b\}$ such that $t \geq \delta$ (namely, $t = c^T (z - \lambda s)$).

Now, forget that we fixed δ. We have thus shown that for every $\delta \in \mathbb{R}$, there exists a $t \in \left\{ c^T x \mid x \in \mathbb{R}^n;\ Ax \le b \right\}$ such that $t \ge \delta$. In other words, the set $\left\{ c^T x \mid x \in \mathbb{R}^n;\ Ax \le b \right\}$ is unbounded from above.

Finally, let us prove that the set $\left\{ y \in \mathbb{R}^m \mid y \ge 0 \text{ and } y^T A = c^T \right\}$ is empty. In fact, let $\psi \in \left\{ y \in \mathbb{R}^m \mid y \ge 0 \text{ and } y^T A = c^T \right\}$ be arbitrary. We will derive a contradiction (thus showing that ψ cannot exist).

Since $\psi \in \left\{ y \in \mathbb{R}^m \mid y \ge 0 \text{ and } y^T A = c^T \right\}$, we have $\psi \in \mathbb{R}^m$, $\psi \ge 0$ and $\psi^T A = c^T$. From $\psi^T A = c^T$, it follows that $\psi^T A s = c^T s$. Since $As = u$ (by (37)), this rewrites as $\psi^T u = c^T s$. Combined with (38), this yields $\psi^T u < 0$. But since $\psi \ge 0$, we have $\psi^T \ge 0$ (since the transpose of a nonnegative vector is always nonnegative), and thus Lemma 2.0v (applied to m, ψ^T and u instead of n, x and y) yields $\psi^T u \ge 0$ (since $u \ge 0$). This contradicts $\psi^T u < 0$.

Now, forget that we fixed ψ. Thus, for every $\psi \in \left\{ y \in \mathbb{R}^m \mid y \ge 0 \text{ and } y^T A = c^T \right\}$, we have derived a contradiction. Hence, there exists no $\psi \in \left\{ y \in \mathbb{R}^m \mid y \ge 0 \text{ and } y^T A = c^T \right\}$. In other words, the set $\left\{ y \in \mathbb{R}^m \mid y \ge 0 \text{ and } y^T A = c^T \right\}$ is empty.

Altogether, we have proven that the set $\left\{ c^T x \mid x \in \mathbb{R}^n;\ Ax \le b \right\}$ is unbounded from above, and the set $\left\{ y \in \mathbb{R}^m \mid y \ge 0 \text{ and } y^T A = c^T \right\}$ is empty. In other words, Assertion J1 holds. Thus, at least one of Assertions J1 and J2 holds.

We have thus proven that in Case 2, at least one of Assertions J1 and J2 holds.

We have thus proven that, in each of the Cases 1 and 2, at least one of Assertions J1 and J2 holds. Since these cases cover all possibilities, this yields that, in every situation, at least one of Assertions J1 and J2 holds. Combined with the fact that the Assertions J1 and J2 cannot hold at the same time, this yields that exactly one of the Assertions J1 and J2 holds. This proves Lemma 2.6d. $\qquad\square$

Lemma 2.6d was "one piece" of Theorem 2.6c; here is another "piece":

Lemma 2.6e. Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$. Let A be an $m \times n$-matrix. Let $b \in \mathbb{R}^m$. Let $c \in \mathbb{R}^n$. Assume that the set $\left\{ y \in \mathbb{R}^m \mid y \ge 0,\ y^T A = 0 \text{ and } y^T b < 0 \right\}$ has at least one element. Then:

(a) The set $\left\{ x \in \mathbb{R}^n \mid Ax \le b \right\}$ is empty.

(b) Exactly one of the following two assertions holds:

Assertion K1: The sets $\left\{ x \in \mathbb{R}^n \mid Ax \le b \right\}$ and $\left\{ y \in \mathbb{R}^m \mid y \ge 0 \text{ and } y^T A = c^T \right\}$ are empty.

Assertion K2: The set $\left\{ y^T b \mid y \in \mathbb{R}^m;\ y \ge 0 \text{ and } y^T A = c^T \right\}$ is unbounded from below, and the set $\left\{ x \in \mathbb{R}^n \mid Ax \le b \right\}$ is empty.

Proof of Lemma 2.6e. The Assertions K1 and K2 cannot hold at the same time[40].

---

[40]*Proof.* Assume the opposite. Then, the Assertions K1 and K2 hold at the same time. The set $\left\{ y^T b \mid y \in \mathbb{R}^m;\ y \ge 0 \text{ and } y^T A = c^T \right\}$ is unbounded from below (by Assertion K2), and thus nonempty. In other words, there exists some $y \in \mathbb{R}^m$ satisfying $y \ge 0$ and $y^T A = c^T$. In other words, the set $\left\{ y \in \mathbb{R}^m \mid y \ge 0 \text{ and } y^T A = c^T \right\}$ is nonempty. But this contradicts Assertion K1. This contradiction shows that our assumption was wrong, qed.

We will now show that at least one of these assertions holds.

The set $\left\{y \in \mathbb{R}^m \ \mid \ y \geq 0,\, y^T A = 0 \text{ and } y^T b < 0\right\}$ has at least one element (by assumption). Denote this element by z. Then, $z \in \left\{y \in \mathbb{R}^m \ \mid \ y \geq 0,\, y^T A = 0 \text{ and } y^T b < 0\right\}$. In other words, $z \in \mathbb{R}^m$, $z \geq 0$, $z^T A = 0$ and $z^T b < 0$. From $z \geq 0$, it follows that $z^T \geq 0$ (since the transpose of any nonnegative vector is nonnegative).

Now, it is easy to see that the set $\{x \in \mathbb{R}^n \ \mid \ Ax \leq b\}$ is empty[41]. This proves Lemma 2.6e **(a)**.

Applying Theorem 2.5d to n, m, A^T and c instead of m, n, A and b, we conclude that exactly one of the following two assertions holds:

> *Assertion $K_1 1$:* The system $A^T x = c$ has a nonnegative solution $x \in \mathbb{R}^m$.

> *Assertion $K_1 2$:* There exists a vector $y \in \mathbb{R}^n$ such that $y^T A^T \geq 0$ and $y^T c < 0$.

Thus, we must be in one of the following two cases:

Case 1: Assertion $K_1 1$ holds.

Case 2: Assertion $K_1 2$ holds.

Let us consider Case 1 first. In this case, Assertion $K_1 1$ holds. In other words, the system $A^T x = c$ has a nonnegative solution $x \in \mathbb{R}^m$. Denote this x by u. Thus, $u \in \mathbb{R}^m$ is nonnegative and satisfies $A^T u = c$.

Comparing $\left(A^T u\right)^T = u^T \underbrace{\left(A^T\right)^T}_{=A} = u^T A$ with $\left(\underbrace{A^T u}_{=c}\right)^T = c^T$, we obtain $u^T A = c^T$. Also, u is nonnegative, i. e., we have $u \geq 0$. Since $u \in \mathbb{R}^m$, $u \geq 0$ and $u^T A = c^T$, we have $u \in \left\{y \in \mathbb{R}^m \ \mid \ y \geq 0 \text{ and } y^T A = c^T\right\}$ and thus

$$u^T b \in \left\{y^T b \ \mid \ y \in \mathbb{R}^m;\ y \geq 0 \text{ and } y^T A = c^T\right\}. \tag{40}$$

Now, let $\delta \in \mathbb{R}$. We will show that there exists a $t \in \left\{y^T b \ \mid \ y \in \mathbb{R}^m;\ y \geq 0 \text{ and } y^T A = c^T\right\}$ such that $t \leq \delta$.

In fact, since $z^T b < 0$, we have $z^T b \neq 0$, so that $z^T b$ is invertible. Now, let $\lambda = \max\left\{0, \dfrac{\delta - u^T b}{z^T b}\right\}$ (this is well-defined since $z^T b$ is invertible). Then, $\lambda \geq 0$ and $\lambda \geq \dfrac{\delta - u^T b}{z^T b}$.

---

[41] *Proof.* Let $w \in \{x \in \mathbb{R}^n \ \mid \ Ax \leq b\}$. Then, $w \in \mathbb{R}^n$ and $Aw \leq b$.

From $Aw \leq b$, we obtain $b \geq Aw$, so that $b - Aw \geq 0$. Hence, Lemma 2.0v (applied to m, z^T and $b - Aw$ instead of n, x and y) yields $z^T (b - Aw) \geq 0$ (since $z^T \geq 0$). This contradicts $z^T (b - Aw) = z^T b - \underbrace{z^T A}_{=0} w = z^T b < 0$.

Now, forget that we fixed w. We have thus shown that any $w \in \{x \in \mathbb{R}^n \ \mid \ Ax \leq b\}$ satisfies a contradiction. In other words, there exists no $w \in \{x \in \mathbb{R}^n \ \mid \ Ax \leq b\}$. In other words, the set $\{x \in \mathbb{R}^n \ \mid \ Ax \leq b\}$ is empty, qed.

Multiplying the inequality $\lambda \geq \dfrac{\delta - u^T b}{z^T b}$ with $z^T b$, we obtain

$$\lambda \cdot z^T b \leq \delta - u^T b \tag{41}$$

(the sign flipped since $z^T b < 0$).

From $u \geq 0$, $\lambda \geq 0$ and $z \geq 0$, we obtain $u + \lambda z \geq 0$.

Also,

$$\underbrace{(u + \lambda z)^T}_{=u^T + \lambda z^T} A = \left(u^T + \lambda z^T\right) A = \underbrace{u^T A}_{=c^T} + \lambda \underbrace{z^T A}_{=0} = c^T.$$

Since $u + \lambda z \in \mathbb{R}^m$, $u + \lambda z \geq 0$ and $(u + \lambda z)^T A = c^T$, we have $u + \lambda z \in \left\{ y \in \mathbb{R}^m \mid y \geq 0 \text{ and } y^T A = c^T \right\}$ and thus

$$(u + \lambda z)^T b \in \left\{ y^T b \mid y \in \mathbb{R}^m;\ y \geq 0 \text{ and } y^T A = c^T \right\}.$$

Since

$$(u + \lambda z)^T b = u^T b + \underbrace{\lambda \cdot z^T b}_{\substack{\leq \delta - u^T b \\ \text{(by (41))}}} \leq u^T b + \delta - u^T b = \delta,$$

this yields that there exists a $t \in \left\{ y^T b \mid y \in \mathbb{R}^m;\ y \geq 0 \text{ and } y^T A = c^T \right\}$ such that $t \leq \delta$ (namely, $t = (u + \lambda z)^T b$).

Now, forget that we fixed δ. We thus have proven that for every $\delta \in \mathbb{R}$, there exists a $t \in \left\{ y^T b \mid y \in \mathbb{R}^m;\ y \geq 0 \text{ and } y^T A = c^T \right\}$ such that $t \leq \delta$. In other words, the set $\left\{ y^T b \mid y \in \mathbb{R}^m;\ y \geq 0 \text{ and } y^T A = c^T \right\}$ is unbounded from below.

We now know that the set $\left\{ y^T b \mid y \in \mathbb{R}^m;\ y \geq 0 \text{ and } y^T A = c^T \right\}$ is unbounded from below, and the set $\{ x \in \mathbb{R}^n \mid Ax \leq b \}$ is empty. In other words, Assertion K2 holds. Thus, at least one of Assertions K1 and K2 holds.

Hence, we know that in Case 1, at least one of Assertions K1 and K2 holds.

Let us now consider Case 2. In this case, Assertion $K_1 2$ holds. In other words, there exists a vector $y \in \mathbb{R}^n$ such that $y^T A^T \geq 0$ and $y^T c < 0$. Denote this vector y by w. Then, $w \in \mathbb{R}^n$ is a vector such that $w^T A^T = 0$ and $w^T c < 0$.

Since $w^T A^T = 0$, we have $\left(w^T A^T\right)^T = 0^T = 0$. This rewrites as $Aw = 0$ (because $\left(w^T A^T\right)^T = \underbrace{\left(A^T\right)^T}_{=A} \underbrace{\left(w^T\right)^T}_{=w} = Aw$).

Since $w^T c$ is a scalar, we have $\left(w^T c\right)^T = w^T c < 0$. Since $\left(w^T c\right)^T = c^T \underbrace{\left(w^T\right)^T}_{=w} = c^T w$, this rewrites as $c^T w < 0$.

Now, it is easy to see that the set $\left\{ y \in \mathbb{R}^m \mid y \geq 0 \text{ and } y^T A = c^T \right\}$ is empty[42].

---

[42]*Proof.* Let $s \in \left\{ y \in \mathbb{R}^m \mid y \geq 0 \text{ and } y^T A = c^T \right\}$. Then, $s \in \mathbb{R}^m$ is a vector such that $s \geq 0$ and $s^T A = c^T$. Thus, $\underbrace{s^T A}_{=c^T} w = c^T w < 0$, contradicting $s^T \underbrace{Aw}_{=0} = 0$.

Now, forget that we fixed s. We thus have shown that every $s \in$

Altogether, we now know that the sets $\{x \in \mathbb{R}^n \mid Ax \leq b\}$ and $\{y \in \mathbb{R}^m \mid y \geq 0 \text{ and } y^T A = c^T\}$ are empty. In other words, Assertion K1 holds. Thus, at least one of Assertions K1 and K2 holds.

Hence, we know that in Case 2, at least one of Assertions K1 and K2 holds.

We have thus proven that, in each of the Cases 1 and 2, at least one of Assertions K1 and K2 holds. Since these cases cover all possibilities, this yields that, in every situation, at least one of Assertions K1 and K2 holds. Combined with the fact that the Assertions K1 and K2 cannot hold at the same time, this yields that exactly one of the Assertions K1 and K2 holds. This proves Lemma 2.6e **(b)**. $\square$

Proof of Theorem 2.6c. Corollary 2.5o yields that exactly one of the two assertions O1 and O2 holds.[43]

Thus, we must be in one of the following two cases:

Case 1: Assertion O1 holds.

Case 2: Assertion O2 holds.

Let us consider Case 1 first. In this case, Assertion O1 holds. In other words, there exists a vector $x \in \mathbb{R}^n$ such that $Ax \leq b$. In other words, the set $\{x \in \mathbb{R}^n \mid Ax \leq b\}$ has at least one element. Hence, Lemma 2.6d yields that exactly one of the two assertions J1 and J2 holds.[44] Since Assertion J1 is identical to Assertion I2, and Assertion J2 is identical to Assertion I4, this rewrites as follows: Exactly one of the assertions I2 and I4 holds.

Assertion I1 is false[45]. Assertion I3 is also false.[46]

So we know that Assertions I1 and I3 are both false. Combining this with the fact that exactly one of the assertions I2 and I4 holds, we conclude that exactly one of the four assertions I1, I2, I3 and I4 holds. In other words, Theorem 2.6c is proven in Case 1.

Now, let us consider Case 2. In this case, Assertion O2 holds. In other words, there exists a vector $y \in \mathbb{R}^m$ such that $y \geq 0$, $y^T A = 0$ and $y^T b < 0$. In other words, the set $\{y \in \mathbb{R}^m \mid y \geq 0, y^T A = 0 \text{ and } y^T b < 0\}$ has at least one element. Thus, Lemma 2.6e **(b)** yields that exactly one of the two assertions K1 and K2 holds.[47] Since Assertion K1 is identical to Assertion I1, and Assertion K2 is identical to Assertion I3, this rewrites as follows: Exactly one of the assertions I1 and I3 holds.

$\{y \in \mathbb{R}^m \mid y \geq 0 \text{ and } y^T A = c^T\}$ satisfies a contradiction. In other words, no $s \in \{y \in \mathbb{R}^m \mid y \geq 0 \text{ and } y^T A = c^T\}$ can exist. In other words, the set $\{y \in \mathbb{R}^m \mid y \geq 0 \text{ and } y^T A = c^T\}$ is empty, qed.

[43]See the statement of Corollary 2.5o for these assertions.

[44]See the statement of Lemma 2.6d for these assertions.

[45]*Proof.* Assume the contrary. Thus, Assertion I1 holds. Hence, the set $\{x \in \mathbb{R}^n \mid Ax \leq b\}$ is empty. This contradicts the fact that the set $\{x \in \mathbb{R}^n \mid Ax \leq b\}$ has at least one element. This contradiction shows that our assumption was false, qed.

[46]*Proof.* Assume the contrary. Thus, Assertion I3 holds. Thus, in particular, the set $\{x \in \mathbb{R}^n \mid Ax \leq b\}$ is empty. This contradicts the fact that the set $\{x \in \mathbb{R}^n \mid Ax \leq b\}$ has at least one element. This contradiction shows that our assumption was false, qed.

[47]See the statement of Lemma 2.6e for these assertions.

Assertion I2 is false[48]. Assertion I4 is also false.[49]

So we know that Assertions I2 and I4 are both false. Combining this with the fact that exactly one of the assertions I1 and I3 holds, we conclude that exactly one of the four assertions I1, I2, I3 and I4 holds. In other words, Theorem 2.6c is proven in Case 2.

So we have proven Theorem 2.6c in each of the cases 1 and 2. Since these cases 1 and 2 cover all possibilities, this yields that Theorem 2.6c is proven in all situations. In other words, the proof of Theorem 2.6c is complete. □

Note that Theorem 2.5p is an easy corollary of Theorem 2.6c:

Alternative proof of Theorem 2.5p. The Assertions P1 and P2 cannot hold at the same time[50]. We will now show that at least one of these assertions holds.

Theorem 2.6c yields that exactly one of the four assertions I1, I2, I3 and I4 holds.[51] Thus, we must be in one of the following four cases:

Case 1: Assertion I1 holds.

Case 2: Assertion I2 holds.

Case 3: Assertion I3 holds.

Case 4: Assertion I4 holds.

Let us consider Case 1 first. In this case, Assertion I1 holds. Thus, in particular, the set $\{x \in \mathbb{R}^n \mid Ax \leq b\}$ is empty. This contradicts the assumption that the set $\{x \in \mathbb{R}^n \mid Ax \leq b\}$ has at least one element. Thus, we have obtained a contradiction. Hence, Theorem 2.5p holds in Case 1 (because ex falso quod libet).

Let us now consider Case 2. In this case, Assertion I2 holds. Hence, in particular, the set $\{c^T x \mid x \in \mathbb{R}^n; \ Ax \leq b\}$ is unbounded from above. Thus, there exists a $t \in \{c^T x \mid x \in \mathbb{R}^n; \ Ax \leq b\}$ such that $t \geq \delta + 1$. Consider this t. Then, $t \in \{c^T x \mid x \in \mathbb{R}^n; \ Ax \leq b\}$, so that there exists some $z \in \mathbb{R}^n$ such that $Az \leq b$ and $t = c^T z$. Consider this z. We have $Az \leq b$ and $c^T z = t \geq \delta + 1 > \delta$. Thus, there exists a vector $x \in \mathbb{R}^n$ such that $Ax \leq b$ and $c^T x > \delta$ (namely, $x = z$). In other words, Assertion P1 holds. Hence, at least one of Assertions P1 and P2

---

[48]*Proof.* Assume the contrary. Thus, Assertion I2 holds. Thus, in particular, the set $\{c^T x \mid x \in \mathbb{R}^n; \ Ax \leq b\}$ is unbounded from above, and therefore nonempty. In other words, there exists some $x \in \mathbb{R}^n$ satisfying $Ax \leq b$.

But Lemma 2.6e **(a)** shows that the set $\{x \in \mathbb{R}^n \mid Ax \leq b\}$ is empty. In other words, there exists no $x \in \mathbb{R}^n$ satisfying $Ax \leq b$. This contradicts the fact that there exists some $x \in \mathbb{R}^n$ satisfying $Ax \leq b$. This contradiction shows that our assumption was false, qed.

[49]*Proof.* Assume the contrary. Thus, Assertion I4 holds. Hence, in particular, the number $\max \{c^T x \mid x \in \mathbb{R}^n; \ Ax \leq b\}$ exists. In other words, the set $\{c^T x \mid x \in \mathbb{R}^n; \ Ax \leq b\}$ has a maximum. Hence, this set is nonempty. In other words, there exists some $x \in \mathbb{R}^n$ such that $Ax \leq b$.

But Lemma 2.6e **(a)** shows that the set $\{x \in \mathbb{R}^n \mid Ax \leq b\}$ is empty. In other words, there exists no $x \in \mathbb{R}^n$ satisfying $Ax \leq b$. This contradicts the fact that there exists some $x \in \mathbb{R}^n$ satisfying $Ax \leq b$. This contradiction shows that our assumption was false, qed.

[50]This can be proven just as in our first proof of Theorem 2.5p above.

[51]See the statement of Theorem 2.6c for these assertions.

holds. Combined with the fact that the Assertions P1 and P2 cannot hold at the same time, this yields that exactly one of Assertions P1 and P2 holds. In other words, Theorem 2.5p holds in Case 2.

In Case 3, we obtain the very same contradiction as in Case 1. Hence, Theorem 2.5p holds in Case 3 (because ex falso quod libet).

Let us finally consider Case 4. In this case, Assertion I4 holds. In other words, the numbers $\max \left\{ c^T x \mid x \in \mathbb{R}^n; \ Ax \le b \right\}$ and $\min \left\{ y^T b \mid y \in \mathbb{R}^m; \ y \ge 0 \text{ and } y^T A = c^T \right\}$ exist and satisfy

$$\max \left\{ c^T x \mid x \in \mathbb{R}^n; \ Ax \le b \right\} = \min \left\{ y^T b \mid y \in \mathbb{R}^m; \ y \ge 0 \text{ and } y^T A = c^T \right\}.$$

Hence, we can define an $\varepsilon \in \mathbb{R}$ by

$$\varepsilon = \max \left\{ c^T x \mid x \in \mathbb{R}^n; \ Ax \le b \right\} = \min \left\{ y^T b \mid y \in \mathbb{R}^m; \ y \ge 0 \text{ and } y^T A = c^T \right\}.$$

Now, we must be in one of the following two subcases:

Subcase 4.1: We have $\delta < \varepsilon$.

Subcase 4.2: We have $\delta \ge \varepsilon$.

Let us first consider Subcase 4.1. In this subcase, $\delta < \varepsilon$. Thus, $\varepsilon > \delta$. Since $\varepsilon = \max \left\{ c^T x \mid x \in \mathbb{R}^n; \ Ax \le b \right\} \in \left\{ c^T x \mid x \in \mathbb{R}^n; \ Ax \le b \right\}$ (because the maximum of a set always lies in this set), there exists a $z \in \mathbb{R}^n$ such that $Az \le b$ and $\varepsilon = c^T z$. Consider this z. Then, $Az \le b$ and $c^T z = \varepsilon > \delta$. Hence, there exists a vector $x \in \mathbb{R}^n$ such that $Ax \le b$ and $c^T x > \delta$ (namely, $x = z$). In other words, Assertion P1 holds. Hence, at least one of Assertions P1 and P2 holds. Combined with the fact that the Assertions P1 and P2 cannot hold at the same time, this yields that exactly one of Assertions P1 and P2 holds. In other words, Theorem 2.5p holds in Subcase 4.1.

Let us now consider Subcase 4.2. In this subcase, $\delta \ge \varepsilon$. Thus, $\varepsilon \le \delta$. Since $\varepsilon = \min \left\{ y^T b \mid y \in \mathbb{R}^m; \ y \ge 0 \text{ and } y^T A = c^T \right\} \in \left\{ y^T b \mid y \in \mathbb{R}^m; \ y \ge 0 \text{ and } y^T A = c^T \right\}$ (because the minimum of a set always lies in this set), there exists a $z \in \mathbb{R}^m$ such that $z \ge 0$, $z^T A = c^T$ and $\varepsilon = z^T b$. Consider this z. Then, $z \ge 0$, $z^T A = c^T$ and $z^T b = \varepsilon \le \delta$. Hence, there exists a vector $y \in \mathbb{R}^m$ such that $y \ge 0$, $y^T A = c^T$ and $y^T b \le \delta$ (namely, $y = z$). In other words, Assertion P2 holds. Hence, at least one of Assertions P1 and P2 holds. Combined with the fact that the Assertions P1 and P2 cannot hold at the same time, this yields that exactly one of Assertions P1 and P2 holds. In other words, Theorem 2.5p holds in Subcase 4.2.

We now know that Theorem 2.5p holds in each of Subcases 4.1 and 4.2. Since these Subcases 4.1 and 4.2 cover all of Case 4, this yields that Theorem 2.5p holds in Case 4.

We thus know that Theorem 2.5p holds in each of Cases 1, 2, 3 and 4. Since these Cases 1, 2, 3 and 4 cover all possibilities, this yields that Theorem 2.5p holds in every situation. Theorem 2.5p is thus proven again. $\square$

15. The symmetric version of the Duality theorem

We are now going to prove a variant of the duality theorem which is "symmetric" in the sense that the set we take the minimum of and the set we take the maximum of are defined similarly (in particular, both are parametrized by nonnegative vectors, which was not the case in Theorem 2.6c):

> **Theorem 2.6f.** Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$. Let A be an $m \times n$-matrix. Let $b \in \mathbb{R}^m$. Let $c \in \mathbb{R}^n$. Then, exactly one of the following four assertions holds:
>
> *Assertion Q1:* The sets $\{x \in \mathbb{R}^n \mid x \geq 0 \text{ and } Ax \leq b\}$ and $\{y \in \mathbb{R}^m \mid y \geq 0 \text{ and } y^T A \geq c^T\}$ are empty.
>
> *Assertion Q2:* The set $\{c^T x \mid x \in \mathbb{R}^n;\ x \geq 0 \text{ and } Ax \leq b\}$ is unbounded from above, and the set $\{y \in \mathbb{R}^m \mid y \geq 0 \text{ and } y^T A \geq c^T\}$ is empty.
>
> *Assertion Q3:* The set $\{y^T b \mid y \in \mathbb{R}^m;\ y \geq 0 \text{ and } y^T A \geq c^T\}$ is unbounded from below, and the set $\{x \in \mathbb{R}^n \mid x \geq 0 \text{ and } Ax \leq b\}$ is empty.
>
> *Assertion Q4:* The numbers $\max\{c^T x \mid x \in \mathbb{R}^n;\ x \geq 0 \text{ and } Ax \leq b\}$ and $\min\{y^T b \mid y \in \mathbb{R}^m;\ y \geq 0 \text{ and } y^T A \geq c^T\}$ exist and satisfy
>
> $$\max\{c^T x \mid x \in \mathbb{R}^n;\ x \geq 0 \text{ and } Ax \leq b\}$$
> $$= \min\{y^T b \mid y \in \mathbb{R}^m;\ y \geq 0 \text{ and } y^T A \geq c^T\}.$$

Theorem 2.6f is stronger than Exercise 2.23 in [Schrij17] (both classically and constructively).

It should be noticed that most of the times when Schrijver applies linear programming duality in [Schrij17], it is being applied not in the form of Theorem 2.6 in [Schrij17] (or our Theorem 2.6c), but in the form of Exercise 2.23 in [Schrij17] (or our Theorem 2.6f). In particular, in the proofs of Corollary 3.7b and Corollary 8.3a in [Schrij17], Schrijver is applying Exercise 2.23 when he says that he is using linear programming duality.

Proof of Theorem 2.6f. For any nonnegative integer γ, let I_γ denote the $\gamma \times \gamma$ identity matrix.

Let 0_n be the zero vector in $\mathbb{R}^n$.

Consider the block matrix $\begin{pmatrix} A \\ -I_n \end{pmatrix}$; this is an $(m+n) \times n$-matrix.

Also, consider the block matrix $\begin{pmatrix} b \\ 0_n \end{pmatrix}$; this is an $(m+n) \times 1$-matrix, i. e., a vector in $\mathbb{R}^{m+n}$ (since we identify matrices having only one column with column vectors).

Applying Theorem 2.6c to $m + n$, $\begin{pmatrix} A \\ -I_n \end{pmatrix}$ and $\begin{pmatrix} b \\ 0_n \end{pmatrix}$ instead of m, A and b, we see that exactly one of the following four assertions holds:

Assertion $Q_1 1$: The sets $\left\{ x \in \mathbb{R}^n \mid \begin{pmatrix} A \\ -I_n \end{pmatrix} x \leq \begin{pmatrix} b \\ 0_n \end{pmatrix} \right\}$ and

$\left\{ y \in \mathbb{R}^{m+n} \mid y \geq 0 \text{ and } y^T \begin{pmatrix} A \\ -I_n \end{pmatrix} = c^T \right\}$ are empty.

Assertion $Q_1 2$: The set $\left\{ c^T x \mid x \in \mathbb{R}^n; \begin{pmatrix} A \\ -I_n \end{pmatrix} x \leq \begin{pmatrix} b \\ 0_n \end{pmatrix} \right\}$ is unbounded from above, and the set $\left\{ y \in \mathbb{R}^{m+n} \mid y \geq 0 \text{ and } y^T \begin{pmatrix} A \\ -I_n \end{pmatrix} = c^T \right\}$ is empty.

Assertion $Q_1 3$: The set $\left\{ y^T \begin{pmatrix} b \\ 0_n \end{pmatrix} \mid y \in \mathbb{R}^{m+n}; y \geq 0 \text{ and } y^T \begin{pmatrix} A \\ -I_n \end{pmatrix} = c^T \right\}$ is unbounded from below, and the set $\left\{ x \in \mathbb{R}^n \mid \begin{pmatrix} A \\ -I_n \end{pmatrix} x \leq \begin{pmatrix} b \\ 0_n \end{pmatrix} \right\}$ is empty.

Assertion $Q_1 4$: The numbers $\max \left\{ c^T x \mid x \in \mathbb{R}^n; \begin{pmatrix} A \\ -I_n \end{pmatrix} x \leq \begin{pmatrix} b \\ 0_n \end{pmatrix} \right\}$ and $\min \left\{ y^T \begin{pmatrix} b \\ 0_n \end{pmatrix} \mid y \in \mathbb{R}^{m+n}; y \geq 0 \text{ and } y^T \begin{pmatrix} A \\ -I_n \end{pmatrix} = c^T \right\}$ exist and satisfy

$$
\max \left\{ c^T x \mid x \in \mathbb{R}^n; \begin{pmatrix} A \\ -I_n \end{pmatrix} x \leq \begin{pmatrix} b \\ 0_n \end{pmatrix} \right\}
$$
$$
= \min \left\{ y^T \begin{pmatrix} b \\ 0_n \end{pmatrix} \mid y \in \mathbb{R}^{m+n}; y \geq 0 \text{ and } y^T \begin{pmatrix} A \\ -I_n \end{pmatrix} = c^T \right\}.
$$

We will now prove that these assertions $Q_1 1$, $Q_1 2$, $Q_1 3$, and $Q_1 4$ are equivalent to the assertions Q1, Q2, Q3, and Q4, respectively.

First, we notice that

$$
\left\{ c^T x \mid x \in \mathbb{R}^n; \begin{pmatrix} A \\ -I_n \end{pmatrix} x \leq \begin{pmatrix} b \\ 0_n \end{pmatrix} \right\}
$$
$$
= \left\{ c^T x \mid x \in \mathbb{R}^n; x \geq 0 \text{ and } Ax \leq b \right\}. \tag{42}
$$

52

Next, we notice that

$$\left\{ y^T \begin{pmatrix} b \\ 0_n \end{pmatrix} \ \mid \ y \in \mathbb{R}^{m+n}; \ y \ge 0 \text{ and } y^T \begin{pmatrix} A \\ -I_n \end{pmatrix} = c^T \right\}$$
$$= \left\{ y^T b \ \mid \ y \in \mathbb{R}^m; \ y \ge 0 \text{ and } y^T A \ge c^T \right\}. \tag{44}$$

53

---

[52] *Proof of (42):* For each $x \in \mathbb{R}^n$, we have

$$\begin{pmatrix} b \\ 0_n \end{pmatrix} - \underbrace{\begin{pmatrix} A \\ -I_n \end{pmatrix} x}_{\substack{= \begin{pmatrix} Ax \\ -x \end{pmatrix} \\ \text{(by the multiplication rule} \\ \text{for block matrices)}}} = \begin{pmatrix} b \\ 0_n \end{pmatrix} - \begin{pmatrix} Ax \\ -x \end{pmatrix} = \begin{pmatrix} b - Ax \\ 0_n - (-x) \end{pmatrix}$$

$$= \begin{pmatrix} b - Ax \\ x \end{pmatrix}. \tag{43}$$

Now, for each $x \in \mathbb{R}^n$, we have the following chain of equivalences:

$$\left(\begin{pmatrix} A \\ -I_n \end{pmatrix} x \le \begin{pmatrix} b \\ 0_n \end{pmatrix} \right)$$
$$\iff \left(\begin{pmatrix} b \\ 0_n \end{pmatrix} \ge \begin{pmatrix} A \\ -I_n \end{pmatrix} x \right) \iff \left(\begin{pmatrix} b \\ 0_n \end{pmatrix} - \begin{pmatrix} Ax \\ -x \end{pmatrix} \ge 0 \right)$$
$$\iff \left(\begin{pmatrix} b - Ax \\ x \end{pmatrix} \ge 0 \right) \qquad \text{(by (43))}$$
$$\iff (b - Ax \ge 0 \text{ and } x \ge 0) \qquad \text{(by Lemma 2.5h)}$$
$$\iff \left(x \ge 0 \text{ and } \underbrace{b - Ax \ge 0}_{\iff (b \ge Ax) \iff (Ax \le b)} \right)$$
$$\iff (x \ge 0 \text{ and } Ax \le b).$$

Hence,

$$\left\{ c^T x \ \mid \ x \in \mathbb{R}^n; \ \begin{pmatrix} A \\ -I_n \end{pmatrix} x \le \begin{pmatrix} b \\ 0_n \end{pmatrix} \right\} = \left\{ c^T x \ \mid \ x \in \mathbb{R}^n; \ x \ge 0 \text{ and } Ax \le b \right\}.$$

This proves (42).

[53] *Proof of (44):* Let $\lambda \in \left\{ y^T \begin{pmatrix} b \\ 0_n \end{pmatrix} \ \mid \ y \in \mathbb{R}^{m+n}; \ y \ge 0 \text{ and } y^T \begin{pmatrix} A \\ -I_n \end{pmatrix} = c^T \right\}$ be arbitrary. Then, there exists an $y \in \mathbb{R}^{m+n}$ such that $y \ge 0$, $y^T \begin{pmatrix} A \\ -I_n \end{pmatrix} = c^T$ and $\lambda = y^T \begin{pmatrix} b \\ 0_n \end{pmatrix}$. Denote this y by ξ. Thus, $\xi \in \mathbb{R}^{m+n}$ satisfies $\xi \ge 0$, $\xi^T \begin{pmatrix} A \\ -I_n \end{pmatrix} = c^T$ and $\lambda = \xi^T \begin{pmatrix} b \\ 0_n \end{pmatrix}$.

Let us write the vector ξ in the form $\begin{pmatrix} u \\ v \end{pmatrix}$, where $u \in \mathbb{R}^m$ and $v \in \mathbb{R}^n$.

We have $\begin{pmatrix} u \\ v \end{pmatrix} = \xi \geq 0$. Thus, Lemma 2.5h **(b)** (applied to $N = m$, $M = n$, $x = u$ and $y = v$) yields that $u \geq 0$ and $v \geq 0$. Thus, $u^T \geq 0$ and $v^T \geq 0$.

Since $\xi = \begin{pmatrix} u \\ v \end{pmatrix}$, we have $\xi^T = \begin{pmatrix} u \\ v \end{pmatrix}^T = \begin{pmatrix} u^T & v^T \end{pmatrix}$, and thus

$$\xi^T \begin{pmatrix} A \\ -I_n \end{pmatrix} = \begin{pmatrix} u^T & v^T \end{pmatrix} \begin{pmatrix} A \\ -I_n \end{pmatrix}$$
$$= u^T A + \underbrace{v^T (-I_n)}_{=-v^T} \qquad \text{(by the multiplication rule for block matrices)}$$
$$= u^T A - v^T.$$

Compared with $\xi^T \begin{pmatrix} A \\ -I_n \end{pmatrix} = c^T$, this yields $c^T = u^T A - v^T$. In other words,

$$u^T A = c^T + v^T.$$

Hence, $u^T A = c^T + \underbrace{v^T}_{\geq 0} \geq c^T$.

So we now know that $u \in \mathbb{R}^m$, $u \geq 0$ and $u^T A \geq c^T$. In other words, $u \in \{y \in \mathbb{R}^m \mid y \geq 0 \text{ and } y^T A \geq c^T\}$, so that

$$u^T b \in \{y^T b \mid y \in \mathbb{R}^m; \ y \geq 0 \text{ and } y^T A \geq c^T\}. \tag{45}$$

Now,

$$\lambda = \underbrace{\xi^T}_{=\begin{pmatrix} u^T & v^T \end{pmatrix}} \begin{pmatrix} b \\ 0_n \end{pmatrix} = \begin{pmatrix} u^T & v^T \end{pmatrix} \begin{pmatrix} b \\ 0_n \end{pmatrix}$$
$$= u^T b + \underbrace{v^T 0_n}_{=0} \qquad \text{(by the multiplication rule for block matrices)}$$
$$= u^T b \in \{y^T b \mid y \in \mathbb{R}^m; \ y \geq 0 \text{ and } y^T A \geq c^T\} \qquad \text{(by (45))}.$$

Now, forget that we fixed λ. We thus have proven that every $\lambda \in \left\{ y^T \begin{pmatrix} b \\ 0_n \end{pmatrix} \mid y \in \mathbb{R}^{m+n}; \ y \geq 0 \text{ and } y^T \begin{pmatrix} A \\ -I_n \end{pmatrix} = c^T \right\}$ satisfies $\lambda \in \{y^T b \mid y \in \mathbb{R}^m; \ y \geq 0 \text{ and } y^T A \geq c^T\}$. In other words,

$$\left\{ y^T \begin{pmatrix} b \\ 0_n \end{pmatrix} \mid y \in \mathbb{R}^{m+n}; \ y \geq 0 \text{ and } y^T \begin{pmatrix} A \\ -I_n \end{pmatrix} = c^T \right\}$$
$$\subseteq \{y^T b \mid y \in \mathbb{R}^m; \ y \geq 0 \text{ and } y^T A \geq c^T\}. \tag{46}$$

Now, let $\mu \in \{y^T b \mid y \in \mathbb{R}^m; \ y \geq 0 \text{ and } y^T A \geq c^T\}$ be arbitrary. Then, there exists a $y \in \mathbb{R}^m$ such that $y \geq 0$, $y^T A \geq c^T$ and $\mu = y^T b$. Denote this y by u. Thus, $u \in \mathbb{R}^m$ satisfies $u \geq 0$, $u^T A \geq c^T$ and $\mu = u^T b$.

We have $u^T A \geq c^T$. In other words, $u^T A - c^T \geq 0$.

Let v be the vector $\left(u^T A - c^T \right)^T$. Then, $v^T = \left(\left(u^T A - c^T \right)^T \right)^T = u^T A - c^T \geq 0$; therefore, $v \geq 0$.

Lemma 2.5h **(a)** (applied to $N = m$, $M = n$, $x = u$ and $y = v$) yields that $\begin{pmatrix} u \\ v \end{pmatrix} \geq 0$ (since $u \geq 0$ and $v \geq 0$). But on the other hand,

$$\underbrace{\begin{pmatrix} u \\ v \end{pmatrix}^T}_{= \left(\begin{array}{cc} u^T & v^T \end{array} \right)} \begin{pmatrix} A \\ -I_n \end{pmatrix} = \left(\begin{array}{cc} u^T & v^T \end{array} \right) \begin{pmatrix} A \\ -I_n \end{pmatrix} = u^T A + \underbrace{v^T \left(-I_n \right)}_{= -v^T}$$

$$\text{(by the multiplication rule for block matrices)}$$

$$= u^T A - \underbrace{v^T}_{= u^T A - c^T} = u^T A - \left(u^T A - c^T \right) = c^T.$$

Now, we know that our $\begin{pmatrix} u \\ v \end{pmatrix} \in \mathbb{R}^{m+n}$ satisfies $\begin{pmatrix} u \\ v \end{pmatrix} \geq 0$ and $\begin{pmatrix} u \\ v \end{pmatrix}^T \begin{pmatrix} A \\ -I_n \end{pmatrix} = c^T$. In other words,

$$\begin{pmatrix} u \\ v \end{pmatrix} \in \left\{ y \in \mathbb{R}^{m+n} \mid y \geq 0 \text{ and } y^T \begin{pmatrix} A \\ -I_n \end{pmatrix} = c^T \right\},$$

and thus

$$\begin{pmatrix} u \\ v \end{pmatrix}^T \begin{pmatrix} b \\ 0_n \end{pmatrix} \in \left\{ y^T \begin{pmatrix} b \\ 0_n \end{pmatrix} \mid y \in \mathbb{R}^{m+n}; \ y \geq 0 \text{ and } y^T \begin{pmatrix} A \\ -I_n \end{pmatrix} = c^T \right\}.$$

Since

$$\underbrace{\begin{pmatrix} u \\ v \end{pmatrix}^T}_{= \left(\begin{array}{cc} u^T & v^T \end{array} \right)} \begin{pmatrix} b \\ 0_n \end{pmatrix} = \left(\begin{array}{cc} u^T & v^T \end{array} \right) \begin{pmatrix} b \\ 0_n \end{pmatrix} = \underbrace{u^T b}_{= \mu} + \underbrace{v^T 0_n}_{= 0}$$

$$\text{(by the multiplication rule for block matrices)}$$

$$= \mu,$$

this rewrites as

$$\mu \in \left\{ y^T \begin{pmatrix} b \\ 0_n \end{pmatrix} \mid y \in \mathbb{R}^{m+n}; \ y \geq 0 \text{ and } y^T \begin{pmatrix} A \\ -I_n \end{pmatrix} = c^T \right\}.$$

Now, forget that we fixed μ. We thus have proven that every $\mu \in \left\{ y^T b \mid y \in \mathbb{R}^m; \ y \geq 0 \text{ and } y^T A \geq c^T \right\}$ satisfies $\mu \in \left\{ y^T \begin{pmatrix} b \\ 0_n \end{pmatrix} \mid y \in \mathbb{R}^{m+n}; \ y \geq 0 \text{ and } y^T \begin{pmatrix} A \\ -I_n \end{pmatrix} = c^T \right\}$. In other words,

$$\left\{ y^T b \mid y \in \mathbb{R}^m; \ y \geq 0 \text{ and } y^T A \geq c^T \right\}$$

$$\subseteq \left\{ y^T \begin{pmatrix} b \\ 0_n \end{pmatrix} \mid y \in \mathbb{R}^{m+n}; \ y \geq 0 \text{ and } y^T \begin{pmatrix} A \\ -I_n \end{pmatrix} = c^T \right\}.$$

Next, we notice that we have the following equivalence of assertions:

$$\left(\text{the set } \left\{ x \in \mathbb{R}^n \;\middle|\; \begin{pmatrix} A \\ -I_n \end{pmatrix} x \leq \begin{pmatrix} b \\ 0_n \end{pmatrix} \right\} \text{ is empty}\right)$$
$$\Longleftrightarrow \left(\text{the set } \{ x \in \mathbb{R}^n \;\mid\; x \geq 0 \text{ and } Ax \leq b \} \text{ is empty}\right). \tag{47}$$

[54]

Also, we have the following equivalence of assertions:

$$\left(\text{the set } \left\{ y \in \mathbb{R}^{m+n} \;\middle|\; y \geq 0 \text{ and } y^T \begin{pmatrix} A \\ -I_n \end{pmatrix} = c^T \right\} \text{ is empty}\right)$$
$$\Longleftrightarrow \left(\text{the set } \{ y \in \mathbb{R}^m \;\mid\; y \geq 0 \text{ and } y^T A \geq c^T \} \text{ is empty}\right). \tag{48}$$

Combined with (46), this yields

$$\left\{ y^T \begin{pmatrix} b \\ 0_n \end{pmatrix} \;\middle|\; y \in \mathbb{R}^{m+n};\ y \geq 0 \text{ and } y^T \begin{pmatrix} A \\ -I_n \end{pmatrix} = c^T \right\}$$
$$= \left\{ y^T b \;\mid\; y \in \mathbb{R}^m;\ y \geq 0 \text{ and } y^T A \geq c^T \right\}.$$

This proves (44).

[54] *Proof of (47):* We have the following equivalence of assertions:

$$\left(\text{the set } \left\{ x \in \mathbb{R}^n \;\middle|\; \begin{pmatrix} A \\ -I_n \end{pmatrix} x \leq \begin{pmatrix} b \\ 0_n \end{pmatrix} \right\} \text{ is empty}\right)$$
$$\Longleftrightarrow \left(\text{there exists no } x \in \mathbb{R}^n \text{ satisfying } \begin{pmatrix} A \\ -I_n \end{pmatrix} x \leq \begin{pmatrix} b \\ 0_n \end{pmatrix}\right)$$
$$\Longleftrightarrow \left(\text{the set } \underbrace{\left\{ c^T x \;\middle|\; x \in \mathbb{R}^n;\ \begin{pmatrix} A \\ -I_n \end{pmatrix} x \leq \begin{pmatrix} b \\ 0_n \end{pmatrix} \right\}}_{\substack{=\left\{ c^T x \;\mid\; x \in \mathbb{R}^n;\ x \geq 0 \text{ and } Ax \leq b \right\} \\ \text{(by (42))}}} \text{ is empty}\right)$$
$$\Longleftrightarrow \left(\text{the set } \left\{ c^T x \;\mid\; x \in \mathbb{R}^n;\ x \geq 0 \text{ and } Ax \leq b \right\} \text{ is empty}\right)$$
$$\Longleftrightarrow \left(\text{there exists no } x \in \mathbb{R}^n \text{ satisfying } x \geq 0 \text{ and } Ax \leq b\right)$$
$$\Longleftrightarrow \left(\text{the set } \{ x \in \mathbb{R}^n \;\mid\; x \geq 0 \text{ and } Ax \leq b \} \text{ is empty}\right).$$

This proves (47).

55

Now, we make the following four observations:

- Assertion $Q_1 1$ is equivalent to Assertion Q1.[56]

- Assertion $Q_1 2$ is equivalent to Assertion Q2.[57]

---

[55]*Proof of (48):* We have the following equivalence of assertions:

$$\left(\text{the set } \left\{ y \in \mathbb{R}^{m+n} \mid y \geq 0 \text{ and } y^T \begin{pmatrix} A \\ -I_n \end{pmatrix} = c^T \right\} \text{ is empty} \right)$$

$$\Longleftrightarrow \left(\text{there exists no } y \in \mathbb{R}^{m+n} \text{ satisfying } y \geq 0 \text{ and } y^T \begin{pmatrix} A \\ -I_n \end{pmatrix} = c^T \right)$$

$$\Longleftrightarrow \left(\text{the set } \underbrace{\left\{ y^T \begin{pmatrix} b \\ 0_n \end{pmatrix} \mid y \in \mathbb{R}^{m+n}; \ y \geq 0 \text{ and } y^T \begin{pmatrix} A \\ -I_n \end{pmatrix} = c^T \right\}}_{= \left\{ y^T b \mid y \in \mathbb{R}^m; \ y \geq 0 \text{ and } y^T A \geq c^T \right\} \atop \text{(by (44))}} \text{ is empty} \right)$$

$$\Longleftrightarrow \left(\text{the set } \left\{ y^T b \mid y \in \mathbb{R}^m; \ y \geq 0 \text{ and } y^T A \geq c^T \right\} \text{ is empty} \right)$$

$$\Longleftrightarrow \left(\text{there exists no } y \in \mathbb{R}^m \text{ satisfying } y \geq 0 \text{ and } y^T A \geq c^T \right)$$

$$\Longleftrightarrow \left(\text{the set } \left\{ y \in \mathbb{R}^m \mid y \geq 0 \text{ and } y^T A \geq c^T \right\} \text{ is empty} \right).$$

This proves (48).

[56]*Proof.* Assertion $Q_1 1$ says that the sets $\left\{ x \in \mathbb{R}^n \mid \begin{pmatrix} A \\ -I_n \end{pmatrix} x \leq \begin{pmatrix} b \\ 0_n \end{pmatrix} \right\}$ and $\left\{ y \in \mathbb{R}^{m+n} \mid y \geq 0 \text{ and } y^T \begin{pmatrix} A \\ -I_n \end{pmatrix} = c^T \right\}$ are empty. Thus, we have the following equivalence of assertions:

(Assertion $Q_1 1$ holds)

$$\Longleftrightarrow \underbrace{\left(\text{the set } \left\{ x \in \mathbb{R}^n \mid \begin{pmatrix} A \\ -I_n \end{pmatrix} x \leq \begin{pmatrix} b \\ 0_n \end{pmatrix} \right\} \text{ is empty} \right)}_{\Longleftrightarrow \ (\text{the set } \{x \in \mathbb{R}^n \mid x \geq 0 \text{ and } Ax \leq b\} \text{ is empty}) \atop \text{(by (47))}}$$

$$\wedge \underbrace{\left(\text{the set } \left\{ y \in \mathbb{R}^{m+n} \mid y \geq 0 \text{ and } y^T \begin{pmatrix} A \\ -I_n \end{pmatrix} = c^T \right\} \text{ is empty} \right)}_{\Longleftrightarrow \ \left(\text{the set } \left\{ y \in \mathbb{R}^m \mid y \geq 0 \text{ and } y^T A \geq c^T \right\} \text{ is empty} \right) \atop \text{(by (48))}}$$

$$\Longleftrightarrow \left(\text{the set } \left\{ x \in \mathbb{R}^n \mid x \geq 0 \text{ and } Ax \leq b \right\} \text{ is empty} \right)$$

$$\wedge \left(\text{the set } \left\{ y \in \mathbb{R}^m \mid y \geq 0 \text{ and } y^T A \geq c^T \right\} \text{ is empty} \right)$$

$$\Longleftrightarrow (\text{Assertion Q1 holds})$$

(because Assertion Q1 says that the sets $\{x \in \mathbb{R}^n \mid x \geq 0 \text{ and } Ax \leq b\}$ and $\{y \in \mathbb{R}^m \mid y \geq 0 \text{ and } y^T A \geq c^T\}$ are empty). In other words, Assertion $Q_1 1$ is equivalent to Assertion Q1.

[57]*Proof.* Assertion $Q_1 2$ says that the set $\left\{ c^T x \mid x \in \mathbb{R}^n; \ \begin{pmatrix} A \\ -I_n \end{pmatrix} x \leq \begin{pmatrix} b \\ 0_n \end{pmatrix} \right\}$ is un-

---

- Assertion $Q_1 3$ is equivalent to Assertion Q3.[58]

---

bounded from above, and that the set $\left\{ y \in \mathbb{R}^{m+n} \;\mid\; y \geq 0 \text{ and } y^T \begin{pmatrix} A \\ -I_n \end{pmatrix} = c^T \right\}$ is empty. Thus, we have the following equivalence of assertions:

(Assertion $Q_1 2$ holds)

$$\Longleftrightarrow \left(\begin{array}{c} \text{the set } \underbrace{\left\{ c^T x \;\mid\; x \in \mathbb{R}^n; \; \begin{pmatrix} A \\ -I_n \end{pmatrix} x \leq \begin{pmatrix} b \\ 0_n \end{pmatrix} \right\}}_{\substack{= \left\{ c^T x \;\mid\; x \in \mathbb{R}^n; \; x \geq 0 \text{ and } Ax \leq b \right\} \\ \text{(by (42))}}} \text{ is unbounded from above} \\ \\ \wedge \; \underbrace{\left(\text{the set } \left\{ y \in \mathbb{R}^{m+n} \;\mid\; y \geq 0 \text{ and } y^T \begin{pmatrix} A \\ -I_n \end{pmatrix} = c^T \right\} \text{ is empty} \right)}_{\substack{\Longleftrightarrow \left(\text{the set } \left\{ y \in \mathbb{R}^m \;\mid\; y \geq 0 \text{ and } y^T A \geq c^T \right\} \text{ is empty} \right) \\ \text{(by (48))}}} \end{array} \right)$$

$\Longleftrightarrow$ (the set $\left\{ c^T x \;\mid\; x \in \mathbb{R}^n; \; x \geq 0 \text{ and } Ax \leq b \right\}$ is unbounded from above)

$\qquad \wedge$ (the set $\left\{ y \in \mathbb{R}^m \;\mid\; y \geq 0 \text{ and } y^T A \geq c^T \right\}$ is empty)

$\Longleftrightarrow$ (Assertion Q2 holds)

(since Assertion Q2 says that the set $\left\{ c^T x \;\mid\; x \in \mathbb{R}^n; \; x \geq 0 \text{ and } Ax \leq b \right\}$ is unbounded from above, and that the set $\left\{ y \in \mathbb{R}^m \;\mid\; y \geq 0 \text{ and } y^T A \geq c^T \right\}$ is empty). In other words, Assertion $Q_1 2$ is equivalent to Assertion Q2.

[58]*Proof.* Assertion $Q_1 3$ says that the set $\left\{ y^T \begin{pmatrix} b \\ 0_n \end{pmatrix} \;\mid\; y \in \mathbb{R}^{m+n}; \; y \geq 0 \text{ and } y^T \begin{pmatrix} A \\ -I_n \end{pmatrix} = c^T \right\}$ is unbounded from below, and that the set $\left\{ x \in \mathbb{R}^n \;\mid\; \begin{pmatrix} A \\ -I_n \end{pmatrix} x \leq \begin{pmatrix} b \\ 0_n \end{pmatrix} \right\}$ is empty.

Thus, we have the following equivalence of assertions:

(Assertion $Q_1 3$ holds)

$$\Longleftrightarrow \left(\begin{array}{c} \text{the set } \underbrace{\left\{ y^T \begin{pmatrix} b \\ 0_n \end{pmatrix} \;\mid\; y \in \mathbb{R}^{m+n}; \; y \geq 0 \text{ and } y^T \begin{pmatrix} A \\ -I_n \end{pmatrix} = c^T \right\}}_{\substack{= \left\{ y^T b \;\mid\; y \in \mathbb{R}^m; \; y \geq 0 \text{ and } y^T A \geq c^T \right\} \\ \text{(by (44))}}} \text{ is unbounded from below} \\ \\ \wedge \; \underbrace{\left(\text{the set } \left\{ x \in \mathbb{R}^n \;\mid\; \begin{pmatrix} A \\ -I_n \end{pmatrix} x \leq \begin{pmatrix} b \\ 0_n \end{pmatrix} \right\} \text{ is empty} \right)}_{\substack{\Longleftrightarrow \left(\text{the set } \left\{ x \in \mathbb{R}^n \;\mid\; x \geq 0 \text{ and } Ax \leq b \right\} \text{ is empty} \right) \\ \text{(by (47))}}} \end{array} \right)$$

$\Longleftrightarrow$ (the set $\left\{ y^T b \;\mid\; y \in \mathbb{R}^m; \; y \geq 0 \text{ and } y^T A \geq c^T \right\}$ is unbounded from below)

$\qquad \wedge$ (the set $\left\{ x \in \mathbb{R}^n \;\mid\; x \geq 0 \text{ and } Ax \leq b \right\}$ is empty)

$\Longleftrightarrow$ (Assertion Q3 holds)

(since Assertion Q3 says that the set $\left\{ y^T b \;\mid\; y \in \mathbb{R}^m; \; y \geq 0 \text{ and } y^T A \geq c^T \right\}$ is unbounded

- Assertion $Q_1 4$ is equivalent to Assertion Q4.[59]

Combining these observations, we see that Assertions $Q_1 1$, $Q_1 2$, $Q_1 3$, $Q_1 4$ are equivalent to Assertions Q1, Q2, Q3, Q4, respectively. Hence, we have the following equivalence of assertions:

$$\text{(exactly one of Assertions } Q_1 1, Q_1 2, Q_1 3, Q_1 4 \text{ holds)}$$
$$\Longleftrightarrow \text{ (exactly one of Assertions } Q1, Q2, Q3, Q4 \text{ holds)} .$$

Since we know that exactly one of Assertions $Q_1 1$, $Q_1 2$, $Q_1 3$, $Q_1 4$ holds, this yields that exactly one of Assertions Q1, Q2, Q3, Q4 holds. In other words, Theorem 2.6f is proven. $\qquad \square$

16. Appendix: Proofs omitted from the early sections

We have promised to give proofs for various statements made in Section 1, in Section 2, in Section 3, in Section 4, and in Section 9. Let us now fulfill this promise.

16.1. Proofs for Section 1

Proof of Proposition 2.0a. We will first show that

$$\left(\begin{array}{c} \text{if } J \text{ is any finite subset of } I, \text{ and if } (\mu_i)_{i \in J} \text{ is a} \\ \text{family of nonnegative reals indexed by } J \text{ such} \\ \text{that } \sum_{i \in J} \mu_i = 1, \text{ then } \sum_{i \in J} \mu_i x_i \in C \end{array} \right). \tag{49}$$

[*Proof of (49):* We shall prove (49) by strong induction over $|J|$:

Induction step:[60] Fix $N \in \mathbb{N}$. Assume that (49) holds in the case when $|J| < N$. We must prove that (49) holds in the case when $|J| = N$.

---

from below, and that the set $\{x \in \mathbb{R}^n \mid x \geq 0 \text{ and } Ax \leq b\}$ is empty). In other words, Assertion $Q_1 3$ is equivalent to Assertion Q3.

[59]*Proof.* If we use (42) to replace every $\{c^T x \mid x \in \mathbb{R}^n; \ x \geq 0 \text{ and } Ax \leq b\}$ by $\left\{ c^T x \mid x \in \mathbb{R}^n; \ \begin{pmatrix} A \\ -I_n \end{pmatrix} x \leq \begin{pmatrix} b \\ 0_n \end{pmatrix} \right\}$, and if we use (44) to replace every $\{y^T b \mid y \in \mathbb{R}^m; \ y \geq 0 \text{ and } y^T A \geq c^T\}$ by $\left\{ y^T \begin{pmatrix} b \\ 0_n \end{pmatrix} \mid y \in \mathbb{R}^{m+n}; \ y \geq 0 \text{ and } y^T \begin{pmatrix} A \\ -I_n \end{pmatrix} = c^T \right\}$, then Assertion Q4 turns into Assertion $Q_1 4$. Hence, Assertion $Q_1 4$ is equivalent to Assertion Q4.

[60]A strong induction needs no induction base.

We have assumed that (49) holds in the case when $|J| < N$. In other words,

$$\begin{pmatrix} \text{if } J \text{ is any finite subset of } I \text{ satisfying } |J| < N, \text{ and} \\ \text{if } (\mu_i)_{i \in J} \text{ is a family of nonnegative reals indexed by } J \\ \text{such that } \sum_{i \in J} \mu_i = 1, \text{ then } \sum_{i \in J} \mu_i x_i \in C \end{pmatrix}. \qquad (50)$$

Now, let us prove that (49) holds in the case when $|J| = N$. Thus, let J be any finite subset of I satisfying $|J| = N$, and let $(\mu_i)_{i \in J}$ be a family of nonnegative reals indexed by J such that $\sum_{i \in J} \mu_i = 1$. We shall show that $\sum_{i \in J} \mu_i x_i \in C$.

If $|J| \leq 1$, then $\sum_{i \in J} \mu_i x_i \in C$ holds[61]. Hence, for the rest of our proof of $\sum_{i \in J} \mu_i x_i \in C$, we WLOG assume that we don't have $|J| \leq 1$.

There exists a $k \in J$ satisfying $\mu_k < 1$ [62]. Consider such a k. We have $\mu_k < 1$, thus $1 - \mu_k > 0$.

For every $i \in J \setminus \{k\}$, set $\mu_i' = \dfrac{\mu_i}{1 - \mu_k}$. This is a nonnegative real (since μ_i is a nonnegative real, and since $1 - \mu_k > 0$). Hence, $(\mu_i')_{i \in J \setminus \{k\}}$ is a family of nonnegative reals indexed by $J \setminus \{k\}$. Moreover, $k \in J$ and thus $|J \setminus \{k\}| = \underbrace{|J|}_{=N} - 1 = N - 1 < N$. Furthermore,

$$1 = \sum_{i \in J} \mu_i = \underbrace{\sum_{\substack{i \in J; \\ i \neq k}} \mu_i}_{= \sum_{i \in J \setminus \{k\}}} + \underbrace{\sum_{\substack{i \in J; \\ i = k}} \mu_i}_{\substack{= \mu_k \\ (\text{since } k \in J)}} = \sum_{i \in J \setminus \{k\}} \mu_i + \mu_k.$$

Solving this equation for $\sum_{i \in J \setminus \{k\}} \mu_i$ gives us $\sum_{i \in J \setminus \{k\}} \mu_i = 1 - \mu_k$. Hence, $\dfrac{\sum_{i \in J \setminus \{k\}} \mu_i}{1 - \mu_k} =$

---

[61] *Proof.* Assume that $|J| \leq 1$.

The sum $\sum_{i \in J} \mu_i$ cannot be empty (since $\sum_{i \in J} \mu_i = 1 \neq 0$). Hence, J cannot be the empty set. Thus, $|J| \geq 1$. Therefore, $|J| = 1$ (since $|J| \leq 1$). In other words, $J = \{k\}$ for some $k \in J$. Consider this k. We have $k \in J \subseteq I$ and thus $x_k \in C$. Now, from $J = \{k\}$, we obtain $\sum_{i \in J} \mu_i = \sum_{i \in \{k\}} \mu_i = \mu_k$, so that $\mu_k = \sum_{i \in J} \mu_i = 1$. From $J = \{k\}$, we also obtain $\sum_{i \in J} \mu_i x_i = \sum_{i \in \{k\}} \mu_i x_i = \underbrace{\mu_k}_{=1} x_k = x_k \in C$, qed.

[62] *Proof.* Assume the contrary. Thus, every $k \in J$ satisfies $\mu_k \geq 1$. In other words, every $i \in J$ satisfies $\mu_i \geq 1$. Now, $\sum_{i \in J} \mu_i = 1$, so that $1 = \sum_{i \in J} \underbrace{\mu_i}_{\geq 1} \geq \sum_{k \in J} 1 = |J| \cdot 1 = |J|$. In other words, $|J| \leq 1$. This contradicts the fact that we don't have $|J| \leq 1$. This contradiction proves that our assumption was wrong, qed.

1. Now,

$$\sum_{i \in J \setminus \{k\}} \underbrace{\mu'_i}_{= \frac{\mu_i}{1 - \mu_k}} = \sum_{i \in J \setminus \{k\}} \frac{\mu_i}{1 - \mu_k} = \frac{\sum\limits_{i \in J \setminus \{k\}} \mu_i}{1 - \mu_k} = 1.$$

Thus, we can apply (50) to $J \setminus \{k\}$ and $(\mu'_i)_{i \in J \setminus \{k\}}$ instead of J and $(\mu_i)_{i \in J}$. As a result, we obtain $\sum\limits_{i \in J \setminus \{k\}} \mu'_i x_i \in C$.

On the other hand,

$$\sum_{i \in J} \mu_i x_i = \underbrace{\sum_{\substack{i \in J; \\ i = k}} \mu_i}_{\substack{= \mu_k x_k \\ (\text{since } k \in J)}} + \underbrace{\sum_{\substack{i \in J; \\ i \neq k}} \mu_i x_i}_{= \sum\limits_{i \in J \setminus \{k\}}} = \mu_k x_k + \sum_{i \in J \setminus \{k\}} \mu_i x_i. \tag{51}$$

We have $k \in J \subseteq I$ and thus $x_k \in C$. Also, $\mu_k \in [0, 1]$ (since μ_k is a nonnegative real and satisfies $\mu_k < 1$).

Now, recall that the set C is convex. In other words, every two elements $x \in C$ and $y \in C$ and every real number $\lambda \in [0, 1]$ satisfy $\lambda x + (1 - \lambda) y \in C$ (because of the definition of convexity). Applying this to $x = x_k$, $y = \sum\limits_{i \in J \setminus \{k\}} \mu'_i x_i$ and $\lambda = \mu_k$ gives us $\mu_k x_k + (1 - \mu_k) \sum\limits_{i \in J \setminus \{k\}} \mu'_i x_i \in C$ (since $x_k \in C$, $\sum\limits_{i \in J \setminus \{k\}} \mu'_i x_i \in C$ and $\mu_k \in [0, 1]$). Since

$$\mu_k x_k + (1 - \mu_k) \sum_{i \in J \setminus \{k\}} \underbrace{\mu'_i}_{= \frac{\mu_i}{1 - \mu_k}} x_i$$

$$= \mu_k x_k + (1 - \mu_k) \underbrace{\sum_{i \in J \setminus \{k\}} \frac{\mu_i}{1 - \mu_k} x_i}_{= \frac{1}{1 - \mu_k} \sum\limits_{i \in J \setminus \{k\}} \mu_i x_i} = \mu_k x_k + \underbrace{(1 - \mu_k) \frac{1}{1 - \mu_k}}_{= 1} \sum_{i \in J \setminus \{k\}} \mu_i x_i$$

$$= \mu_k x_k + \sum_{i \in J \setminus \{k\}} \mu_i x_i = \sum_{i \in J} \mu_i x_i \qquad (\text{by (51)}),$$

this rewrites as $\sum\limits_{i \in J} \mu_i x_i \in C$.

Now, let us forget that we fixed J and $(\mu_i)_{i \in J}$. We thus have proven that if J is any finite subset of I satisfying $|J| = N$, and if $(\mu_i)_{i \in J}$ is a family of nonnegative reals indexed by J such that $\sum\limits_{i \in J} \mu_i = 1$, then $\sum\limits_{i \in J} \mu_i x_i \in C$. In other words, (49) holds in the case when $|J| = N$. This completes the induction step.

Thus, (49) is proven by induction.]

Now, recall that all but finitely many $i \in I$ satisfy $\lambda_i = 0$. In other words, there exists a finite subset J of I such that

$$\text{every } i \in I \setminus J \text{ satisfies } \lambda_i = 0. \tag{52}$$

Consider this J. We have $\sum\limits_{i \in I} \lambda_i = 1$. Since

$$\sum_{i \in I} \lambda_i = \underbrace{\sum_{\substack{i \in I; \\ i \in J}} \lambda_i}_{\substack{= \sum\limits_{i \in J} \\ (\text{since } J \subseteq I)}} + \underbrace{\sum_{\substack{i \in I; \\ i \notin J}} \lambda_i}_{= \sum\limits_{i \in I \setminus J}} = \sum_{i \in J} \lambda_i + \sum_{i \in I \setminus J} \underbrace{\lambda_i}_{\substack{=0 \\ (\text{by (52)})}}$$

$$= \sum_{i \in J} \lambda_i + \underbrace{\sum_{i \in I \setminus J} 0}_{=0} = \sum_{i \in J} \lambda_i,$$

this rewrites as $\sum\limits_{i \in J} \lambda_i = 1$. Thus, (49) (applied to $(\mu_i)_{i \in J} = (\lambda_i)_{i \in J}$) shows that $\sum\limits_{i \in J} \lambda_i x_i \in C$. Now,

$$\sum_{i \in I} \lambda_i x_i = \underbrace{\sum_{\substack{i \in I; \\ i \in J}} \lambda_i x_i}_{\substack{= \sum\limits_{i \in J} \\ (\text{since } J \subseteq I)}} + \underbrace{\sum_{\substack{i \in I; \\ i \notin J}} \lambda_i x_i}_{= \sum\limits_{i \in I \setminus J}} = \sum_{i \in J} \lambda_i x_i + \sum_{i \in I \setminus J} \underbrace{\lambda_i}_{\substack{=0 \\ (\text{by (52)})}}$$

$$= \sum_{i \in J} \lambda_i x_i + \underbrace{\sum_{i \in I \setminus J} 0 x_i}_{=0} = \sum_{i \in J} \lambda_i x_i \in C.$$

This proves Proposition 2.0a. $\qquad\square$

Proof of Proposition 2.0f. **(a)** Let E be an $\mathbb{R}$-vector space. Let S be a subset of E.

Let C_1 be the intersection of all convex subsets of E which contain S as a subset. Definition 2.0c defined the convex hull of S to be this set C_1. In other words,

$$(\text{the convex hull of } S \text{ defined according to Definition 2.0c}) = C_1. \tag{53}$$

Let C_2 be the set of all convex combinations of the vectors s for $s \in S$. Definition 2.0d defined the convex hull of S to be this set C_2. In other words,

$$(\text{the convex hull of } S \text{ defined according to Definition 2.0d}) = C_2. \tag{54}$$

Let C_3 be the set

$$\left\{ x \in E \;\middle|\; \left(\begin{array}{c} \text{there exist some } t \in \mathbb{N}, \\ \text{a } t\text{-tuple } (x_1, x_2, \ldots, x_t) \text{ of elements of } S \\ \text{and a } t\text{-tuple } (\lambda_1, \lambda_2, \ldots, \lambda_t) \text{ of nonnegative reals} \\ \text{such that } \sum_{i=1}^{t} \lambda_i = 1 \text{ and } \sum_{i=1}^{t} \lambda_i x_i = x \end{array} \right) \right\}.$$

Definition 2.0e defined the convex hull of S to be this set C_3. In other words,

$$(\text{the convex hull of } S \text{ defined according to Definition 2.0e}) = C_3. \tag{55}$$

The set C_3 is convex[63]. Moreover, $S \subseteq C_3$ [64]. Hence, C_3 is a convex subset

---

[63]*Proof.* Let $y \in C_3$, $z \in C_3$ and $\lambda \in [0,1]$. We shall show that $\lambda y + (1 - \lambda) z \in C_3$.

Notice that λ and $1 - \lambda$ both are nonnegative reals (since $\lambda \in [0,1]$).

We have

$$y \in C_3 = \left\{ x \in E \ \middle| \ \left(\begin{array}{c} \text{there exist some } t \in \mathbb{N}, \\ \text{a } t\text{-tuple } (x_1, x_2, \ldots, x_t) \text{ of elements of } S \\ \text{and a } t\text{-tuple } (\lambda_1, \lambda_2, \ldots, \lambda_t) \text{ of nonnegative reals} \\ \text{such that } \sum_{i=1}^{t} \lambda_i = 1 \text{ and } \sum_{i=1}^{t} \lambda_i x_i = x \end{array} \right) \right\}.$$

In other words, we can write y in the form $y = \sum_{i=1}^{t} \lambda_i x_i$ for some $t \in \mathbb{N}$, some t-tuple $(x_1, x_2, \ldots, x_t)$ of elements of S and some t-tuple $(\lambda_1, \lambda_2, \ldots, \lambda_t)$ of nonnegative reals such that $\sum_{i=1}^{t} \lambda_i = 1$. Let us denote this t, this $(x_1, x_2, \ldots, x_t)$ and this $(\lambda_1, \lambda_2, \ldots, \lambda_t)$ by p, $(y_1, y_2, \ldots, y_p)$ and $(\mu_1, \mu_2, \ldots, \mu_p)$, respectively. Thus, p is an element of $\mathbb{N}$, and $(y_1, y_2, \ldots, y_p)$ is a p-tuple of elements of S, and $(\mu_1, \mu_2, \ldots, \mu_p)$ is a p-tuple of nonnegative reals such that $\sum_{i=1}^{p} \mu_i = 1$ and $y = \sum_{i=1}^{p} \mu_i y_i$.

Similarly, use the assumption $z \in C_3$ to write z in the form $z = \sum_{i=1}^{q} \nu_i z_i$, where q is an element of $\mathbb{N}$, and where $(z_1, z_2, \ldots, z_q)$ is a q-tuple of elements of S, and where $(\nu_1, \nu_2, \ldots, \nu_q)$ is a q-tuple of nonnegative reals such that $\sum_{i=1}^{q} \nu_i = 1$.

Define a $(p+q)$-tuple $(w_1, w_2, \ldots, w_{p+q})$ of elements of S by

$$(w_1, w_2, \ldots, w_{p+q}) = (y_1, y_2, \ldots, y_p, z_1, z_2, \ldots, z_q).$$

Then:

- For every $i \in \{1, 2, \ldots, p\}$, we have

$$w_i = y_i. \tag{56}$$

- For every $i \in \{p+1, p+2, \ldots, p+q\}$, we have

$$w_i = z_{i-p}. \tag{57}$$

Furthermore, define a $(p+q)$-tuple $(\eta_1, \eta_2, \ldots, \eta_{p+q})$ of nonnegative reals by

$$(\eta_1, \eta_2, \ldots, \eta_{p+q}) = (\lambda\mu_1, \lambda\mu_2, \ldots, \lambda\mu_p, (1-\lambda)\nu_1, (1-\lambda)\nu_2, \ldots, (1-\lambda)\nu_q)$$

(this is well-defined because both λ and $1 - \lambda$ are nonnegative reals and because $(\mu_1, \mu_2, \ldots, \mu_p)$ and $(\nu_1, \nu_2, \ldots, \nu_q)$ are tuples of nonnegative reals). Then:

- For every $i \in \{1, 2, \ldots, p\}$, we have

$$\eta_i = \lambda\mu_i. \tag{58}$$

- For every $i \in \{p+1, p+2, \ldots, p+q\}$, we have

$$\eta_i = (1-\lambda)\nu_{i-p}. \tag{59}$$

Now,

$$\sum_{i=1}^{p+q} \eta_i = \sum_{i=1}^{p} \underbrace{\eta_i}_{\substack{=\lambda\mu_i \\ \text{(by (58))}}} + \sum_{i=p+1}^{p+q} \underbrace{\eta_i}_{\substack{=(1-\lambda)\nu_{i-p} \\ \text{(by (59))}}} = \underbrace{\sum_{i=1}^{p} \lambda\mu_i}_{=\lambda\sum\limits_{i=1}^{p}\mu_i} + \underbrace{\sum_{i=p+1}^{p+q} (1-\lambda)\,\nu_{i-p}}_{=(1-\lambda)\sum\limits_{i=p+1}^{p+q}\nu_{i-p}}$$

$$= \lambda\sum_{i=1}^{p}\mu_i + (1-\lambda)\sum_{i=p+1}^{p+q}\nu_{i-p} = \lambda\underbrace{\sum_{i=1}^{p}\mu_i}_{=1} + (1-\lambda)\underbrace{\sum_{i=1}^{q}\nu_i}_{=1}$$

(here, we substituted i for $i-p$ in the second sum)

$$= \lambda + (1-\lambda) = 1$$

and

$$\sum_{i=1}^{p+q} \eta_i w_i = \sum_{i=1}^{p} \underbrace{\eta_i}_{\substack{=\lambda\mu_i \\ \text{(by (58))}}}\underbrace{w_i}_{\substack{=y_i \\ \text{(by (56))}}} + \sum_{i=p+1}^{p+q} \underbrace{\eta_i}_{\substack{=(1-\lambda)\nu_{i-p} \\ \text{(by (59))}}}\underbrace{w_i}_{\substack{=z_{i-p} \\ \text{(by (57))}}}$$

$$= \underbrace{\sum_{i=1}^{p} \lambda\mu_i y_i}_{=\lambda\sum\limits_{i=1}^{p}\mu_i y_i} + \underbrace{\sum_{i=p+1}^{p+q} (1-\lambda)\,\nu_{i-p}z_{i-p}}_{=(1-\lambda)\sum\limits_{i=p+1}^{p+q}\nu_{i-p}z_{i-p}}$$

$$= \lambda\sum_{i=1}^{p}\mu_i y_i + (1-\lambda)\sum_{i=p+1}^{p+q}\nu_{i-p}z_{i-p} = \lambda\underbrace{\sum_{i=1}^{p}\mu_i y_i}_{=y} + (1-\lambda)\underbrace{\sum_{i=1}^{q}\nu_i z_i}_{=z}$$

(here, we substituted i for $i-p$ in the second sum)

$$= \lambda y + (1-\lambda)\,z.$$

Hence, there exist some $t \in \mathbb{N}$, a t-tuple $(x_1, x_2, \ldots, x_t)$ of elements of S and a t-tuple $(\lambda_1, \lambda_2, \ldots, \lambda_t)$ of nonnegative reals such that $\sum_{i=1}^{t} \lambda_i = 1$ and $\sum_{i=1}^{t} \lambda_i x_i = \lambda y + (1-\lambda)\,z$ (namely, $t = p + q$, $(x_1, x_2, \ldots, x_t) = (w_1, w_2, \ldots, w_{p+q})$ and $(\lambda_1, \lambda_2, \ldots, \lambda_t) = (\eta_1, \eta_2, \ldots, \eta_{p+q})$). In other words,

$$\lambda y + (1-\lambda)\,z \in \left\{ x \in E \;\middle|\; \left(\begin{array}{c} \text{there exist some } t \in \mathbb{N}, \\ \text{a } t\text{-tuple } (x_1, x_2, \ldots, x_t) \text{ of elements of } S \\ \text{and a } t\text{-tuple } (\lambda_1, \lambda_2, \ldots, \lambda_t) \text{ of nonnegative reals} \\ \text{such that } \sum\limits_{i=1}^{t} \lambda_i = 1 \text{ and } \sum\limits_{i=1}^{t} \lambda_i x_i = x \end{array} \right) \right\} = C_3.$$

Now, let us forget that we fixed y, z and λ. We thus have proven that every two elements $y \in C_3$ and $z \in C_3$ and every real number $\lambda \in [0,1]$ satisfy $\lambda y + (1-\lambda)\,z \in C_3$. If we rename y and z as x and y in this statement, we obtain the following: Every two elements $x \in C_3$ and $y \in C_3$ and every real number $\lambda \in [0,1]$ satisfy $\lambda x + (1-\lambda)\,y \in C_3$. In other words, the set C_3 is convex (by the definition of convexity). Qed.

of E which contains S as a subset. Thus, the intersection of all convex subsets of E which contain S as a subset is a subset of C_3. In other words, C_1 is a subset of C_3 (since C_1 is the intersection of all convex subsets of E which contain S as a subset). In other words, $C_1 \subseteq C_3$.

On the other hand, $C_3 \subseteq C_1$ [65]. Combined with $C_1 \subseteq C_3$, this shows that

---

[64] *Proof.* Let $w \in S$. Then, $\sum\limits_{i=1}^{1} 1 = 1$ and $\sum\limits_{i=1}^{1} 1w = 1w = w$. Hence, there exist some $t \in \mathbb{N}$, a t-tuple $(x_1, x_2, \ldots, x_t)$ of elements of S and a t-tuple $(\lambda_1, \lambda_2, \ldots, \lambda_t)$ of nonnegative reals such that $\sum\limits_{i=1}^{t} \lambda_i = 1$ and $\sum\limits_{i=1}^{t} \lambda_i x_i = w$ (namely, $t = 1$, $(x_1, x_2, \ldots, x_t) = (w)$ and $(\lambda_1, \lambda_2, \ldots, \lambda_t) = (1)$). In other words,

$$w \in \left\{ x \in E \ \middle| \ \begin{pmatrix} \text{there exist some } t \in \mathbb{N}, \\ \text{a } t\text{-tuple } (x_1, x_2, \ldots, x_t) \text{ of elements of } S \\ \text{and a } t\text{-tuple } (\lambda_1, \lambda_2, \ldots, \lambda_t) \text{ of nonnegative reals} \\ \text{such that } \sum\limits_{i=1}^{t} \lambda_i = 1 \text{ and } \sum\limits_{i=1}^{t} \lambda_i x_i = x \end{pmatrix} \right\} = C_3.$$

Let us now forget that we fixed w. We thus have proven that $w \in C_3$ for every $w \in S$. In other words, $S \subseteq C_3$, qed.

[65] *Proof.* Let $w \in C_3$. We shall show that $w \in C_1$.

Let D be any convex subset of E which contains S as a subset. Thus, D is a convex subset of E and satisfies $S \subseteq D$.

We have

$$w \in C_3 = \left\{ x \in E \ \middle| \ \begin{pmatrix} \text{there exist some } t \in \mathbb{N}, \\ \text{a } t\text{-tuple } (x_1, x_2, \ldots, x_t) \text{ of elements of } S \\ \text{and a } t\text{-tuple } (\lambda_1, \lambda_2, \ldots, \lambda_t) \text{ of nonnegative reals} \\ \text{such that } \sum\limits_{i=1}^{t} \lambda_i = 1 \text{ and } \sum\limits_{i=1}^{t} \lambda_i x_i = x \end{pmatrix} \right\}.$$

In other words, there exist some $t \in \mathbb{N}$, a t-tuple $(x_1, x_2, \ldots, x_t)$ of elements of S and a t-tuple $(\lambda_1, \lambda_2, \ldots, \lambda_t)$ of nonnegative reals such that $\sum\limits_{i=1}^{t} \lambda_i = 1$ and $\sum\limits_{i=1}^{t} \lambda_i x_i = w$. Consider this t, this $(x_1, x_2, \ldots, x_t)$ and this $(\lambda_1, \lambda_2, \ldots, \lambda_t)$. For every $i \in \{1, 2, \ldots, t\}$, we have $x_i \in S \subseteq D$. Thus, $(x_i)_{i \in \{1,2,\ldots,t\}}$ is a family of elements of D. Also, clearly, $(\lambda_i)_{i \in \{1,2,\ldots,t\}}$ is a family of nonnegative reals (since $(\lambda_1, \lambda_2, \ldots, \lambda_t)$ is a t-tuple of nonnegative reals). Furthermore, all but finitely many $i \in \{1, 2, \ldots, t\}$ satisfy $\lambda_i = 0$ (since there are only finitely many $i \in \{1, 2, \ldots, t\}$). Finally, $\sum\limits_{i \in \{1,2,\ldots,t\}} \lambda_i = \sum\limits_{i=1}^{t} \lambda_i = 1$. Hence, Proposition 2.0a (applied to D and $\{1, 2, \ldots, t\}$ instead of C and I) shows that $\sum\limits_{i \in \{1,2,\ldots,t\}} \lambda_i x_i \in D$. Thus,

$$w = \sum\limits_{i=1}^{t} \lambda_i x_i = \sum\limits_{i \in \{1,2,\ldots,t\}} \lambda_i x_i \in D.$$

Let us now forget that we fixed D. We thus have proven that $w \in D$ whenever D is any convex subset of E which contains S as a subset. In other words, w lies in every convex subset of E which contains S as a subset. Hence, w lies in the intersection of all convex subsets of E which contain S as a subset. In other words, w lies in C_1 (since C_1 is the intersection of all convex subsets of E which contain S as a subset). In other words, $w \in C_1$.

Let us now forget that we fixed w. We thus have proven that $w \in C_1$ for every $w \in C_3$.

$C_1 = C_3$.

In other words, $C_3 \subseteq C_1$, qed.

$C_1 = C_3$.

The set C_2 is convex[66]. Moreover, $S \subseteq C_2$ [67]. Hence, C_2 is a convex subset of E which contains S as a subset. Thus, the intersection of all convex subsets of

---

[66]*Proof.* Let $y \in C_2$, $z \in C_2$ and $\lambda \in [0, 1]$. We shall show that $\lambda y + (1 - \lambda) z \in C_2$.

Notice that λ and $1 - \lambda$ both are nonnegative reals (since $\lambda \in [0, 1]$).

We have $y \in C_2$. In other words, y is a convex combination of the vectors s for $s \in S$ (since C_2 is the set of all convex combinations of the vectors s for $s \in S$). In other words, y has the form $y = \sum_{i \in S} \mu_i i$ for some family $(\mu_i)_{i \in S}$ of nonnegative reals indexed by elements of S and having the property that all but finitely many $i \in S$ satisfy $\mu_i = 0$, and the property that $\sum_{i \in S} \mu_i = 1$. Similarly, z has the form $z = \sum_{i \in S} \nu_i i$ for some family $(\nu_i)_{i \in S}$ of nonnegative reals indexed by elements of S and having the property that all but finitely many $i \in S$ satisfy $\nu_i = 0$, and the property that $\sum_{i \in S} \nu_i = 1$. Consider these two families $(\mu_i)_{i \in S}$ and $(\nu_i)_{i \in S}$.

For every $i \in S$, both $\lambda \mu_i$ and $(1 - \lambda) \nu_i$ are nonnegative reals (since λ, μ_i, $1 - \lambda$ and ν_i are nonnegative reals), and hence the sum $\lambda \mu_i + (1 - \lambda) \nu_i$ is a nonnegative real. Thus, we can define a family $(\eta_i)_{i \in S}$ of nonnegative reals by setting

$$(\eta_i = \lambda \mu_i + (1 - \lambda) \nu_i \qquad \text{for every } i \in S).$$

Consider this family $(\eta_i)_{i \in S}$. We know that all but finitely many $i \in S$ satisfy $\mu_i = 0$, and we also know that all but finitely many $i \in S$ satisfy $\nu_i = 0$. Using these two facts, we see that all but finitely many $i \in S$ satisfy $\eta_i = \lambda \underbrace{\mu_i}_{=0} + (1 - \lambda) \underbrace{\nu_i}_{=0} = 0 + 0 = 0$. We have

$$\sum_{i \in S} \underbrace{\eta_i}_{=\lambda \mu_i + (1-\lambda)\nu_i} = \sum_{i \in S} (\lambda \mu_i + (1 - \lambda) \nu_i) = \lambda \underbrace{\sum_{i \in S} \mu_i}_{=1} + (1 - \lambda) \underbrace{\sum_{i \in S} \nu_i}_{=1} = \lambda + (1 - \lambda) = 1$$

and

$$\sum_{i \in S} \underbrace{\eta_i}_{=\lambda \mu_i + (1-\lambda)\nu_i} i = \sum_{i \in S} (\lambda \mu_i + (1 - \lambda) \nu_i) i = \lambda \underbrace{\sum_{i \in S} \mu_i i}_{=y} + (1 - \lambda) \underbrace{\sum_{i \in S} \nu_i z}_{=z} = \lambda y + (1 - \lambda) z.$$

Hence, $\lambda y + (1 - \lambda) z = \sum_{i \in S} \eta_i i$. Therefore, $\lambda y + (1 - \lambda) z$ has the form $\sum_{i \in S} \lambda_i i$ for some family $(\lambda_i)_{i \in S}$ of nonnegative reals indexed by elements of S and having the property that all but finitely many $i \in S$ satisfy $\lambda_i = 0$, and the property that $\sum_{i \in S} \lambda_i = 1$ (namely, for the family $(\lambda_i)_{i \in S} = (\eta_i)_{i \in S}$). In other words, $\lambda y + (1 - \lambda) z$ is a convex combination of the vectors s for $s \in S$. In other words, $\lambda y + (1 - \lambda) z \in C_2$ (since C_2 is the set of all convex combinations of the vectors s for $s \in S$).

Now, let us forget that we fixed y, z and λ. We thus have proven that every two elements $y \in C_2$ and $z \in C_2$ and every real number $\lambda \in [0, 1]$ satisfy $\lambda y + (1 - \lambda) z \in C_2$. If we rename y and z as x and y in this statement, we obtain the following: Every two elements $x \in C_2$ and $y \in C_2$ and every real number $\lambda \in [0, 1]$ satisfy $\lambda x + (1 - \lambda) y \in C_2$. In other words, the set C_2 is convex (by the definition of convexity).

[67]*Proof.* Let $w \in S$. Define a family $(\mu_i)_{i \in S}$ of nonnegative reals by setting

$$\left(\mu_i = \begin{cases} 1, & \text{if } i = w; \\ 0, & \text{if } i \neq w \end{cases} \qquad \text{for every } i \in S \right).$$

E which contain S as a subset is a subset of C_2. In other words, C_1 is a subset of C_2 (since C_1 is the intersection of all convex subsets of E which contain S as a subset). In other words, $C_1 \subseteq C_2$.

On the other hand, $C_2 \subseteq C_1$ [68]. Combined with $C_1 \subseteq C_2$, this shows that $C_1 = C_2$.

Combining $C_1 = C_2$ with $C_1 = C_3$, we obtain $C_1 = C_2 = C_3$. Now, (53) becomes

(the convex hull of S defined according to Definition 2.0c)

$= C_1 = C_2 = $ (the convex hull of S defined according to Definition 2.0d)   (60)

Then, it is easy to see that:

- All but finitely many $i \in S$ satisfy $\mu_i = 0$.

- We have $\sum\limits_{i \in S} \mu_i = 1$.

- We have $\sum\limits_{i \in S} \mu_i i = w$.

Hence, $w = \sum\limits_{i \in S} \mu_i i$. Thus, w has the form $\sum\limits_{i \in S} \lambda_i i$ for some family $(\lambda_i)_{i \in S}$ of nonnegative reals indexed by elements of S and having the property that all but finitely many $i \in S$ satisfy $\lambda_i = 0$, and the property that $\sum\limits_{i \in S} \lambda_i = 1$ (namely, for the family $(\lambda_i)_{i \in S} = (\mu_i)_{i \in S}$).

In other words, w is a convex combination of the vectors s for $s \in S$. In other words, $w \in C_2$ (since C_2 is the set of all convex combinations of the vectors s for $s \in S$).

Now, let us forget that we fixed w. We thus have proven that $w \in C_2$ for every $w \in S$. In other words, $S \subseteq C_2$, qed.

[68] *Proof.* Let $w \in C_2$. We shall show that $w \in C_1$.

Let D be any convex subset of E which contains S as a subset. Thus, D is a convex subset of E and satisfies $S \subseteq D$.

We have $w \in C_2$. In other words, w is a convex combination of the vectors s for $s \in S$ (since C_2 is the set of all convex combinations of the vectors s for $s \in S$). In other words, w has the form $w = \sum\limits_{i \in S} \lambda_i i$ for some family $(\lambda_i)_{i \in S}$ of nonnegative reals indexed by elements of S and having the property that all but finitely many $i \in S$ satisfy $\lambda_i = 0$, and the property that $\sum\limits_{i \in S} \lambda_i = 1$. Consider this family $(\lambda_i)_{i \in S}$.

For every $i \in S$, we have $i \in S \subseteq D$. Thus, $(i)_{i \in S}$ is a family of elements of D. Hence, Proposition 2.0a (applied to D, S, $(i)_{i \in S}$ and $(\lambda_i)_{i \in S}$ instead of C, I, $(x_i)_{i \in I}$ and $(\lambda_i)_{i \in I}$) shows that $\sum\limits_{i \in S} \lambda_i i \in D$. Thus, $w = \sum\limits_{i \in S} \lambda_i i \in D$.

Let us now forget that we fixed D. We thus have proven that $w \in D$ whenever D is any convex subset of E which contains S as a subset. In other words, w lies in every convex subset of E which contains S as a subset. Hence, w lies in the intersection of all convex subsets of E which contain S as a subset. In other words, w lies in C_1 (since C_1 is the intersection of all convex subsets of E which contain S as a subset). In other words, $w \in C_1$.

Let us now forget that we fixed w. We thus have proven that $w \in C_1$ for every $w \in C_2$. In other words, $C_2 \subseteq C_1$, qed.

(by (54)). Also, (54) becomes

(the convex hull of S defined according to Definition 2.0d)

$= C_2 = C_3 = $ (the convex hull of S defined according to Definition 2.0e) (61)

(by (55)).

Now, let us forget that we fixed S. We thus have shown that if S is any subset of E, then (60) and (61) hold. In other words, if S is any subset of E, then

(the convex hull of S defined according to Definition 2.0c)

$=$ (the convex hull of S defined according to Definition 2.0d)

$=$ (the convex hull of S defined according to Definition 2.0e).

In other words, Definitions 2.0c, 2.0d and 2.0e are equivalent. This proves Proposition 2.0f **(a)**.

(c) We define the set C_3 as in our proof of Proposition 2.0f **(a)**. Then, (55) shows that

$C_3 = $ (the convex hull of S defined according to Definition 2.0e)

$=$ (the convex hull of S) $=$ conv . hull S.

But in our proof of Proposition 2.0f **(a)**, we have shown that $S \subseteq C_3$. Thus, $S \subseteq C_3 = $ conv . hull S. This proves Proposition 2.0f **(c)**.

(d) We define the set C_3 as in our proof of Proposition 2.0f **(a)**. Then, (55) shows that

$C_3 = $ (the convex hull of S defined according to Definition 2.0e)

$=$ (the convex hull of S) $=$ conv . hull S.

But in our proof of Proposition 2.0f **(a)**, we have shown that the set C_3 is convex. Since $C_3 = $ conv . hull S, this rewrites as follows: The set conv . hull S is convex. This proves Proposition 2.0f **(d)**.

(f) Let D be a convex subset of E which contains S as a subset. Thus, the intersection of all convex subsets of E which contain S as a subset is a subset of D. In other words, conv . hull S is a subset of D (since conv . hull S is the intersection of all convex subsets of E which contain S as a subset (by Definition 2.0d)). In other words, D contains conv . hull S as a subset.

Now, let us forget that we fixed D. We thus have shown that if D is a convex subset of E which contains S as a subset, then D contains conv . hull S as a subset. In other words, every convex subset of E which contains S as a subset also contains conv . hull S as a subset. This proves Proposition 2.0f **(f)**.

(b) Every two elements $x \in \varnothing$ and $y \in \varnothing$ and every real number $\lambda \in [0, 1]$ satisfy $\lambda x + (1 - \lambda) y \in \varnothing$ [69]. In other words, the set $\varnothing$ is convex (by the

---

[69]Indeed, this is vacuously true (since there exist no $x \in \varnothing$).

definition of convexity). Thus, $\varnothing$ is a convex subset of E which contains $\varnothing$ as a subset. Now, Proposition 2.0f **(f)** (applied to $S = \varnothing$) shows that every convex subset of E which contains $\varnothing$ as a subset also contains $\operatorname{conv.hull}\varnothing$ as a subset. Thus, the set $\varnothing$ contains $\operatorname{conv.hull}\varnothing$ as a subset (since $\varnothing$ is a convex subset of E which contains $\varnothing$ as a subset). In other words, $\operatorname{conv.hull}\varnothing \subseteq \varnothing$, so that $\operatorname{conv.hull}\varnothing = \varnothing$. This proves Proposition 2.0f **(b)**.

(g) Let T be a subset of $\operatorname{conv.hull}S$. The set $\operatorname{conv.hull}S$ is a convex set (by Proposition 2.0f **(d)**) and contains T as a subset (since $T \subseteq \operatorname{conv.hull}S$). In other words, $\operatorname{conv.hull}S$ is a convex subset of E which contains T as a subset.

But Proposition 2.0f **(f)** (applied to T instead of S) shows that every convex subset of E which contains T as a subset also contains $\operatorname{conv.hull}T$ as a subset. Thus, the set $\operatorname{conv.hull}S$ contains $\operatorname{conv.hull}T$ as a subset (since $\operatorname{conv.hull}S$ is a convex subset of E which contains T as a subset). In other words, $\operatorname{conv.hull}T \subseteq \operatorname{conv.hull}S$. This proves Proposition 2.0f **(g)**.

(e) Let T be a subset of S. Thus, $T \subseteq S \subseteq \operatorname{conv.hull}S$ (by Proposition 2.0f **(c)**). In other words, T is a subset of $\operatorname{conv.hull}S$. Hence, Proposition 2.0f **(g)** yields $\operatorname{conv.hull}T \subseteq \operatorname{conv.hull}S$. This proves Proposition 2.0f **(e)**. $\square$

Proof of Proposition 2.0g. **(a)** Let S be a subset of E. Definition 2.0e yields

$\operatorname{conv.hull}S$

$$= \left\{ x \in E \;\middle|\; \left(\begin{array}{c} \text{there exist some } t \in \mathbb{N}, \\ \text{a } t\text{-tuple } (x_1, x_2, \ldots, x_t) \text{ of elements of } S \\ \text{and a } t\text{-tuple } (\lambda_1, \lambda_2, \ldots, \lambda_t) \text{ of nonnegative reals} \\ \text{such that } \sum_{i=1}^{t} \lambda_i = 1 \text{ and } \sum_{i=1}^{t} \lambda_i x_i = x \end{array} \right) \right\}$$

$$= \left\{ x \in E \;\middle|\; \left(\begin{array}{c} x \text{ can be written in the form } x = \sum_{i=1}^{t} \lambda_i x_i, \\ \text{where } t \text{ is an element of } \mathbb{N}, \\ \text{and } (x_1, x_2, \ldots, x_t) \text{ is a } t\text{-tuple of elements of } S, \\ \text{and } (\lambda_1, \lambda_2, \ldots, \lambda_t) \text{ is a } t\text{-tuple of nonnegative reals} \\ \text{such that } \sum_{i=1}^{t} \lambda_i = 1 \end{array} \right) \right\}$$

$$= \left\{ \sum_{i=1}^{t} \lambda_i x_i \;\middle|\; \left(\begin{array}{c} t \text{ is an element of } \mathbb{N}, \\ \text{and } (x_1, x_2, \ldots, x_t) \text{ is a } t\text{-tuple of elements of } S, \\ \text{and } (\lambda_1, \lambda_2, \ldots, \lambda_t) \text{ is a } t\text{-tuple of nonnegative reals} \\ \text{such that } \sum_{i=1}^{t} \lambda_i = 1 \end{array} \right) \right\}.$$

This proves Proposition 2.0g **(a)**.

(b) Proposition 2.0g **(a)** provides an expression for $\operatorname{conv.hull}S$ which clearly does not depend on whether we consider S as a subset of F or as a subset of E (since it only refers to the elements of S). Hence, $\operatorname{conv.hull}S$ does not depend

on whether we consider S as a subset of F or as a subset of E. This proves Proposition 2.0g **(b)**. □

Proof of Proposition 2.0h. **(a)** Let $S = \{x_1, x_2, \ldots, x_n\}$. Let Q be the set of all convex combinations of the vectors x_1, x_2, …, x_n.

Recall that Q is the set of all convex combinations of the vectors x_1, x_2, …, x_n. In other words, Q is the set of all convex combinations of the vectors x_i for $i \in \{1, 2, \ldots, n\}$.

The set Q is convex[70]. Moreover, $S \subseteq Q$ [71]. Hence, Q is a convex subset

---

[70] *Proof.* Let $y \in Q$, $z \in Q$ and $\lambda \in [0,1]$. We shall show that $\lambda y + (1-\lambda) z \in Q$.

Notice that λ and $1-\lambda$ both are nonnegative reals (since $\lambda \in [0,1]$).

We have $y \in Q$. In other words, y is a convex combination of the vectors x_i for $i \in \{1, 2, \ldots, n\}$ (since Q is the set of all convex combinations of the vectors x_i for $i \in \{1, 2, \ldots, n\}$). In other words, y has the form $y = \sum_{i \in \{1,2,\ldots,n\}} \mu_i x_i$ for some family $(\mu_i)_{i \in \{1,2,\ldots,n\}}$ of nonnegative reals indexed by elements of $\{1, 2, \ldots, n\}$ and having the property that all but finitely many $i \in \{1, 2, \ldots, n\}$ satisfy $\mu_i = 0$, and the property that $\sum_{i \in \{1,2,\ldots,n\}} \mu_i = 1$. Similarly, z has the form $z = \sum_{i \in \{1,2,\ldots,n\}} \nu_i x_i$ for some family $(\nu_i)_{i \in \{1,2,\ldots,n\}}$ of nonnegative reals indexed by elements of $\{1, 2, \ldots, n\}$ and having the property that all but finitely many $i \in \{1, 2, \ldots, n\}$ satisfy $\nu_i = 0$, and the property that $\sum_{i \in \{1,2,\ldots,n\}} \nu_i = 1$. Consider these two families $(\mu_i)_{i \in \{1,2,\ldots,n\}}$ and $(\nu_i)_{i \in \{1,2,\ldots,n\}}$.

For every $i \in \{1, 2, \ldots, n\}$, both $\lambda \mu_i$ and $(1-\lambda) \nu_i$ are nonnegative reals (since λ, μ_i, $1-\lambda$ and ν_i are nonnegative reals), and hence the sum $\lambda \mu_i + (1-\lambda) \nu_i$ is a nonnegative real. Thus, we can define a family $(\eta_i)_{i \in \{1,2,\ldots,n\}}$ of nonnegative reals by setting

$$\left(\eta_i = \lambda \mu_i + (1-\lambda) \nu_i \qquad \text{for every } i \in \{1, 2, \ldots, n\} \right).$$

Consider this family $(\eta_i)_{i \in \{1,2,\ldots,n\}}$. We know that all but finitely many $i \in \{1, 2, \ldots, n\}$ satisfy $\eta_i = 0$ (since there are only finitely many $i \in \{1, 2, \ldots, n\}$). We have

$$\sum_{i \in \{1,2,\ldots,n\}} \underbrace{\eta_i}_{= \lambda \mu_i + (1-\lambda)\nu_i} = \sum_{i \in \{1,2,\ldots,n\}} (\lambda \mu_i + (1-\lambda) \nu_i)$$
$$= \lambda \underbrace{\sum_{i \in \{1,2,\ldots,n\}} \mu_i}_{=1} + (1-\lambda) \underbrace{\sum_{i \in \{1,2,\ldots,n\}} \nu_i}_{=1} = \lambda + (1-\lambda) = 1$$

and

$$\sum_{i \in \{1,2,\ldots,n\}} \underbrace{\eta_i}_{= \lambda \mu_i + (1-\lambda)\nu_i} x_i = \sum_{i \in \{1,2,\ldots,n\}} (\lambda \mu_i + (1-\lambda) \nu_i) x_i$$
$$= \lambda \underbrace{\sum_{i \in \{1,2,\ldots,n\}} \mu_i x_i}_{=y} + (1-\lambda) \underbrace{\sum_{i \in \{1,2,\ldots,n\}} \nu_i x_i}_{=z} = \lambda y + (1-\lambda) z.$$

Hence, $\lambda y + (1-\lambda) z = \sum_{i \in \{1,2,\ldots,n\}} \eta_i x_i$. Therefore, $\lambda y + (1-\lambda) z$ has the form $\sum_{i \in \{1,2,\ldots,n\}} \lambda_i x_i$ for some family $(\lambda_i)_{i \in \{1,2,\ldots,n\}}$ of nonnegative reals indexed by elements of $\{1, 2, \ldots, n\}$ and having the property that all but finitely many $i \in \{1, 2, \ldots, n\}$ satisfy $\lambda_i = 0$, and the property that $\sum_{i \in \{1,2,\ldots,n\}} \lambda_i = 1$ (namely, for the family $(\lambda_i)_{i \in \{1,2,\ldots,n\}} = (\eta_i)_{i \in \{1,2,\ldots,n\}}$). In other words, $\lambda y + (1-\lambda) z$ is a convex combination of the vectors x_i or $i \in \{1, 2, \ldots, n\}$. In other words, $\lambda y + (1-\lambda) z \in Q$ (since Q is the set of all convex combinations of the vectors x_i for $i \in \{1, 2, \ldots, n\}$).

Now, let us forget that we fixed y, z and λ. We thus have proven that every two elements $y \in Q$ and $z \in Q$ and every real number $\lambda \in [0,1]$ satisfy $\lambda y + (1-\lambda) z \in Q$. If we rename

of E which contains S as a subset. Thus, Proposition 2.0f **(f)** (applied to the convex subset Q of E) shows that Q contains conv.hull S as a subset. In other words, conv.hull $S \subseteq Q$.

On the other hand, $Q \subseteq$ conv.hull S [72]. Combining this with conv.hull $S \subseteq$

---

y and z as x and y in this statement, we obtain the following: Every two elements $x \in Q$ and $y \in Q$ and every real number $\lambda \in [0,1]$ satisfy $\lambda x + (1-\lambda) y \in Q$. In other words, the set Q is convex (by the definition of convexity). Qed.

[71] *Proof.* Let $w \in S$. Then, $w \in S = \{x_1, x_2, \ldots, x_n\}$. Thus, there exists some $k \in \{1, 2, \ldots, n\}$ such that $w = x_k$. Consider this k.

Define a family $(\mu_i)_{i \in \{1,2,\ldots,n\}}$ of nonnegative reals by setting

$$\left(\mu_i = \begin{cases} 1, & \text{if } i = k; \\ 0, & \text{if } i \neq k \end{cases} \qquad \text{for every } i \in \{1, 2, \ldots, n\} \right).$$

Then, it is easy to see that:

- All but finitely many $i \in \{1, 2, \ldots, n\}$ satisfy $\mu_i = 0$.

- We have $\sum\limits_{i \in \{1,2,\ldots,n\}} \mu_i = 1$.

- We have $\sum\limits_{i \in \{1,2,\ldots,n\}} \mu_i x_i = x_k$.

Hence, $w = x_k = \sum\limits_{i \in \{1,2,\ldots,n\}} \mu_i x_i$. Thus, w has the form $\sum\limits_{i \in \{1,2,\ldots,n\}} \lambda_i x_i$ for some family $(\lambda_i)_{i \in \{1,2,\ldots,n\}}$ of nonnegative reals indexed by elements of $\{1, 2, \ldots, n\}$ and having the property that all but finitely many $i \in \{1, 2, \ldots, n\}$ satisfy $\lambda_i = 0$, and the property that $\sum\limits_{i \in \{1,2,\ldots,n\}} \lambda_i = 1$ (namely, for the family $(\lambda_i)_{i \in \{1,2,\ldots,n\}} = (\mu_i)_{i \in \{1,2,\ldots,n\}}$). In other words, w is a convex combination of the vectors x_i for $i \in \{1, 2, \ldots, n\}$. In other words, $w \in Q$ (since Q is the set of all convex combinations of the vectors x_i for $i \in \{1, 2, \ldots, n\}$).

Now, let us forget that we fixed w. We thus have proven that $w \in Q$ for every $w \in S$. In other words, $S \subseteq Q$, qed.

[72] *Proof.* Let $w \in Q$. We shall show that $w \in$ conv.hull S.

Let D be any convex subset of E which contains S as a subset. Thus, D is a convex subset of E and satisfies $S \subseteq D$.

We have $w \in Q$. In other words, w is a convex combination of the vectors x_i for $i \in \{1, 2, \ldots, n\}$ (since Q is the set of all convex combinations of the vectors x_i for $i \in \{1, 2, \ldots, n\}$). In other words, w has the form $w = \sum\limits_{i \in \{1,2,\ldots,n\}} \lambda_i x_i$ for some family $(\lambda_i)_{i \in \{1,2,\ldots,n\}}$ of nonnegative reals indexed by elements of $\{1, 2, \ldots, n\}$ and having the property that all but finitely many $i \in \{1, 2, \ldots, n\}$ satisfy $\lambda_i = 0$, and the property that $\sum\limits_{i \in \{1,2,\ldots,n\}} \lambda_i = 1$. Consider this family $(\lambda_i)_{i \in \{1,2,\ldots,n\}}$.

For every $i \in \{1, 2, \ldots, n\}$, we have $x_i \in \{x_1, x_2, \ldots, x_n\} = S \subseteq D$. Thus, $(x_i)_{i \in \{1,2,\ldots,n\}}$ is a family of elements of D. Hence, Proposition 2.0a (applied to D, $\{1, 2, \ldots, n\}$, $(x_i)_{i \in \{1,2,\ldots,n\}}$ and $(\lambda_i)_{i \in \{1,2,\ldots,n\}}$ instead of C, I, $(x_i)_{i \in I}$ and $(\lambda_i)_{i \in I}$) shows that $\sum\limits_{i \in \{1,2,\ldots,n\}} \lambda_i x_i \in D$. Thus, $w = \sum\limits_{i \in \{1,2,\ldots,n\}} \lambda_i x_i \in D$.

Let us now forget that we fixed D. We thus have proven that $w \in D$ whenever D is any convex subset of E which contains S as a subset. In other words, w lies in every convex subset of E which contains S as a subset. Hence, w lies in the intersection of all

---

Q, we obtain conv . hull $S = Q$. Thus,

$$\text{conv . hull } \underbrace{\{x_1, x_2, \ldots, x_n\}}_{=S}$$
$$= \text{conv . hull } S = Q$$
$$= (\text{the set of all convex combinations of the vectors } x_1, x_2, \ldots, x_n)$$

(by the definition of Q). This proves Proposition 2.0h **(a)**.

(b) We have the following chain of logical equivalences:

$$\left(x \in \underbrace{\text{conv . hull } \{x_1, x_2, \ldots, x_n\}}_{\substack{=(\text{the set of all convex combinations of the vectors } x_1, x_2, \ldots, x_n) \\ (\text{by Proposition 2.0h } (\mathbf{a}))}} \right)$$

$$\Longleftrightarrow (x \in (\text{the set of all convex combinations of the vectors } x_1, x_2, \ldots, x_n))$$

$$\Longleftrightarrow (x \text{ is a convex combination of the vectors } x_1, x_2, \ldots, x_n).$$

In other words, we have $x \in \text{conv . hull } \{x_1, x_2, \ldots, x_n\}$ if and only if x is a convex combination of the vectors $x_1, x_2, \ldots, x_n$. This proves Proposition 2.0h **(b)**. $\square$

16.2. Proofs for Section 2

Proof of Proposition 2.0i. We will first show that

$$\left(\begin{array}{c} \text{if } J \text{ is any finite subset of } I, \text{ and if } (\mu_i)_{i \in J} \text{ is a} \\ \text{family of nonnegative reals indexed by } J, \\ \text{then } \sum_{i \in J} \mu_i x_i \in C \end{array} \right). \tag{62}$$

[*Proof of (62):* We shall prove (62) by strong induction over $|J|$:

Induction step:[73] Fix $N \in \mathbb{N}$. Assume that (62) holds in the case when $|J| < N$. We must prove that (62) holds in the case when $|J| = N$.

We have assumed that (62) holds in the case when $|J| < N$. In other words,

$$\left(\begin{array}{c} \text{if } J \text{ is any finite subset of } I \text{ satisfying } |J| < N, \text{ and} \\ \text{if } (\mu_i)_{i \in J} \text{ is a family of nonnegative reals indexed} \\ \text{by } J, \text{ then } \sum_{i \in J} \mu_i x_i \in C \end{array} \right). \tag{63}$$

---

convex subsets of E which contain S as a subset. In other words, w lies in conv . hull S (since conv . hull S is the intersection of all convex subsets of E which contain S as a subset (because of Definition 2.0c)). In other words, $w \in \text{conv . hull } S$.

Let us now forget that we fixed w. We thus have proven that $w \in \text{conv . hull } S$ for every $w \in Q$. In other words, $Q \subseteq \text{conv . hull } S$, qed.

[73]A strong induction needs no induction base.

Now, let us prove that (62) holds in the case when $|J| = N$. Thus, let J be any finite subset of I satisfying $|J| = N$, and let $(\mu_i)_{i \in J}$ be a family of nonnegative reals indexed by J. We shall show that $\sum_{i \in J} \mu_i x_i \in C$.

If $|J| \leq 0$, then $\sum_{i \in J} \mu_i x_i \in C$ holds[74]. Hence, for the rest of our proof of $\sum_{i \in J} \mu_i x_i \in C$, we can WLOG assume that we don't have $|J| \leq 0$. Assume this.

We have $|J| > 0$ (since we don't have $|J| \leq 0$). Thus, $J \neq \varnothing$. Hence, there exists a $k \in J$. Consider such a k.

Now, $(\mu_i)_{i \in J \setminus \{k\}}$ is a family of nonnegative reals indexed by $J \setminus \{k\}$. Moreover, $k \in J$ and thus $|J \setminus \{k\}| = \underbrace{|J|}_{=N} - 1 = N - 1 < N$. Thus, we can apply (63) to $J \setminus \{k\}$ and $(\mu_i)_{i \in J \setminus \{k\}}$ instead of J and $(\mu_i)_{i \in J}$. As a result, we obtain $\sum_{i \in J \setminus \{k\}} \mu_i x_i \in C$. Also, $x_k \in C$ (since $(x_i)_{i \in I}$ is a family of elements of C).

Recall that C is a convex cone. Thus, every two elements $x \in C$ and $y \in C$ and every nonnegative reals λ and μ satisfy $\lambda x + \mu y \in C$ (by the definition of a convex cone). Applying this to $x = x_k$, $y = \sum_{i \in J \setminus \{k\}} \mu_i x_i$, $\lambda = \mu_k$ and $\mu = 1$, we obtain $\mu_k x_k + 1 \sum_{i \in J \setminus \{k\}} \mu_i x_i \in C$. Now,

$$\sum_{i \in J} \mu_i x_i = \underbrace{\sum_{\substack{i \in J; \\ i = k}} \mu_i}_{\substack{= \mu_k x_k \\ (\text{since } k \in J)}} + \underbrace{\sum_{\substack{i \in J; \\ i \neq k}} \mu_i x_i}_{= \sum_{i \in J \setminus \{k\}}} = \mu_k x_k + \underbrace{\sum_{i \in J \setminus \{k\}} \mu_i x_i}_{= 1 \sum_{i \in J \setminus \{k\}} \mu_i x_i}$$
$$= \mu_k x_k + 1 \sum_{i \in J \setminus \{k\}} \mu_i x_i \in C.$$

Now, let us forget that we fixed J and $(\mu_i)_{i \in J}$. We thus have proven that if J is any finite subset of I satisfying $|J| = N$, and if $(\mu_i)_{i \in J}$ is a family of nonnegative reals indexed by J, then $\sum_{i \in J} \mu_i x_i \in C$. In other words, (62) holds in the case when $|J| = N$. This completes the induction step.

Thus, (62) is proven by induction.]

Now, recall that all but finitely many $i \in I$ satisfy $\lambda_i = 0$. In other words, there exists a finite subset J of I such that

$$\text{every } i \in I \setminus J \text{ satisfies } \lambda_i = 0. \tag{64}$$

---

[74]*Proof.* Assume that $|J| \leq 0$. We need to prove $\sum_{i \in J} \mu_i x_i \in C$.

From $|J| \leq 0$, we obtain $|J| = 0$. Thus, $J = \varnothing$.

Recall that C is a convex cone. Hence, $0 \in C$ (according to the definition of a convex cone).

But $J = \varnothing$, and thus $\sum_{i \in J} \mu_i x_i = (\text{empty sum}) = 0 \in C$, qed.

Consider this J. Then, (62) (applied to $(\mu_i)_{i \in J} = (\lambda_i)_{i \in J}$) shows that $\sum\limits_{i \in J} \lambda_i x_i \in C$. Now,

$$\sum_{i \in I} \lambda_i x_i = \underbrace{\sum_{\substack{i \in I; \\ i \in J}} \lambda_i x_i}_{\substack{=\sum\limits_{i \in J} \\ (\text{since } J \subseteq I)}} + \underbrace{\sum_{\substack{i \in I; \\ i \notin J}} \lambda_i x_i}_{=\sum\limits_{i \in I \setminus J}} = \sum_{i \in J} \lambda_i x_i + \sum_{i \in I \setminus J} \underbrace{\lambda_i}_{\substack{=0 \\ (\text{by } (64))}}$$

$$= \sum_{i \in J} \lambda_i x_i + \underbrace{\sum_{i \in I \setminus J} 0 x_i}_{=0} = \sum_{i \in J} \lambda_i x_i \in C.$$

This proves Proposition 2.0i. $\qquad\square$

Proof of Proposition 2.0m. **(a)** Let E be an $\mathbb{R}$-vector space. Let S be a subset of E.

Let C_1 be the intersection of all convex cones in E which contain S as a subset. Definition 2.0j defined the convex conic hull of S to be this set C_1. In other words,

$$(\text{the convex conic hull of } S \text{ defined according to Definition 2.0j}) = C_1. \quad (65)$$

Let C_2 be the set of all linear combinations of the vectors s for $s \in S$ with nonnegative coefficients. Definition 2.0k defined the convex conic hull of S to be this set C_2. In other words,

$$(\text{the convex conic hull of } S \text{ defined according to Definition 2.0k}) = C_2. \quad (66)$$

Let C_3 be the set

$$\left\{ x \in E \;\middle|\; \left(\begin{array}{c} \text{there exist some } t \in \mathbb{N}, \\ \text{a } t\text{-tuple } (x_1, x_2, \ldots, x_t) \text{ of elements of } S \\ \text{and a } t\text{-tuple } (\lambda_1, \lambda_2, \ldots, \lambda_t) \text{ of nonnegative reals} \\ \text{such that } \sum\limits_{i=1}^{t} \lambda_i x_i = x \end{array} \right) \right\}.$$

Definition 2.0l defined the convex conic hull of S to be this set C_3. In other words,

$$(\text{the convex conic hull of } S \text{ defined according to Definition 2.0l}) = C_3. \quad (67)$$

The set C_3 is a convex cone[75]. Moreover, $S \subseteq C_3$ [76]. Hence, C_3 is a convex

---

[75]*Proof.* As usual, we let $()$ denote an empty 0-tuple. If $(x_1, x_2, \ldots, x_0) = ()$ and $(\lambda_1, \lambda_2, \ldots, \lambda_0) = ()$, then $\sum_{i=1}^{0} \lambda_i x_i = (\text{empty sum}) = 0$. Hence, 0 is an element of E such that there exist some $t \in \mathbb{N}$, a t-tuple $(x_1, x_2, \ldots, x_t)$ of elements of S and a t-tuple $(\lambda_1, \lambda_2, \ldots, \lambda_t)$ of nonnegative reals such that $\sum_{i=1}^{t} \lambda_i x_i = 0$ (namely, $t = 0$, $(x_1, x_2, \ldots, x_t) = ()$ and $(\lambda_1, \lambda_2, \ldots, \lambda_t) = ()$). In other words,

$$0 \in \left\{ x \in E \;\middle|\; \left(\begin{array}{c} \text{there exist some } t \in \mathbb{N}, \\ \text{a } t\text{-tuple } (x_1, x_2, \ldots, x_t) \text{ of elements of } S \\ \text{and a } t\text{-tuple } (\lambda_1, \lambda_2, \ldots, \lambda_t) \text{ of nonnegative reals} \\ \text{such that } \sum_{i=1}^{t} \lambda_i x_i = x \end{array} \right) \right\} = C_3.$$

Let $y \in C_3$ and $z \in C_3$. Let λ and κ be two nonnegative reals. We shall show that $\lambda y + \kappa z \in C_3$.

We have

$$y \in C_3 = \left\{ x \in E \;\middle|\; \left(\begin{array}{c} \text{there exist some } t \in \mathbb{N}, \\ \text{a } t\text{-tuple } (x_1, x_2, \ldots, x_t) \text{ of elements of } S \\ \text{and a } t\text{-tuple } (\lambda_1, \lambda_2, \ldots, \lambda_t) \text{ of nonnegative reals} \\ \text{such that } \sum_{i=1}^{t} \lambda_i x_i = x \end{array} \right) \right\}.$$

In other words, we can write y in the form $y = \sum_{i=1}^{t} \lambda_i x_i$ for some $t \in \mathbb{N}$, some t-tuple $(x_1, x_2, \ldots, x_t)$ of elements of S and some t-tuple $(\lambda_1, \lambda_2, \ldots, \lambda_t)$ of nonnegative reals. Let us denote this t, this $(x_1, x_2, \ldots, x_t)$ and this $(\lambda_1, \lambda_2, \ldots, \lambda_t)$ by p, $(y_1, y_2, \ldots, y_p)$ and $(\mu_1, \mu_2, \ldots, \mu_p)$, respectively. Thus, p is an element of $\mathbb{N}$, and $(y_1, y_2, \ldots, y_p)$ is a p-tuple of elements of S, and $(\mu_1, \mu_2, \ldots, \mu_p)$ is a p-tuple of nonnegative reals such that $y = \sum_{i=1}^{p} \mu_i y_i$.

Similarly, use the assumption $z \in C_3$ to write z in the form $z = \sum_{i=1}^{q} \nu_i z_i$, where q is an element of $\mathbb{N}$, and where $(z_1, z_2, \ldots, z_q)$ is a q-tuple of elements of S, and where $(\nu_1, \nu_2, \ldots, \nu_q)$ is a q-tuple of nonnegative reals.

Define a $(p+q)$-tuple $(w_1, w_2, \ldots, w_{p+q})$ of elements of S by

$$(w_1, w_2, \ldots, w_{p+q}) = (y_1, y_2, \ldots, y_p, z_1, z_2, \ldots, z_q).$$

Then:

- For every $i \in \{1, 2, \ldots, p\}$, we have

$$w_i = y_i. \tag{68}$$

- For every $i \in \{p+1, p+2, \ldots, p+q\}$, we have

$$w_i = z_{i-p}. \tag{69}$$

Furthermore, define a $(p+q)$-tuple $(\eta_1, \eta_2, \ldots, \eta_{p+q})$ of nonnegative reals by

$$(\eta_1, \eta_2, \ldots, \eta_{p+q}) = (\lambda \mu_1, \lambda \mu_2, \ldots, \lambda \mu_p, \kappa \nu_1, \kappa \nu_2, \ldots, \kappa \nu_q)$$

cone in E which contains S as a subset. Thus, the intersection of all convex cones

(this is well-defined because both λ and κ are nonnegative reals and because $(\mu_1, \mu_2, \ldots, \mu_p)$ and $(\nu_1, \nu_2, \ldots, \nu_q)$ are tuples of nonnegative reals). Then:

- For every $i \in \{1, 2, \ldots, p\}$, we have

$$\eta_i = \lambda\mu_i. \tag{70}$$

- For every $i \in \{p+1, p+2, \ldots, p+q\}$, we have

$$\eta_i = \kappa\nu_{i-p}. \tag{71}$$

Now,

$$\sum_{i=1}^{p+q} \eta_i w_i = \sum_{i=1}^{p} \underbrace{\eta_i}_{\substack{=\lambda\mu_i \\ (\text{by }(70))}} \underbrace{w_i}_{\substack{=y_i \\ (\text{by }(68))}} + \sum_{i=p+1}^{p+q} \underbrace{\eta_i}_{\substack{=\kappa\nu_{i-p} \\ (\text{by }(71))}} \underbrace{w_i}_{\substack{=z_{i-p} \\ (\text{by }(69))}} = \underbrace{\sum_{i=1}^{p} \lambda\mu_i y_i}_{=\lambda\sum_{i=1}^{p}\mu_i y_i} + \underbrace{\sum_{i=p+1}^{p+q} \kappa\nu_{i-p}z_{i-p}}_{=\kappa\sum_{i=p+1}^{p+q}\nu_{i-p}z_{i-p}}$$

$$= \lambda\sum_{i=1}^{p}\mu_i y_i + \kappa\sum_{i=p+1}^{p+q}\nu_{i-p}z_{i-p} = \lambda\underbrace{\sum_{i=1}^{p}\mu_i y_i}_{=y} + \kappa\underbrace{\sum_{i=1}^{q}\nu_i z_i}_{=z}$$

(here, we substituted i for $i-p$ in the second sum)

$$= \lambda y + \kappa z.$$

Hence, there exist some $t \in \mathbb{N}$, a t-tuple $(x_1, x_2, \ldots, x_t)$ of elements of S and a t-tuple $(\lambda_1, \lambda_2, \ldots, \lambda_t)$ of nonnegative reals such that $\sum_{i=1}^{t} \lambda_i x_i = \lambda y + \kappa z$ (namely, $t = p+q$, $(x_1, x_2, \ldots, x_t) = (w_1, w_2, \ldots, w_{p+q})$ and $(\lambda_1, \lambda_2, \ldots, \lambda_t) = (\eta_1, \eta_2, \ldots, \eta_{p+q})$). In other words,

$$\lambda y + \kappa z \in \left\{ x \in E \;\middle|\; \left(\begin{array}{c} \text{there exist some } t \in \mathbb{N}, \\ \text{a } t\text{-tuple } (x_1, x_2, \ldots, x_t) \text{ of elements of } S \\ \text{and a } t\text{-tuple } (\lambda_1, \lambda_2, \ldots, \lambda_t) \text{ of nonnegative reals} \\ \text{such that } \sum_{i=1}^{t} \lambda_i x_i = x \end{array} \right) \right\} = C_3.$$

Now, let us forget that we fixed y, z, λ and κ. We thus have proven that every two elements $y \in C_3$ and $z \in C_3$ and every nonnegative reals λ and κ satisfy $\lambda y + \kappa z \in C_3$. If we rename y, z and κ as x, y and μ in this statement, we obtain the following: Every two elements $x \in C_3$ and $y \in C_3$ and every nonnegative reals λ and μ satisfy $\lambda x + \mu y \in C_3$.

Thus, we have proven the following two statements:

- We have $0 \in C_3$.

- Every two elements $x \in C_3$ and $y \in C_3$ and every nonnegative reals λ and μ satisfy $\lambda x + \mu y \in C_3$.

In other words, the set C_3 is a convex cone (by the definition of a convex cone).

[76]*Proof.* Let $w \in S$. Then, $\sum_{i=1}^{1} 1w = 1w = w$. Hence, there exist some $t \in \mathbb{N}$, a t-tuple

in E which contain S as a subset is a subset of C_3. In other words, C_1 is a subset of C_3 (since C_1 is the intersection of all convex cones in E which contain S as a subset). In other words, $C_1 \subseteq C_3$.

On the other hand, $C_3 \subseteq C_1$ [77]. Combined with $C_1 \subseteq C_3$, this shows that $C_1 = C_3$.

$(x_1, x_2, \ldots, x_t)$ of elements of S and a t-tuple $(\lambda_1, \lambda_2, \ldots, \lambda_t)$ of nonnegative reals such that $\sum_{i=1}^{t} \lambda_i x_i = w$ (namely, $t = 1$, $(x_1, x_2, \ldots, x_t) = (w)$ and $(\lambda_1, \lambda_2, \ldots, \lambda_t) = (1)$). In other words,

$$w \in \left\{ x \in E \mid \begin{pmatrix} \text{there exist some } t \in \mathbb{N}, \\ \text{a } t\text{-tuple } (x_1, x_2, \ldots, x_t) \text{ of elements of } S \\ \text{and a } t\text{-tuple } (\lambda_1, \lambda_2, \ldots, \lambda_t) \text{ of nonnegative reals} \\ \text{such that } \sum_{i=1}^{t} \lambda_i x_i = x \end{pmatrix} \right\} = C_3.$$

Let us now forget that we fixed w. We thus have proven that $w \in C_3$ for every $w \in S$. In other words, $S \subseteq C_3$.

[77]*Proof.* Let $w \in C_3$. We shall show that $w \in C_1$.

Let D be any convex cone in E which contains S as a subset. Thus, D is a convex cone in E and satisfies $S \subseteq D$.

We have

$$w \in C_3 = \left\{ x \in E \mid \begin{pmatrix} \text{there exist some } t \in \mathbb{N}, \\ \text{a } t\text{-tuple } (x_1, x_2, \ldots, x_t) \text{ of elements of } S \\ \text{and a } t\text{-tuple } (\lambda_1, \lambda_2, \ldots, \lambda_t) \text{ of nonnegative reals} \\ \text{such that } \sum_{i=1}^{t} \lambda_i x_i = x \end{pmatrix} \right\}.$$

In other words, there exist some $t \in \mathbb{N}$, a t-tuple $(x_1, x_2, \ldots, x_t)$ of elements of S and a t-tuple $(\lambda_1, \lambda_2, \ldots, \lambda_t)$ of nonnegative reals such that $\sum_{i=1}^{t} \lambda_i x_i = w$. Consider this t, this $(x_1, x_2, \ldots, x_t)$ and this $(\lambda_1, \lambda_2, \ldots, \lambda_t)$. For every $i \in \{1, 2, \ldots, t\}$, we have $x_i \in S \subseteq D$. Thus, $(x_i)_{i \in \{1, 2, \ldots, t\}}$ is a family of elements of D. Also, clearly, $(\lambda_i)_{i \in \{1, 2, \ldots, t\}}$ is a family of nonnegative reals (since $(\lambda_1, \lambda_2, \ldots, \lambda_t)$ is a t-tuple of nonnegative reals). Clearly, all but finitely many $i \in \{1, 2, \ldots, t\}$ satisfy $\lambda_i = 0$ (since there are only finitely many $i \in \{1, 2, \ldots, t\}$). Hence, Proposition 2.0i (applied to D and $\{1, 2, \ldots, t\}$ instead of C and I) shows that $\sum_{i \in \{1, 2, \ldots, t\}} \lambda_i x_i \in D$. Thus, $w = \sum_{i=1}^{t} \lambda_i x_i = \sum_{i \in \{1, 2, \ldots, t\}} \lambda_i x_i \in D$.

Let us now forget that we fixed D. We thus have proven that $w \in D$ whenever D is any convex cone in E which contains S as a subset. In other words, w lies in every convex cone in E which contains S as a subset. Hence, w lies in the intersection of all convex cones in E which contain S as a subset. In other words, w lies in C_1 (since C_1 is the intersection of all convex cones in E which contain S as a subset). In other words, $w \in C_1$.

Let us now forget that we fixed w. We thus have proven that $w \in C_1$ for every $w \in C_3$. In other words, $C_3 \subseteq C_1$, qed.

The set C_2 is a convex cone[78]. Moreover, $S \subseteq C_2$ [79]. Hence, C_2 is a convex

---

[78]*Proof.* We have $\sum\limits_{i \in S} \underbrace{0i}_{=0} = \sum\limits_{i \in S} 0 = 0$. Thus, 0 has the form $\sum\limits_{i \in S} \lambda_i i$ for some family $(\lambda_i)_{i \in S}$
of nonnegative reals indexed by elements of S and having the property that all but finitely
many $i \in S$ satisfy $\lambda_i = 0$ (namely, for $(\lambda_i)_{i \in S} = (0)_{i \in S}$). In other words, 0 is a linear
combination of the vectors s for $s \in S$ with nonnegative coefficients. In other words, $0 \in C_2$
(since C_2 is the set of all linear combinations of the vectors s for $s \in S$).

Now, let $y \in C_2$ and $z \in C_2$. Let λ and κ be two nonnegative reals. We shall show that
$\lambda y + \kappa z \in C_2$.

We have $y \in C_2$. In other words, y is a linear combination of the vectors s for $s \in S$ with
nonnegative coefficients (since C_2 is the set of all linear combinations of the vectors s for
$s \in S$ with nonnegative coefficients). In other words, y has the form $y = \sum\limits_{i \in S} \mu_i i$ for some
family $(\mu_i)_{i \in S}$ of nonnegative reals indexed by elements of S and having the property that
all but finitely many $i \in S$ satisfy $\mu_i = 0$. Similarly, z has the form $z = \sum\limits_{i \in S} \nu_i i$ for some
family $(\nu_i)_{i \in S}$ of nonnegative reals indexed by elements of S and having the property that
all but finitely many $i \in S$ satisfy $\nu_i = 0$. Consider these two families $(\mu_i)_{i \in S}$ and $(\nu_i)_{i \in S}$.

For every $i \in S$, both $\lambda \mu_i$ and $\kappa \nu_i$ are nonnegative reals (since λ, μ_i, κ and ν_i are
nonnegative reals), and hence the sum $\lambda \mu_i + \kappa \nu_i$ is a nonnegative real. Thus, we can define
a family $(\eta_i)_{i \in S}$ of nonnegative reals by setting

$$\left(\eta_i = \lambda \mu_i + \kappa \nu_i \qquad \text{for every } i \in S \right).$$

Consider this family $(\eta_i)_{i \in S}$. We know that all but finitely many $i \in S$ satisfy $\mu_i = 0$, and
we also know that all but finitely many $i \in S$ satisfy $\nu_i = 0$. Using these two facts, we see
that all but finitely many $i \in S$ satisfy $\eta_i = \lambda \underbrace{\mu_i}_{=0} + \kappa \underbrace{\nu_i}_{=0} = 0 + 0 = 0$. We have

$$\sum_{i \in S} \underbrace{\eta_i}_{=\lambda \mu_i + \kappa \nu_i} i = \sum_{i \in S} (\lambda \mu_i + \kappa \nu_i) i = \lambda \underbrace{\sum_{i \in S} \mu_i i}_{=y} + \kappa \underbrace{\sum_{i \in S} \nu_i z}_{=z} = \lambda y + \kappa z.$$

Hence, $\lambda y + \kappa z = \sum\limits_{i \in S} \eta_i i$. Therefore, $\lambda y + \kappa z$ has the form $\sum\limits_{i \in S} \lambda_i i$ for some family $(\lambda_i)_{i \in S}$
of nonnegative reals indexed by elements of S and having the property that all but finitely
many $i \in S$ satisfy $\lambda_i = 0$ (namely, for the family $(\lambda_i)_{i \in S} = (\eta_i)_{i \in S}$). In other words, $\lambda y + \kappa z$
is a linear combination of the vectors s for $s \in S$ with nonnegative coefficients. In other
words, $\lambda y + \kappa z \in C_2$ (since C_2 is the set of all linear combinations of the vectors s for $s \in S$
with nonnegative coefficients).

Now, let us forget that we fixed y, z, λ and κ. We thus have proven that every two
elements $y \in C_2$ and $z \in C_2$ and every nonnegative reals λ and κ satisfy $\lambda y + \kappa z \in C_2$. If
we rename y, z and κ as x, y and μ in this statement, we obtain the following: Every two
elements $x \in C_2$ and $y \in C_2$ and every nonnegative reals λ and μ satisfy $\lambda x + \mu y \in C_2$.

Thus, we have proven the following two statements:

- We have $0 \in C_2$.

- Every two elements $x \in C_2$ and $y \in C_2$ and every nonnegative reals λ and μ satisfy
  $\lambda x + \mu y \in C_2$.

In other words, the set C_2 is a convex cone (by the definition of a convex cone). Qed.

---

cone in E which contains S as a subset. Thus, the intersection of all convex cones in E which contain S as a subset is a subset of C_2. In other words, C_1 is a subset of C_2 (since C_1 is the intersection of all convex cones in E which contain S as a subset). In other words, $C_1 \subseteq C_2$.

On the other hand, $C_2 \subseteq C_1$ [80]. Combined with $C_1 \subseteq C_2$, this shows that $C_1 = C_2$.

Combining $C_1 = C_2$ with $C_1 = C_3$, we obtain $C_1 = C_2 = C_3$. Now, (65)

---

[79]*Proof.* Let $w \in S$. Define a family $(\mu_i)_{i \in S}$ of nonnegative reals by setting

$$\left(\mu_i = \begin{cases} 1, & \text{if } i = w; \\ 0, & \text{if } i \neq w \end{cases} \qquad \text{for every } i \in S \right).$$

It is easy to see that:

- All but finitely many $i \in S$ satisfy $\mu_i = 0$.

- We have $\sum\limits_{i \in S} \mu_i i = w$.

Hence, $w = \sum\limits_{i \in S} \mu_i i$. Thus, w has the form $\sum\limits_{i \in S} \lambda_i i$ for some family $(\lambda_i)_{i \in S}$ of nonnegative reals indexed by elements of S and having the property that all but finitely many $i \in S$ satisfy $\lambda_i = 0$ (namely, for the family $(\lambda_i)_{i \in S} = (\mu_i)_{i \in S}$). In other words, w is a linear combination of the vectors s for $s \in S$ with nonnegative coefficients. In other words, $w \in C_2$ (since C_2 is the set of all linear combinations of the vectors s for $s \in S$ with nonnegative coefficients).

Now, let us forget that we fixed w. We thus have proven that $w \in C_2$ for every $w \in S$. In other words, $S \subseteq C_2$.

[80]*Proof.* Let $w \in C_2$. We shall show that $w \in C_1$.

Let D be any convex cone in E which contains S as a subset. Thus, D is a convex cone in E and satisfies $S \subseteq D$.

We have $w \in C_2$. In other words, w belongs to the set C_2. In other words, w is a linear combination of the vectors s for $s \in S$ with nonnegative coefficients (since C_2 is the set of all linear combinations of the vectors s for $s \in S$ with nonnegative coefficients). In other words, w has the form $w = \sum\limits_{i \in S} \lambda_i i$ for some family $(\lambda_i)_{i \in S}$ of nonnegative reals indexed by elements of S and having the property that all but finitely many $i \in S$ satisfy $\lambda_i = 0$. Consider this family $(\lambda_i)_{i \in S}$.

For every $i \in S$, we have $i \in S \subseteq D$. Thus, $(i)_{i \in S}$ is a family of elements of D. Hence, Proposition 2.0i (applied to D, S, $(i)_{i \in S}$ and $(\lambda_i)_{i \in S}$ instead of C, I, $(x_i)_{i \in I}$ and $(\lambda_i)_{i \in I}$) shows that $\sum\limits_{i \in S} \lambda_i i \in D$. Thus, $w = \sum\limits_{i \in S} \lambda_i i \in D$.

Let us now forget that we fixed D. We thus have proven that $w \in D$ whenever D is any convex cone in E which contains S as a subset. In other words, w lies in every convex cone in E which contains S as a subset. Hence, w lies in the intersection of all convex cones in E which contain S as a subset. In other words, w lies in C_1 (since C_1 is the intersection of all convex cones in E which contain S as a subset). In other words, $w \in C_1$.

Let us now forget that we fixed w. We thus have proven that $w \in C_1$ for every $w \in C_2$. In other words, $C_2 \subseteq C_1$.

---

becomes

(the convex conic hull of S defined according to Definition 2.0j)

$= C_1 = C_2 =$ (the convex conic hull of S defined according to Definition 2.0k)

$$(72)$$

(by (66)). Also, (66) becomes

(the convex conic hull of S defined according to Definition 2.0k)

$= C_2 = C_3 =$ (the convex conic hull of S defined according to Definition 2.0l)

$$(73)$$

(by (67)).

Now, let us forget that we fixed S. We thus have shown that if S is any subset of E, then (72) and (73) hold. In other words, if S is any subset of E, then

(the convex conic hull of S defined according to Definition 2.0j)

$=$ (the convex conic hull of S defined according to Definition 2.0k)

$=$ (the convex conic hull of S defined according to Definition 2.0l) .

In other words, Definitions 2.0j, 2.0k and 2.0l are equivalent. This proves Proposition 2.0m **(a)**.

(c) We define the set C_3 as in our proof of Proposition 2.0m **(a)**. Then, (67) shows that

$C_3 =$ (the convex conic hull of S defined according to Definition 2.0l)

$=$ (the convex conic hull of S) $= \operatorname{cone} S$.

But in our proof of Proposition 2.0m **(a)**, we have shown that $S \subseteq C_3$. Thus, $S \subseteq C_3 = \operatorname{cone} S$. This proves Proposition 2.0m **(c)**.

(d) We define the set C_3 as in our proof of Proposition 2.0m **(a)**. Then, (67) shows that

$C_3 =$ (the convex conic hull of S defined according to Definition 2.0l)

$=$ (the convex conic hull of S) $= \operatorname{cone} S$.

But in our proof of Proposition 2.0m **(a)**, we have shown that the set C_3 is a convex cone. Since $C_3 = \operatorname{cone} S$, this rewrites as follows: The set $\operatorname{cone} S$ is a convex cone. This proves Proposition 2.0m **(d)**.

(f) Let D be a convex cone in E which contains S as a subset. Thus, the intersection of all convex cones in E which contain S as a subset is a subset of D. In other words, $\operatorname{cone} S$ is a subset of D (since $\operatorname{cone} S$ is the intersection of all convex cones in E which contain S as a subset (by Definition 2.0k)). In other words, D contains $\operatorname{cone} S$ as a subset.

Now, let us forget that we fixed D. We thus have shown that if D is a convex cone in E which contains S as a subset, then D contains $\operatorname{cone} S$ as a subset. In other words, every convex cone in E which contains S as a subset also contains $\operatorname{cone} S$ as a subset. This proves Proposition 2.0m **(f)**.

(b) Every two elements $x \in 0$ and $y \in 0$ and every nonnegative reals λ and μ satisfy $\lambda x + \mu y \in 0$ [81]. Furthermore, $0 \in 0$. Thus, we have proven the following two statements:

- We have $0 \in 0$.

- Every two elements $x \in 0$ and $y \in 0$ and every nonnegative reals λ and μ satisfy $\lambda x + \mu y \in 0$.

In other words, the set 0 is a convex cone (by the definition of a convex cone).

Thus, 0 is a convex cone in E which contains $\varnothing$ as a subset. Now, Proposition 2.0m **(f)** (applied to $S = \varnothing$) shows that every convex cone in E which contains $\varnothing$ as a subset also contains $\operatorname{cone} \varnothing$ as a subset. Thus, the set 0 contains $\operatorname{cone} \varnothing$ as a subset (since 0 is a convex cone in E which contains $\varnothing$ as a subset). In other words, $\operatorname{cone} \varnothing \subseteq 0$.

On the other hand, $0 \subseteq \operatorname{cone} \varnothing$ [82]. Combining this with $\operatorname{cone} \varnothing \subseteq 0$, we obtain $\operatorname{cone} \varnothing = 0$. This proves Proposition 2.0m **(b)**.

(g) Let T be a subset of $\operatorname{cone} S$. The set $\operatorname{cone} S$ is a convex cone (by Proposition 2.0m **(d)**) and contains T as a subset (since $T \subseteq \operatorname{cone} S$). In other words, $\operatorname{cone} S$ is a convex cone in E which contains T as a subset.

But Proposition 2.0m **(f)** (applied to T instead of S) shows that every convex cone in E which contains T as a subset also contains $\operatorname{cone} T$ as a subset. Thus, the set $\operatorname{cone} S$ contains $\operatorname{cone} T$ as a subset (since $\operatorname{cone} S$ is a convex cone in E which contains T as a subset). In other words, $\operatorname{cone} T \subseteq \operatorname{cone} S$. This proves Proposition 2.0m **(g)**.

(e) Let T be a subset of S. Thus, $T \subseteq S \subseteq \operatorname{cone} S$ (by Proposition 2.0m **(c)**). In other words, T is a subset of $\operatorname{cone} S$. Hence, Proposition 2.0m **(g)** yields $\operatorname{cone} T \subseteq \operatorname{cone} S$. This proves Proposition 2.0m **(e)**.

(h) Proposition 2.0m **(d)** shows that $\operatorname{cone} S$ is a convex cone.

---

[81] *Proof.* Let $x \in 0$ and $y \in 0$ be two elements. Let λ and μ be two nonnegative reals. Then, $x = 0$ (since $x \in 0$) and $y = 0$ (since $y \in 0$), so that $\lambda \underbrace{x}_{=0} + \mu \underbrace{y}_{=0} = \lambda 0 + \mu 0 = 0 \in 0$, qed.

[82] *Proof.* Proposition 2.0m **(d)** (applied to $S = \varnothing$) shows that the set $\operatorname{cone} \varnothing$ is a convex cone. In other words, it satisfies the following two statements:

- We have $0 \in \operatorname{cone} \varnothing$.

- Every two elements $x \in \operatorname{cone} \varnothing$ and $y \in \operatorname{cone} \varnothing$ and every nonnegative reals λ and μ satisfy $\lambda x + \mu y \in \operatorname{cone} \varnothing$.

(This is because of our definition of a convex cone.)
In particular, $0 \in \operatorname{cone} \varnothing$, so that $\{0\} \subseteq \operatorname{cone} \varnothing$. Thus, $0 = \{0\} \subseteq \operatorname{cone} \varnothing$, qed.

Recall that any convex cone is a convex set. Applying this to the convex cone cone S, we conclude that cone S is a convex set (since cone S is a convex cone). In other words, cone S is a convex subset of E.

Proposition 2.0m **(c)** shows that $S \subseteq$ cone S. Thus, the set cone S contains S as a subset. Hence, cone S is a convex subset of E which contains S as a subset.

But Proposition 2.0f **(f)** shows that every convex subset of E which contains S as a subset also contains conv . hull S as a subset. Thus, the set cone S contains conv . hull S as a subset (since cone S is a convex subset of E which contains S as a subset). In other words, conv . hull $S \subseteq$ cone S. This proves Proposition 2.0m **(h)**. $\square$

Proof of Proposition 2.0n. The proof of Proposition 2.0n is analogous to the proof of Proposition 2.0g. (Of course, instead of using properties of convex sets, we now need to use the corresponding properties of convex cones.) $\square$

Proof of Proposition 2.0o. **(a)** Let $S = \{x_1, x_2, \ldots, x_n\}$. Let Q be the set of all linear combinations of the vectors x_1, x_2, ..., x_n with nonnegative coefficients.

Recall that Q is the set of all linear combinations of the vectors x_1, x_2, ..., x_n with nonnegative coefficients. In other words, Q is the set of all linear combinations of the vectors x_i for $i \in \{1, 2, \ldots, n\}$ with nonnegative coefficients.

The set Q is a convex cone[83]. Moreover, $S \subseteq Q$ [84]. Hence, Q is a convex cone

---

[83] *Proof.* We have $\sum\limits_{i \in \{1,2,\ldots,n\}} \underbrace{0x_i}_{=0} = \sum\limits_{i \in \{1,2,\ldots,n\}} 0 = 0$. Hence, $0 = \sum\limits_{i \in \{1,2,\ldots,n\}} 0x_i$. Thus, 0

has the form $\sum\limits_{i \in \{1,2,\ldots,n\}} \lambda_i x_i$ for some family $(\lambda_i)_{i \in \{1,2,\ldots,n\}}$ of nonnegative reals indexed by
elements of $\{1, 2, \ldots, n\}$ and having the property that all but finitely many $i \in \{1, 2, \ldots, n\}$ satisfy $\lambda_i = 0$ (namely, for $(\lambda_i)_{i \in S} = (0)_{i \in S}$). In other words, 0 is a linear combination of the vectors x_i for $i \in \{1, 2, \ldots, n\}$ with nonnegative coefficients. In other words, $0 \in Q$ (since Q is the set of all linear combinations of the vectors x_i for $i \in \{1, 2, \ldots, n\}$ with nonnegative coefficients).

Now, let $y \in Q$ and $z \in Q$. Let λ and κ be two nonnegative reals. We shall show that $\lambda y + \kappa z \in Q$.

We have $y \in Q$. In other words, y is a linear combination of the vectors x_i for $i \in \{1, 2, \ldots, n\}$ with nonnegative coefficients (since Q is the set of all linear combinations of the vectors x_i for $i \in \{1, 2, \ldots, n\}$ with nonnegative coefficients). In other words, y has the form $y = \sum\limits_{i \in \{1,2,\ldots,n\}} \mu_i x_i$ for some family $(\mu_i)_{i \in \{1,2,\ldots,n\}}$ of nonnegative reals indexed by
elements of $\{1, 2, \ldots, n\}$ and having the property that all but finitely many $i \in \{1, 2, \ldots, n\}$ satisfy $\mu_i = 0$. Similarly, z has the form $z = \sum\limits_{i \in \{1,2,\ldots,n\}} \nu_i x_i$ for some family $(\nu_i)_{i \in \{1,2,\ldots,n\}}$
of nonnegative reals indexed by elements of $\{1, 2, \ldots, n\}$ and having the property that all but finitely many $i \in \{1, 2, \ldots, n\}$ satisfy $\nu_i = 0$. Consider these two families $(\mu_i)_{i \in \{1,2,\ldots,n\}}$ and $(\nu_i)_{i \in \{1,2,\ldots,n\}}$.

For every $i \in \{1, 2, \ldots, n\}$, both $\lambda \mu_i$ and $\kappa \nu_i$ are nonnegative reals (since λ, μ_i, κ and ν_i are nonnegative reals), and hence the sum $\lambda \mu_i + \kappa \nu_i$ is a nonnegative real. Thus, we can define a family $(\eta_i)_{i \in \{1,2,\ldots,n\}}$ of nonnegative reals by setting

$$(\eta_i = \lambda \mu_i + \kappa \nu_i \qquad \text{for every } i \in \{1, 2, \ldots, n\}).$$

Consider this family $(\eta_i)_{i \in \{1,2,\ldots,n\}}$. We know that all but finitely many $i \in \{1, 2, \ldots, n\}$ satisfy $\eta_i = 0$ (since there are only finitely many $i \in \{1, 2, \ldots, n\}$). We have

$$\sum\limits_{i \in \{1,2,\ldots,n\}} \underbrace{\eta_i}_{=\lambda\mu_i + \kappa\nu_i} x_i = \sum\limits_{i \in \{1,2,\ldots,n\}} (\lambda\mu_i + \kappa\nu_i) x_i$$

$$= \lambda \underbrace{\sum\limits_{i \in \{1,2,\ldots,n\}} \mu_i x_i}_{=y} + \kappa \underbrace{\sum\limits_{i \in \{1,2,\ldots,n\}} \nu_i x_i}_{=z} = \lambda y + \kappa z.$$

Hence, $\lambda y + \kappa z = \sum\limits_{i \in \{1,2,\ldots,n\}} \eta_i x_i$. Therefore, $\lambda y + \kappa z$ has the form $\sum\limits_{i \in \{1,2,\ldots,n\}} \lambda_i x_i$ for some
family $(\lambda_i)_{i \in \{1,2,\ldots,n\}}$ of nonnegative reals indexed by elements of $\{1, 2, \ldots, n\}$ and having the property that all but finitely many $i \in \{1, 2, \ldots, n\}$ satisfy $\lambda_i = 0$ (namely, for the family $(\lambda_i)_{i \in \{1,2,\ldots,n\}} = (\eta_i)_{i \in \{1,2,\ldots,n\}}$). In other words, $\lambda y + \kappa z$ is a linear combination of the vectors x_i or $i \in \{1, 2, \ldots, n\}$ with nonnegative coefficients. In other words, $\lambda y + \kappa z \in Q$ (since Q is the set of all linear combinations of the vectors x_i for $i \in \{1, 2, \ldots, n\}$ with nonnegative coefficients).

Now, let us forget that we fixed y, z, λ and κ. We thus have proven that every two elements $y \in Q$ and $z \in Q$ and every nonnegative reals λ and κ satisfy $\lambda y + \kappa z \in Q$. If we rename y, z and κ as x, y and μ in this statement, we obtain the following: Every two elements $x \in Q$ and $y \in Q$ and every nonnegative reals λ and μ satisfy $\lambda x + \mu y \in Q$.

---

in E which contains S as a subset. Thus, the intersection of all convex cones in E which contain S as a subset is a subset of Q. In other words, $\operatorname{cone} S$ is a subset of Q (since $\operatorname{cone} S$ is the intersection of all convex cones in E which contain S as a subset (because of Definition 2.0j)). In other words, $\operatorname{cone} S \subseteq Q$.

On the other hand, $Q \subseteq \operatorname{cone} S$ [85]. Combining this with $\operatorname{cone} S \subseteq Q$, we

---

Thus, we have proven the following two statements:

- We have $0 \in Q$.

- Every two elements $x \in Q$ and $y \in Q$ and every nonnegative reals λ and μ satisfy $\lambda x + \mu y \in Q$.

In other words, the set Q is a convex cone (by the definition of a convex cone). Qed.

[84] *Proof.* Let $w \in S$. Then, $w \in S = \{x_1, x_2, \ldots, x_n\}$. Thus, there exists some $k \in \{1, 2, \ldots, n\}$ such that $w = x_k$. Consider this k.

Define a family $(\mu_i)_{i \in \{1,2,\ldots,n\}}$ of nonnegative reals by setting

$$\left(\mu_i = \begin{cases} 1, & \text{if } i = k; \\ 0, & \text{if } i \neq k \end{cases} \qquad \text{for every } i \in \{1, 2, \ldots, n\} \right).$$

Then, it is easy to see that:

- All but finitely many $i \in \{1, 2, \ldots, n\}$ satisfy $\mu_i = 0$.

- We have $\displaystyle\sum_{i \in \{1,2,\ldots,n\}} \mu_i x_i = x_k$.

Hence, $w = x_k = \displaystyle\sum_{i \in \{1,2,\ldots,n\}} \mu_i x_i$. Thus, w has the form $\displaystyle\sum_{i \in \{1,2,\ldots,n\}} \lambda_i x_i$ for some family $(\lambda_i)_{i \in \{1,2,\ldots,n\}}$ of nonnegative reals indexed by elements of $\{1, 2, \ldots, n\}$ and having the property that all but finitely many $i \in \{1, 2, \ldots, n\}$ satisfy $\lambda_i = 0$ (namely, for the family $(\lambda_i)_{i \in \{1,2,\ldots,n\}} = (\mu_i)_{i \in \{1,2,\ldots,n\}}$). In other words, w is a linear combination of the vectors x_i for $i \in \{1, 2, \ldots, n\}$ with nonnegative coefficients. In other words, $w \in Q$ (since Q is the set of all linear combinations of the vectors x_i for $i \in \{1, 2, \ldots, n\}$ with nonnegative coefficients).

Now, let us forget that we fixed w. We thus have proven that $w \in Q$ for every $w \in S$. In other words, $S \subseteq Q$.

[85] *Proof.* Let $w \in Q$. We shall show that $w \in \operatorname{cone} S$.

Let D be any convex cone in E which contains S as a subset. Thus, D is a convex cone in E and satisfies $S \subseteq D$.

We have $w \in Q$. In other words, w is a linear combination of the vectors x_i for $i \in \{1, 2, \ldots, n\}$ with nonnegative coefficients (since Q is the set of all linear combinations of the vectors x_i for $i \in \{1, 2, \ldots, n\}$ with nonnegative coefficients). In other words, w has the form $w = \displaystyle\sum_{i \in \{1,2,\ldots,n\}} \lambda_i x_i$ for some family $(\lambda_i)_{i \in \{1,2,\ldots,n\}}$ of nonnegative reals indexed by elements of $\{1, 2, \ldots, n\}$ and having the property that all but finitely many $i \in \{1, 2, \ldots, n\}$ satisfy $\lambda_i = 0$. Consider this family $(\lambda_i)_{i \in \{1,2,\ldots,n\}}$.

For every $i \in \{1, 2, \ldots, n\}$, we have $x_i \in \{x_1, x_2, \ldots, x_n\} = S \subseteq D$. Thus, $(x_i)_{i \in \{1,2,\ldots,n\}}$ is a family of elements of D. Hence, Proposition 2.0i (applied to D, $\{1, 2, \ldots, n\}$, $(x_i)_{i \in \{1,2,\ldots,n\}}$ and $(\lambda_i)_{i \in \{1,2,\ldots,n\}}$ instead of C, I, $(x_i)_{i \in I}$ and $(\lambda_i)_{i \in I}$) shows that $\displaystyle\sum_{i \in \{1,2,\ldots,n\}} \lambda_i x_i \in D$.

Thus, $w = \displaystyle\sum_{i \in \{1,2,\ldots,n\}} \lambda_i x_i \in D$.

obtain $\operatorname{cone} S = Q$. Thus,

$$\operatorname{cone}\underbrace{\{x_1, x_2, \ldots, x_n\}}_{=S}$$
$$= \operatorname{cone} S = Q$$
$$= (\text{the set of all linear combinations of the}$$
$$\text{vectors } x_1, x_2, \ldots, x_n \text{ with nonnegative coefficients})$$

(by the definition of Q). This proves Proposition 2.0o **(a)**.

(b) We have the following chain of logical equivalences:

$$\left(x \in \underbrace{\operatorname{cone}\{x_1, x_2, \ldots, x_n\}}_{\substack{=(\text{the set of all linear combinations of the} \\ \text{vectors } x_1, x_2, \ldots, x_n \text{ with nonnegative coefficients}) \\ (\text{by Proposition 2.0o } (\mathbf{a}))} \right)$$

$$\Longleftrightarrow \left(\begin{array}{c} x \in (\text{the set of all linear combinations of the} \\ \text{vectors } x_1, x_2, \ldots, x_n \text{ with nonnegative coefficients}) \end{array} \right)$$

$$\Longleftrightarrow (x \text{ is a linear combination of the vectors } x_1, x_2, \ldots, x_n$$
$$\text{with nonnegative coefficients}).$$

In other words, we have $x \in \operatorname{cone}\{x_1, x_2, \ldots, x_n\}$ if and only if x is a linear combination of the vectors x_1, x_2, ..., x_n with nonnegative coefficients. This proves Proposition 2.0o **(b)**. $\qquad\square$

Proof of Proposition 2.0p. **(b)** Let $(\nu_s)_{s\in S}$ is a family of nonnegative reals indexed by elements of S.

Clearly, there are only finitely many $i \in S$ (since the set S is finite). Hence, all but finitely many $i \in S$ satisfy $\nu_i = 0$.

We know that $(\nu_s)_{s\in S}$ is a family of nonnegative reals. Renaming the index s as i in this statement, we conclude that $(\nu_i)_{i\in S}$ is a family of nonnegative reals.

We have $\sum_{s\in S} \nu_s s = \sum_{i\in S} \nu_i i$ (here, we renamed the summation index s as i). Thus, $\sum_{s\in S} \nu_s s$ has the form $\sum_{i\in S} \lambda_i i$ for some family $(\lambda_i)_{i\in S}$ of nonnegative reals indexed by the elements of S such that all but finitely many $i \in S$ satisfy $\lambda_i = 0$ (namely, for

---

Let us now forget that we fixed D. We thus have proven that $w \in D$ whenever D is any convex cone in E which contains S as a subset. In other words, w lies in every convex cone in E which contains S as a subset. Hence, w lies in the intersection of all convex cones in E which contain S as a subset. In other words, w lies in $\operatorname{cone} S$ (since $\operatorname{cone} S$ is the intersection of all convex cones in E which contain S as a subset (because of Definition 2.0j)). In other words, $w \in \operatorname{cone} S$.

Let us now forget that we fixed w. We thus have proven that $w \in \operatorname{cone} S$ for every $w \in Q$. In other words, $Q \subseteq \operatorname{cone} S$, qed.

the family $(\lambda_i)_{i \in S} = (\nu_i)_{i \in S}$. In other words, $\sum_{s \in S} \nu_s s$ is a linear combination of the vectors s for $s \in S$ with nonnegative coefficients. In other words, $\sum_{s \in S} \nu_s s$ belongs to the set of all linear combinations of the vectors s for $s \in S$ with nonnegative coefficients. In other words, $\sum_{s \in S} \nu_s s$ belongs to cone S (since cone S is the set of all linear combinations of the vectors s for $s \in S$ with nonnegative coefficients (according to Definition 2.0k)). In other words, $\sum_{s \in S} \nu_s s \in$ cone S. This proves Proposition 2.0p **(b)**.

(c) Let p be an element of cone S. Thus, $p \in$ cone S. In other words, p is a linear combination of the vectors s for $s \in S$ with nonnegative coefficients (since cone S is the set of all linear combinations of the vectors s for $s \in S$ with nonnegative coefficients (according to Definition 2.0k)). In other words, p can be written in the form $\sum_{i \in S} \lambda_i i$ for some family $(\lambda_i)_{i \in S}$ of nonnegative reals indexed by the elements of S such that all but finitely many $i \in S$ satisfy $\lambda_i = 0$. Consider this $(\lambda_i)_{i \in S}$. Thus, $p = \sum_{i \in S} \lambda_i i$.

We know that $(\lambda_i)_{i \in S}$ is a family of nonnegative reals. Renaming the index i as s in this statement, we conclude that $(\lambda_s)_{s \in S}$ is a family of nonnegative reals. Now, $p = \sum_{i \in S} \lambda_i i = \sum_{s \in S} \lambda_s s$ (here, we renamed the summation index i as s). Thus, there exists a family $(\nu_s)_{s \in S}$ of nonnegative reals indexed by elements of S such that $p = \sum_{s \in S} \nu_s s$ (namely, $(\nu_s)_{s \in S} = (\lambda_s)_{s \in S}$). This proves Proposition 2.0p **(c)**.

(a) If $(\nu_s)_{s \in S}$ is a family of nonnegative reals, then $\sum_{s \in S} \nu_s s \in$ cone S (according to Proposition 2.0p **(b)**). In other words,

$$\left\{ \sum_{s \in S} \nu_s s \;\mid\; (\nu_s)_{s \in S} \text{ is a family of nonnegative reals} \right\} \subseteq \text{cone } S. \qquad (74)$$

On the other hand,

$$\text{cone } S \subseteq \left\{ \sum_{s \in S} \nu_s s \;\mid\; (\nu_s)_{s \in S} \text{ is a family of nonnegative reals} \right\}$$

[86]. Combining this with (74), we obtain

$$\operatorname{cone} S = \left\{ \sum_{s \in S} \nu_s s \;\middle|\; (\nu_s)_{s \in S} \text{ is a family of nonnegative reals} \right\}.$$

This proves Proposition 2.0p **(a)**. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

16.3. Proofs for Section 3

Proof of Proposition 2.0r. Let $x \in C$.

We have $x \in C = \operatorname{conv.hull} \{x_1, x_2, \ldots, x_t\}$, so that x is a convex combination of the vectors x_1, x_2, $\ldots$, x_t. In other words, there exist t nonnegative elements λ_1, λ_2, $\ldots$, λ_t of $\mathbb{R}$ such that $\sum_{i=1}^{t} \lambda_i = 1$ and $\sum_{i=1}^{t} \lambda_i x_i = x$. Consider these λ_1, λ_2, $\ldots$, λ_t.

There exists at least one $j \in \{1, 2, \ldots, t\}$ such that $\lambda_j \neq 0$ (since otherwise, we would have $\lambda_i = 0$ for every $i \in \{1, 2, \ldots, t\}$, so that $\sum_{i=1}^{t} \underbrace{\lambda_i}_{=0} = \sum_{i=1}^{t} 0 = 0$, which would contradict $\sum_{i=1}^{t} \lambda_i = 1 \neq 0$). Consider this j. Then, λ_j is nonnegative and satisfies $\lambda_j \neq 0$. Thus, $\lambda_j > 0$.

Combining $f(x_j) < \delta$ (by (2), applied to $i = j$) with $\lambda_j > 0$, we obtain $\lambda_j f(x_j) < \lambda_j \delta$. In other words,

$$\lambda_j f(x_j) - \lambda_j \delta < 0. \tag{75}$$

---

[86]*Proof.* Let $p \in \operatorname{cone} S$. Thus, there exists a family $(\nu_s)_{s \in S}$ of nonnegative reals indexed by elements of S such that $p = \sum_{s \in S} \nu_s s$ (according to Proposition 2.0p **(c)**). In other words, p has the form $\sum_{s \in S} \nu_s s$ for some family $(\nu_s)_{s \in S}$ of nonnegative reals. In other words,

$$p \in \left\{ \sum_{s \in S} \nu_s s \;\middle|\; (\nu_s)_{s \in S} \text{ is a family of nonnegative reals} \right\}.$$

Now, let us forget that we fixed p. We thus have proven that every $p \in \operatorname{cone} S$ satisfies $p \in \left\{ \sum_{s \in S} \nu_s s \;\middle|\; (\nu_s)_{s \in S} \text{ is a family of nonnegative reals} \right\}$. In other words,

$$\operatorname{cone} S \subseteq \left\{ \sum_{s \in S} \nu_s s \;\middle|\; (\nu_s)_{s \in S} \text{ is a family of nonnegative reals} \right\},$$

qed.

Since $x = \sum\limits_{i=1}^{t} \lambda_i x_i$, we have

$$f(x) = f\left(\sum_{i=1}^{t} \lambda_i x_i\right) = \sum_{i=1}^{t} \lambda_i f(x_i) \qquad \text{(since the map } f \text{ is } \mathbb{R}\text{-linear)}$$

$$= \underbrace{\sum_{i=1}^{t} \lambda_i f(x_i) - \sum_{i=1}^{t} \lambda_i \delta}_{= \sum\limits_{i=1}^{t} (\lambda_i f(x_i) - \lambda_i \delta)} + \underbrace{\sum_{i=1}^{t} \lambda_i}_{=1} \delta$$

$$= \sum_{i=1}^{t} (\lambda_i f(x_i) - \lambda_i \delta) + \delta. \qquad (76)$$

We know that every $i \in \{1, 2, \ldots, t\}$ satisfies $\lambda_i f(x_i) - \lambda_i \delta \leq 0$ (since (2) yields $f(x_i) < \delta$, so that $\lambda_i f(x_i) \leq \lambda_i \delta$ (since $\lambda_i \geq 0$), so that $\lambda_i f(x_i) - \lambda_i \delta \leq 0$). Hence, every addend of the sum $\sum\limits_{i=1}^{t} (\lambda_i f(x_i) - \lambda_i \delta)$ is nonpositive. Since we know that at least one addend of this sum is actually negative (namely, the addend for $i = j$, because of (75)), this sum must thus be < 0. Now, (76) becomes

$$f(x) = \underbrace{\sum_{i=1}^{t} (\lambda_i f(x_i) - \lambda_i \delta)}_{<0} + \delta < \delta.$$

This proves Proposition 2.0r. $\qquad\qquad\square$

Proof of Proposition 2.0s. Let $x \in \text{cone } S$. Then,

$$x \in \text{cone } S = \left\{ \sum_{s \in S} \nu_s s \mid (\nu_s)_{s \in S} \text{ is a family of nonnegative reals} \right\}$$

(by the definition of $\text{cone } S$). Hence, there exists a family $(\nu_s)_{s \in S}$ of nonnegative reals such that $x = \sum\limits_{s \in S} \nu_s s$. Consider this $(\nu_s)_{s \in S}$.

Each $s \in S$ satisfies

$$\nu_s f(s) \leq 0 \qquad \text{(since } \nu_s \geq 0 \text{ and } f(s) \leq 0 \text{ (by (3)))}. \qquad (77)$$

Since $x = \sum\limits_{s \in S} \nu_s s$, we have

$$f(x) = f\left(\sum_{s \in S} \nu_s s\right) = \sum_{s \in S} \underbrace{\nu_s f(s)}_{\substack{\leq 0 \\ \text{(by (3))}}} \qquad \text{(since the map } f \text{ is } \mathbb{R}\text{-linear)}$$

$$\leq \sum_{s \in S} 0 = 0.$$

This proves Proposition 2.0s. $\qquad\qquad\square$

Proof of Lemma 2.0t. **(a)** Let $v \in \mathbb{R}^n$ be a column vector satisfying $v \geq 0$. We must prove that all coordinates of the column vector v are nonnegative.

We have $v \geq 0$. In other words, every $i \in \{1, 2, \ldots, n\}$ satisfies

$$(\text{the } i\text{-th coordinate of } v) \geq (\text{the } i\text{-th coordinate of } 0)$$

(by the definition of $v \geq 0$). Hence, every $i \in \{1, 2, \ldots, n\}$ satisfies

$$(\text{the } i\text{-th coordinate of } v) \geq (\text{the } i\text{-th coordinate of } 0) = 0.$$

In other words, for every $i \in \{1, 2, \ldots, n\}$, the i-th coordinate of v is nonnegative. In other words, the coordinates of the column vector v are nonnegative. This proves Lemma 2.0t **(a)**.

(b) The proof of Lemma 2.0t **(b)** is completely identical to the proof of Lemma 2.0t **(a)** given above (except that "$\mathbb{R}^n$" must be replaced by "$(\mathbb{R}^n)^*$", and the word "column" must be replaced by "row"). $\qquad \square$

Proof of Lemma 2.0u. Write the vector x in the form $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ with x_1, x_2, $\ldots$, x_n being real numbers. Then, for every $i \in \{1, 2, \ldots, n\}$, we have

$$(\text{the } i\text{-th coordinate of } x) = x_i. \tag{78}$$

But now, let u be the vector $\begin{pmatrix} \max\{x_1, 0\} \\ \max\{x_2, 0\} \\ \vdots \\ \max\{x_n, 0\} \end{pmatrix}$. Then, for every $i \in \{1, 2, \ldots, n\}$, we have

$$(\text{the } i\text{-th coordinate of } u) = \max\{x_i, 0\} \tag{79}$$
$$\geq 0 = (\text{the } i\text{-th coordinate of } 0).$$

Thus, $u \geq 0$.

Furthermore, let v be the vector $\begin{pmatrix} -\min\{x_1, 0\} \\ -\min\{x_2, 0\} \\ \vdots \\ -\min\{x_n, 0\} \end{pmatrix}$. Then, for every $i \in \{1, 2, \ldots, n\}$, we have

$$(\text{the } i\text{-th coordinate of } v) = - \underbrace{\min\{x_i, 0\}}_{\leq 0} \tag{80}$$
$$\geq -0 = 0 = (\text{the } i\text{-th coordinate of } 0).$$

Thus, $v \geq 0$.

Now, every $\lambda \in \mathbb{R}$ satisfies

$$\max\{\lambda, 0\} + \min\{\lambda, 0\} = \lambda. \tag{81}$$

[87]

Now, every $i \in \{1, 2, \ldots, n\}$ satisfies

(the i-th coordinate of $u - v$)

$= \underbrace{(\text{the } i\text{-th coordinate of } u)}_{\substack{=\max\{x_i, 0\} \\ (\text{by } (79))}} - \underbrace{(\text{the } i\text{-th coordinate of } v)}_{\substack{=-\min\{x_i, 0\} \\ (\text{by } (80))}}$

$= \max\{x_i, 0\} - (-\min\{x_i, 0\}) = \max\{x_i, 0\} + \min\{x_i, 0\} = x_i$

(by (81), applied to $\lambda = x_i$)

$= (\text{the } i\text{-th coordinate of } x) \qquad (\text{by } (78))$.

Thus, $u - v = x$.

So we know that $u \geq 0$, $v \geq 0$ and $x = u - v$. Hence, there exist two vectors y and z in $\mathbb{R}^n$ such that $y \geq 0$, $z \geq 0$ and $x = y - z$ (namely, $y = u$ and $z = v$). This proves Lemma 2.0u. $\qquad \square$

Proof of Lemma 2.0v. The row vector $x \in (\mathbb{R}^n)^*$ satisfies $x \geq 0$. Thus, Lemma 2.0t **(b)** (applied to $v = x$) shows that the coordinates of the row vector x are nonnegative. Let λ_1, λ_2, ..., λ_n be the coordinates of the row vector x. Then, λ_1, λ_2, ..., λ_n are nonnegative (since the coordinates of the row vector x are nonnegative).

The column vector $y \in \mathbb{R}^n$ satisfies $y \geq 0$. Thus, Lemma 2.0t **(a)** (applied to $v = y$) shows that the coordinates of the column vector y are nonnegative. Let μ_1, μ_2, ..., μ_n be the coordinates of the column vector y. Then, μ_1, μ_2, ..., μ_n are nonnegative (since the coordinates of the column vector y are nonnegative).

By the definition of the product of a row vector with a column vector, we have $xy = \sum_{i=1}^{n} \lambda_i \mu_i$ (since the coordinates of the row vector x are λ_1, λ_2, ..., λ_n, while the coordinates of the column vector y are μ_1, μ_2, ..., μ_n).

---

[87] *Proof of (81):* Let $\lambda \in \mathbb{R}$. Then, we must be in one of the following two cases:

    Case 1: We have $\lambda \geq 0$.

    Case 2: We have $\lambda < 0$.

    Let us first consider Case 1. In this case, $\lambda \geq 0$, so that $\max\{\lambda, 0\} = \lambda$ and $\min\{\lambda, 0\} = 0$, and thus $\underbrace{\max\{\lambda, 0\}}_{=\lambda} + \underbrace{\min\{\lambda, 0\}}_{=0} = \lambda$. Thus, (81) is proven in Case 1.

    Let us next consider Case 2. In this case, $\lambda < 0$, so that $\max\{\lambda, 0\} = 0$ and $\min\{\lambda, 0\} = \lambda$, and thus $\underbrace{\max\{\lambda, 0\}}_{=0} + \underbrace{\min\{\lambda, 0\}}_{=\lambda} = \lambda$. Thus, (81) is proven in Case 2.

    Hence, we have proven (81) in each of the cases 1 and 2. Thus, (81) always holds (since cases 1 and 2 cover all possibilities). Qed.

But for every $i \in \{1, 2, \ldots, n\}$, the real $\lambda_i \mu_i$ is nonnegative (since λ_i is nonnegative and μ_i is nonnegative), i. e., we have $\lambda_i \mu_i \geq 0$. Thus, $xy = \sum_{i=1}^{n} \underbrace{\lambda_i \mu_i}_{\geq 0} \geq \sum_{i=1}^{n} 0 = 0$. This proves Lemma 2.0v. $\qquad\square$

16.4. Proofs for Section 4

Proof of Lemma 2.0x. Let $n \in \mathbb{N}$, and let I_1, I_2, $\ldots$, I_n be n closed intervals. We are going to prove that $I_1 \cap I_2 \cap \cdots \cap I_n$ is a closed interval.

In order to show that, we will prove that

for every $k \in \{0, 1, \ldots, n\}$, the set $I_1 \cap I_2 \cap \cdots \cap I_k$ is a closed interval. (82)

Proof of (82): We will prove (82) by induction over k:

Induction base: For $k = 0$, the set $I_1 \cap I_2 \cap \cdots \cap I_k$ is a closed interval (because for $k = 0$, the set $I_1 \cap I_2 \cap \cdots \cap I_k$ equals $I_1 \cap I_2 \cap \cdots \cap I_0 = \mathbb{R}$, and we know that $\mathbb{R}$ is a closed interval). In other words, (82) holds for $k = 0$. This completes the induction base.

Induction step: Let $i \in \{1, 2, \ldots, n\}$. Assume that (82) holds for $k = i - 1$. We now must show that (82) holds for $k = i$.

Since (82) holds for $k = i - 1$, we know that the set $I_1 \cap I_2 \cap \cdots \cap I_{i-1}$ is a closed interval. Thus, the set $I_1 \cap I_2 \cap \cdots \cap I_{i-1}$ has the form $\{x \in \mathbb{R} \mid a \leq x \leq b\}$ for some elements a and b of $\mathbb{R} \cup \{-\infty, \infty\}$ [88]. In other words, there exist some elements a and b of $\mathbb{R} \cup \{-\infty, \infty\}$ such that $I_1 \cap I_2 \cap \cdots \cap I_{i-1} = \{x \in \mathbb{R} \mid a \leq x \leq b\}$. Denote this a and b as c and d, respectively. Then, $I_1 \cap I_2 \cap \cdots \cap I_{i-1} = \{x \in \mathbb{R} \mid c \leq x \leq d\}$.

The set I_i is a closed interval. Thus, I_i has the form $\{x \in \mathbb{R} \mid a \leq x \leq b\}$ for some elements a and b of $\mathbb{R} \cup \{-\infty, \infty\}$ [89]. In other words, there exist some elements a and b of $\mathbb{R} \cup \{-\infty, \infty\}$ such that $I_i = \{x \in \mathbb{R} \mid a \leq x \leq b\}$. Denote this a and b as u and v, respectively. Then, $I_i = \{x \in \mathbb{R} \mid u \leq x \leq v\}$.

Now, combining the relations

$$\{x \in \mathbb{R} \mid \max\{c, u\} \leq x \leq \min\{d, v\}\} \subseteq I_1 \cap I_2 \cap \cdots \cap I_i$$

[90] and

$$I_1 \cap I_2 \cap \cdots \cap I_i \subseteq \{x \in \mathbb{R} \mid \max\{c, u\} \leq x \leq \min\{d, v\}\}$$

---

[88] This is because we have defined a closed interval to mean a set which has the form $\{x \in \mathbb{R} \mid a \leq x \leq b\}$ for some elements a and b of $\mathbb{R} \cup \{-\infty, \infty\}$.

[89] This is because we have defined a closed interval to mean a set which has the form $\{x \in \mathbb{R} \mid a \leq x \leq b\}$ for some elements a and b of $\mathbb{R} \cup \{-\infty, \infty\}$.

[90] *Proof.* Let $y \in \{x \in \mathbb{R} \mid \max\{c, u\} \leq x \leq \min\{d, v\}\}$. Then, $y \in \mathbb{R}$ and $\max\{c, u\} \leq y \leq \min\{d, v\}$.

Since $c \leq \max\{c, u\} \leq y$ and $y \leq \min\{d, v\} \leq d$, we have $c \leq y \leq d$, so that $y \in \{x \in \mathbb{R} \mid c \leq x \leq d\} = I_1 \cap I_2 \cap \cdots \cap I_{i-1}$.

[91], we obtain

$$I_1 \cap I_2 \cap \cdots \cap I_i = \{x \in \mathbb{R} \mid \max\{c, u\} \le x \le \min\{d, v\}\}.$$

Thus, $I_1 \cap I_2 \cap \cdots \cap I_i$ is a set of the form $\{x \in \mathbb{R} \mid a \le x \le b\}$ for some elements a and b of $\mathbb{R} \cup \{-\infty, \infty\}$ (namely, for $a = \max\{c, u\}$ and $b = \min\{d, v\}$). In other words, $I_1 \cap I_2 \cap \cdots \cap I_i$ is a closed interval[92]. In other words, (82) holds for $k = i$. This completes the induction step. The induction proof of (82) is thus complete.

Now, applying (82) to $k = n$, we conclude that $I_1 \cap I_2 \cap \cdots \cap I_n$ is a closed interval.

Now forget that we fixed n and I_1, I_2, ..., I_n. We have thus shown that whenever $n \in \mathbb{N}$ and whenever I_1, I_2, ..., I_n are n closed intervals, the set $I_1 \cap I_2 \cap \cdots \cap I_n$ is a closed interval. In other words, we have proven that the intersection of finitely many closed intervals always is a closed interval. This proves Lemma 2.0x. $\qquad \square$

Proof of Lemma 2.0y. We distinguish between three cases:

Case 1: We have $\alpha < 0$.

Case 2: We have $\alpha = 0$.

Case 3: We have $\alpha > 0$.

Since $u \le \max\{c, u\} \le y$ and $y \le \min\{d, v\} \le v$, we have $u \le y \le v$, so that $y \in \{x \in \mathbb{R} \mid u \le x \le v\} = I_i$.

Combining $y \in I_1 \cap I_2 \cap \cdots \cap I_{i-1}$ and $y \in I_i$, we obtain $y \in (I_1 \cap I_2 \cap \cdots \cap I_{i-1}) \cap I_i = I_1 \cap I_2 \cap \cdots \cap I_i$.

Now, forget that we fixed y. We thus have proven that every $y \in \{x \in \mathbb{R} \mid \max\{c, u\} \le x \le \min\{d, v\}\}$ satisfies $y \in I_1 \cap I_2 \cap \cdots \cap I_i$. In other words, $\{x \in \mathbb{R} \mid \max\{c, u\} \le x \le \min\{d, v\}\} \subseteq I_1 \cap I_2 \cap \cdots \cap I_i$, qed.

[91]*Proof.* Let $y \in I_1 \cap I_2 \cap \cdots \cap I_i$. Then,

$$y \in I_1 \cap I_2 \cap \cdots \cap I_i \subseteq I_1 \cap I_2 \cap \cdots \cap I_{i-1} = \{x \in \mathbb{R} \mid c \le x \le d\}.$$

Hence, $y \in \mathbb{R}$ and $c \le y \le d$. Also,

$$y \in I_1 \cap I_2 \cap \cdots \cap I_i \subseteq I_i = \{x \in \mathbb{R} \mid u \le x \le v\},$$

so that $u \le y \le v$.

Whenever α, β, γ are three reals satisfying $\alpha \le \gamma$ and $\beta \le \gamma$, we have $\max\{\alpha, \beta\} \le \gamma$. Applied to $\alpha = c$, $\beta = u$ and $\gamma = y$, this yields $\max\{c, u\} \le y$ (since $c \le y$ and $u \le y$).

Whenever α, β, γ are three reals satisfying $\alpha \le \beta$ and $\alpha \le \gamma$, we have $\alpha \le \min\{\beta, \gamma\}$. Applied to $\alpha = y$, $\beta = d$ and $\gamma = v$, this yields $y \le \min\{d, v\}$ (since $y \le d$ and $y \le v$).

Since $\max\{c, u\} \le y \le \min\{d, v\}$, we have $y \in \{x \in \mathbb{R} \mid \max\{c, u\} \le x \le \min\{d, v\}\}$.

Now, forget that we fixed y. We thus have proven that every $y \in I_1 \cap I_2 \cap \cdots \cap I_i$ satisfies $y \in \{x \in \mathbb{R} \mid \max\{c, u\} \le x \le \min\{d, v\}\}$. In other words, $I_1 \cap I_2 \cap \cdots \cap I_i \subseteq \{x \in \mathbb{R} \mid \max\{c, u\} \le x \le \min\{d, v\}\}$, qed.

[92]This is because we have defined a closed interval to mean a set which has the form $\{x \in \mathbb{R} \mid a \le x \le b\}$ for some elements a and b of $\mathbb{R} \cup \{-\infty, \infty\}$.

Let us first consider Case 1. In this case, $\alpha < 0$. Hence, for every $x \in \mathbb{R}$, the assertion $\alpha x \leq \beta$ is equivalent to $x \geq \dfrac{\beta}{\alpha}$ (because dividing an inequality by a negative real number reverses the sign of this inequality). Thus, for every $x \in \mathbb{R}$, we have the following equivalence of assertions:

$$(\alpha x \leq \beta) \Longleftrightarrow \left(x \geq \frac{\beta}{\alpha} \right) \Longleftrightarrow \left(\frac{\beta}{\alpha} \leq x \right) \Longleftrightarrow \left(\frac{\beta}{\alpha} \leq x \leq \infty \right).$$

Hence,

$$\{ x \in \mathbb{R} \mid \alpha x \leq \beta \} = \left\{ x \in \mathbb{R} \mid \frac{\beta}{\alpha} \leq x \leq \infty \right\}.$$

Thus, $\{ x \in \mathbb{R} \mid \alpha x \leq \beta \}$ is a set of the form $\{ x \in \mathbb{R} \mid a \leq x \leq b \}$ for some elements a and b of $\mathbb{R} \cup \{-\infty, \infty\}$ (namely, for $a = \dfrac{\beta}{\alpha}$ and $b = \infty$). In other words, $\{ x \in \mathbb{R} \mid \alpha x \leq \beta \}$ is a closed interval[93]. This proves Lemma 2.0y in Case 1.

Let us now consider Case 2. In this case, $\alpha = 0$. We must be in one of the following two subcases:

Subcase 2.1: We have $\beta \geq 0$.

Subcase 2.2: We have $\beta < 0$.

First, let us consider Subcase 2.1. In this subcase, $\beta \geq 0$. Thus, every $x \in \mathbb{R}$ satisfies $\alpha x \leq \beta$ (since every $x \in \mathbb{R}$ satisfies $\underbrace{\alpha}_{=0} x = 0 \leq \beta$). Hence, $\{ x \in \mathbb{R} \mid \alpha x \leq \beta \} = \mathbb{R}$. Thus, $\{ x \in \mathbb{R} \mid \alpha x \leq \beta \}$ is a closed interval (since $\mathbb{R}$ is a closed interval). This proves Lemma 2.0y in Subcase 2.1.

Next, let us consider Subcase 2.2. In this subcase, $\beta < 0$. Thus, every $x \in \mathbb{R}$ satisfies $\underbrace{\alpha}_{=0} x = 0 > \beta$. Hence, no $x \in \mathbb{R}$ satisfies $\alpha x \leq \beta$. Therefore, $\{ x \in \mathbb{R} \mid \alpha x \leq \beta \} = \varnothing$. Thus, $\{ x \in \mathbb{R} \mid \alpha x \leq \beta \}$ is a closed interval (since $\varnothing$ is a closed interval). This proves Lemma 2.0y in Subcase 2.2.

Since Lemma 2.0y is proven in both Subcases 2.1 and 2.2, it follows that Lemma 2.0y always holds in Case 2 (because Subcases 2.1 and 2.2 cover all possibilities within Case 2).

Finally, let us consider Case 3. In this case, $\alpha > 0$. Hence, for every $x \in \mathbb{R}$, the assertion $\alpha x \leq \beta$ is equivalent to $x \leq \dfrac{\beta}{\alpha}$ (because dividing an inequality by a positive real number leaves the sign of this inequality invariant). Thus, for every $x \in \mathbb{R}$, we have the following equivalence of assertions:

$$(\alpha x \leq \beta) \Longleftrightarrow \left(x \leq \frac{\beta}{\alpha} \right) \Longleftrightarrow \left(-\infty \leq x \leq \frac{\beta}{\alpha} \right).$$

---

[93]This is because we have defined a closed interval to mean a set which has the form $\{ x \in \mathbb{R} \mid a \leq x \leq b \}$ for some elements a and b of $\mathbb{R} \cup \{-\infty, \infty\}$.

Hence,

$$\{x \in \mathbb{R} \mid \alpha x \le \beta\} = \left\{x \in \mathbb{R} \mid -\infty \le x \le \frac{\beta}{\alpha}\right\}.$$

Thus, $\{x \in \mathbb{R} \mid \alpha x \le \beta\}$ is a set of the form $\{x \in \mathbb{R} \mid a \le x \le b\}$ for some elements a and b of $\mathbb{R} \cup \{-\infty, \infty\}$ (namely, for $a = -\infty$ and $b = \dfrac{\beta}{\alpha}$). In other words, $\{x \in \mathbb{R} \mid \alpha x \le \beta\}$ is a closed interval[94]. This proves Lemma 2.0y in Case 3.

Thus, Lemma 2.0y is proven in each of Cases 1, 2 and 3. Since these cases cover all possibilities, this yields that Lemma 2.0y is always proven. $\qquad \square$

Proof of Lemma 2.0z. Let $b_1, b_2, \ldots, b_m$ be the m coordinates of the column vector b. Thus, we have

$$\text{(the } i\text{-th coordinate of } b) = b_i \tag{83}$$

for each $i \in \{1, 2, \ldots, m\}$.

Let $a_1, a_2, \ldots, a_m$ be the m rows of the matrix A. Then, for each column vector $x \in \mathbb{R}^n$, and for each $i \in \{1, 2, \ldots, m\}$, we have

$$\text{(the } i\text{-th coordinate of } Ax) = a_i x \tag{84}$$

(by the definition of the product of a matrix with a column vector).

For each $x \in \mathbb{R}^n$, we have the following chain of equivalences:

$(Ax \le b)$
$\Longleftrightarrow$ (every $i \in \{1, 2, \ldots, m\}$ satisfies

$$\underbrace{\text{(the } i\text{-th coordinate of } Ax)}_{\substack{=a_i x \\ \text{(by (84))}}} \le \underbrace{\text{(the } i\text{-th coordinate of } b)}_{\substack{=b_i \\ \text{(by (83))}}}$$

(by the definition of $Ax \le b$)
$\Longleftrightarrow$ (every $i \in \{1, 2, \ldots, m\}$ satisfies $a_i x \le b_i$)
$\Longleftrightarrow$ ($a_i x \le b_i$ for every $i \in \{1, 2, \ldots, m\}$).

Hence,

$$\{x \in \mathbb{R}^n \mid Ax \le b\}$$
$$= \{x \in \mathbb{R}^n \mid a_i x \le b_i \text{ for every } i \in \{1, 2, \ldots, m\}\}.$$

---

[94]This is because we have defined a closed interval to mean a set which has the form $\{x \in \mathbb{R} \mid a \le x \le b\}$ for some elements a and b of $\mathbb{R} \cup \{-\infty, \infty\}$.

Thus,

$$P = \{x \in \mathbb{R}^n \mid Ax \le b\}$$
$$= \{x \in \mathbb{R}^n \mid a_i x \le b_i \text{ for every } i \in \{1, 2, \ldots, m\}\}.$$

For every $i \in \{1, 2, \ldots, m\}$, the set $\{\mu \in \mathbb{R} \mid a_i c\mu \le b_i - a_i z\}$ is a closed interval (by Lemma 2.0y, applied to $\alpha = a_i c$ and $\beta = b_i - a_i z$).

For any element $\mu \in \mathbb{R}$, we have the following equivalence of assertions:

$$(z + \mu c \in P) \iff \left(\underbrace{a_i (z + \mu c)}_{=a_i z + a_i c\mu} \le b_i \text{ for every } i \in \{1, 2, \ldots, m\} \right)$$

$$(\text{since } P = \{x \in \mathbb{R}^n \mid a_i x \le b_i \text{ for every } i \in \{1, 2, \ldots, m\}\})$$

$$\iff \left(\underbrace{a_i z + a_i c\mu \le b_i}_{\substack{\text{this is equivalent to} \\ a_i c\mu \le b_i - a_i z}} \text{ for every } i \in \{1, 2, \ldots, m\} \right)$$

$$\iff (a_i c\mu \le b_i - a_i z \text{ for every } i \in \{1, 2, \ldots, m\}).$$

Thus,

$$\{\mu \in \mathbb{R} \mid z + \mu c \in P\}$$
$$= \{\mu \in \mathbb{R} \mid a_i c\mu \le b_i - a_i z \text{ for every } i \in \{1, 2, \ldots, m\}\}$$
$$= \bigcap_{i \in \{1,2,\ldots,m\}} \{\mu \in \mathbb{R} \mid a_i c\mu \le b_i - a_i z\}.$$

This shows that $\{\mu \in \mathbb{R} \mid z + \mu c \in P\}$ is an intersection of finitely many closed intervals (since $\{\mu \in \mathbb{R} \mid a_i c\mu \le b_i - a_i z\}$ is a closed interval for every $i \in \{1, 2, \ldots, m\}$). Thus, $\{\mu \in \mathbb{R} \mid z + \mu c \in P\}$ is a closed interval (since Lemma 2.0x says that the intersection of finitely many closed intervals always is a closed interval). This proves Lemma 2.0z. $\square$

16.5. Proofs for Section 9

Proof of Lemma 2.2a. **(a)** Let $v \in \mathbb{R}^n$ be a column vector satisfying $v > 0$. We must prove that all coordinates of the column vector v are positive.

We have $v > 0$. In other words, every $i \in \{1, 2, \ldots, n\}$ satisfies

$$(\text{the } i\text{-th coordinate of } v) > (\text{the } i\text{-th coordinate of } 0)$$

(by the definition of $v > 0$). Hence, every $i \in \{1, 2, \ldots, n\}$ satisfies

$$(\text{the } i\text{-th coordinate of } v) > (\text{the } i\text{-th coordinate of } 0) = 0.$$

In other words, for every $i \in \{1, 2, \ldots, n\}$, the i-th coordinate of v is positive. In other words, the coordinates of the column vector v are positive. This proves Lemma 2.2a **(a)**.

(b) The proof of Lemma 2.2a **(b)** is completely identical to the proof of Lemma 2.2a **(a)** given above (except that "$\mathbb{R}^n$" must be replaced by "$(\mathbb{R}^n)^*$", and the word "column" must be replaced by "row"). □

Proof of Lemma 2.2b. The row vector $x \in (\mathbb{R}^n)^*$ satisfies $x \geq 0$. Thus, Lemma 2.0t **(b)** (applied to $v = x$) shows that the coordinates of the row vector x are nonnegative. Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the coordinates of the row vector x. Then, $\lambda_1, \lambda_2, \ldots, \lambda_n$ are nonnegative (since the coordinates of the row vector x are nonnegative).

The column vector $y \in \mathbb{R}^n$ satisfies $y > 0$. Thus, Lemma 2.2a **(a)** (applied to $v = y$) shows that the coordinates of the column vector y are positive. Let $\mu_1, \mu_2, \ldots, \mu_n$ be the coordinates of the column vector y. Then, $\mu_1, \mu_2, \ldots, \mu_n$ are positive (since the coordinates of the column vector y are positive).

By the definition of the product of a row vector with a column vector, we have $xy = \sum\limits_{i=1}^{n} \lambda_i \mu_i$ (since the coordinates of the row vector x are $\lambda_1, \lambda_2, \ldots, \lambda_n$, while the coordinates of the column vector y are $\mu_1, \mu_2, \ldots, \mu_n$). But for every $i \in \{1, 2, \ldots, n\}$, the real $\lambda_i \mu_i$ is nonnegative (since λ_i is nonnegative[95] and μ_i is positive[96]). In other words, for every $i \in \{1, 2, \ldots, n\}$, we have $\lambda_i \mu_i \geq 0$. Hence, $\sum\limits_{i=1}^{n} \underbrace{\lambda_i \mu_i}_{\geq 0} \geq \sum\limits_{i=1}^{n} 0 = 0$. Hence, $xy = \sum\limits_{i=1}^{n} \lambda_i \mu_i \geq 0$.

Assume (for the sake of contradiction) that $xy = 0$.

Recall that, for every $i \in \{1, 2, \ldots, n\}$, the real $\lambda_i \mu_i$ is nonnegative. Thus, the sum $\sum\limits_{i=1}^{n} \lambda_i \mu_i$ is a sum of nonnegative reals. Since this sum is 0 (because $\sum\limits_{i=1}^{n} \lambda_i \mu_i = 0$), this yields that

$$\text{every } i \in \{1, 2, \ldots, n\} \text{ satisfies } \lambda_i \mu_i = 0 \tag{85}$$

(because if a sum of nonnegative reals is 0, then each of these reals must equal 0). Thus, every $i \in \{1, 2, \ldots, n\}$ satisfies $\lambda_i = 0$ [97]. In other words, all the numbers $\lambda_1, \lambda_2, \ldots, \lambda_n$ equal zero. Since $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the coordinates of the row vector x, this yields that all the coordinates of the row vector x equal zero. Hence, $x = 0$. This contradicts the fact that $x \neq 0$ (since x is nonzero). This contradiction shows that our assumption (that $xy = 0$) was false. Hence,

---

[95]because $\lambda_1, \lambda_2, \ldots, \lambda_n$ are nonnegative

[96]because $\mu_1, \mu_2, \ldots, \mu_n$ are positive

[97]*Proof.* Let $i \in \{1, 2, \ldots, n\}$. Recall that $\mu_1, \mu_2, \ldots, \mu_n$ are positive; thus, μ_i is positive. Hence, $\mu_i \neq 0$. But (85) yields $\lambda_i \mu_i = 0$. We can divide this equality by μ_i (since $\mu_i \neq 0$), and thus obtain $\lambda_i = 0$. Qed.

we cannot have $xy = 0$. Thus, we have $xy \neq 0$. Combining this with $xy \geq 0$, we obtain $xy > 0$. This proves Lemma 2.2b. $\qquad \square$

Proof of Lemma 2.2c. Lemma 2.2c is proven in the same way as Lemma 2.2b, except that the roles of the vectors x and y are partly reversed. (The details are left to the reader.) $\qquad \square$

Proof of Lemma 2.2d. Assume the contrary. Thus, $x \neq 0$. Hence, x is nonzero. Lemma 2.2b thus yields $xy > 0$. This contradicts $xy = 0$. This contradiction shows that our assumption was false. This proves Lemma 2.2d. $\qquad \square$

16.6. Proofs for Section 11

Proof of Lemma 2.5h. We have the following chain of logical equivalences:

$$\left(\begin{pmatrix} x \\ y \end{pmatrix} \geq 0 \right)$$

$$\Longleftrightarrow \left(\text{all coordinates of the vector } \begin{pmatrix} x \\ y \end{pmatrix} \text{ are } \geq 0 \right)$$

$$\Longleftrightarrow \left(\text{all coordinates of the vectors } x \text{ and } y \text{ are } \geq 0 \right)$$

$$\left(\begin{array}{c} \text{since the coordinates of the vector } \begin{pmatrix} x \\ y \end{pmatrix} \text{ are} \\ \text{precisely the coordinates of the vectors } x \text{ and } y \end{array} \right)$$

$$\Longleftrightarrow \underbrace{(\text{all coordinates of the vector } x \text{ are } \geq 0)}_{\Longleftrightarrow \ (x \geq 0)}$$

$$\wedge \underbrace{(\text{all coordinates of the vector } y \text{ are } \geq 0)}_{\Longleftrightarrow \ (y \geq 0)}$$

$$\Longleftrightarrow (x \geq 0) \wedge (y \geq 0).$$

In other words, $\begin{pmatrix} x \\ y \end{pmatrix} \geq 0$ holds if and only if $(x \geq 0$ and $y \geq 0)$. This yields both parts **(a)** and **(b)** of Lemma 2.5h. $\qquad \square$

Proof of Lemma 2.5i. Let $x_1, x_2, \ldots, x_n$ be the n coordinates of the vector x. Let $y_1, y_2, \ldots, y_n$ be the n coordinates of the vector y. Then, the $2n$ coordinates of the vector $\begin{pmatrix} x \\ -x \end{pmatrix}$ are $x_1, x_2, \ldots, x_n, -x_1, -x_2, \ldots, -x_n$, whereas the $2n$ coordinates of the vector $\begin{pmatrix} y \\ -y \end{pmatrix}$ are $y_1, y_2, \ldots, y_n, -y_1, -y_2, \ldots, -y_n$. Hence, the inequality $\begin{pmatrix} x \\ -x \end{pmatrix} \geq \begin{pmatrix} y \\ -y \end{pmatrix}$ says that each of the $2n$ coordinates $x_1, x_2, \ldots, x_n, -x_1, -x_2, \ldots, -x_n$

is $\geq$ to the corresponding coordinate among $y_1, y_2, \ldots, y_n, -y_1, -y_2, \ldots, -y_n$. In other words,

$$x_i \geq y_i \qquad \text{for each } i \in \{1, 2, \ldots, n\}, \tag{86}$$

and

$$-x_i \geq -y_i \qquad \text{for each } i \in \{1, 2, \ldots, n\}. \tag{87}$$

But (87) rewrites as

$$x_i \leq y_i \qquad \text{for each } i \in \{1, 2, \ldots, n\}.$$

Combining this with (86), we obtain $x_i = y_i$ for each $i \in \{1, 2, \ldots, n\}$. Thus, $x = y$. This proves Lemma 2.5i. $\qquad \square$

17. Appendix: Old (and ugly) proofs of Theorems 2.1c and 2.5c

This section gives alternative proofs for Theorem 2.1c and 2.5c. These are the proofs I have found myself, before I became aware of the neat proofs given in Section 7; they are clumsy and long-winded, but maybe there is something of interest in them (although I don't know what that would be).

We shall prove Theorem 2.1c first, and then derive Theorem 2.5c from it. But first of all, we will need a basic fact (which is often used as an exercise in linear algebra courses):

> **Lemma 2.1d.** Let k be an infinite field, and let E be a k-vector space. Let x_1, x_2, ..., x_t be vectors in E. Assume that every $i \in \{1, 2, \ldots, t\}$ satisfies $x_i \neq 0$. Then, there exists an $f \in E^*$ such that every $i \in \{1, 2, \ldots, t\}$ satisfies $f(x_i) \neq 0$.

Proof of Lemma 2.1d. Let us prove that for every $j \in \{0, 1, \ldots, t\}$,

there exists some $f \in E^*$ such that every $i \in \{1, 2, \ldots, j\}$ satisfies $f(x_i) \neq 0$. (88)

Proof of (88): We will prove (88) by induction over j:

Induction base: There exists some $f \in E^*$ such that every $i \in \{1, 2, \ldots, 0\}$ satisfies $f(x_i) \neq 0$ (namely, $f = 0$ (because the assertion that every $i \in \{1, 2, \ldots, 0\}$ satisfies $f(x_i) \neq 0$ is vacuously true)).

Induction step: Let $J \in \{1, 2, \ldots, t\}$ be such that (88) holds for $j = J - 1$. We then have to prove that (88) holds for $j = J$.

We have $x_J \neq 0$ (because every $i \in \{1, 2, \ldots, t\}$ satisfies $x_i \neq 0$).

There is a known linear-algebraic fact that if v is a vector in a k-vector space F such that $v \neq 0$, then there exists an $h \in F^*$ such that $h(v) \neq 0$. Applying this fact to $F = E$ and $v = x_J$, we see that there exists an $h \in E^*$ such that $h(x_J) \neq 0$ (since $x_J \neq 0$). Consider this h.

Since (88) holds for $j = J - 1$, there exists some $f \in E^*$ such that every $i \in \{1, 2, \ldots, J - 1\}$ satisfies $f(x_i) \neq 0$. Denote this f by g. Then, every $i \in \{1, 2, \ldots, J - 1\}$ satisfies $g(x_i) \neq 0$.

Let r be the element $\begin{cases} \dfrac{-h(x_J)}{g(x_J)}, & \text{if } g(x_J) \neq 0; \\ 0, & \text{if } g(x_J) = 0 \end{cases}$ of k. Let M be the subset $\left\{ \dfrac{-h(x_1)}{g(x_1)}, \dfrac{-h(x_2)}{g(x_2)}, \ldots, \dfrac{-h(x_{J-1})}{g(x_{J-1})} \right\} \cup \{r\}$ of k. Then,

$$
\begin{aligned}
|M| &= \left| \left\{ \frac{-h(x_1)}{g(x_1)}, \frac{-h(x_2)}{g(x_2)}, \ldots, \frac{-h(x_{J-1})}{g(x_{J-1})} \right\} \cup \{r\} \right| \\
&\leq \underbrace{\left| \left\{ \frac{-h(x_1)}{g(x_1)}, \frac{-h(x_2)}{g(x_2)}, \ldots, \frac{-h(x_{J-1})}{g(x_{J-1})} \right\} \right|}_{\leq J-1} + \underbrace{|\{r\}|}_{=1} \leq J - 1 + 1 = J,
\end{aligned}
$$

so that we cannot have $k \subseteq M$ (because if we had $k \subseteq M$, then we would have $|k| \leq |M| \leq J$, which contradicts the fact that k is infinite). As a consequence, there exists an $s \in k$ such that $s \notin M$. Consider this s.

Now, every $i \in \{1, 2, \ldots, J\}$ satisfies $(sg + h)(x_i) \neq 0$. [98] Thus, there exists some $f \in E^*$ such that every $i \in \{1, 2, \ldots, J\}$ satisfies $f(x_i) \neq 0$ (namely, $f = sg + h$). In other words, (88) holds for $j = J$. This completes the induction step. The induction proof of (88) is thus complete.

Applying (88) to $j = t$, we obtain that there exists an $f \in E^*$ such that every $i \in \{1, 2, \ldots, t\}$ satisfies $f(x_i) \neq 0$. This proves Lemma 2.1d. $\qquad\square$

Second proof of Theorem 2.1c. We will prove Theorem 2.1c by induction over $\dim E$.

Induction base: In the case when $\dim E = 0$, Theorem 2.1c is obvious.[99] This completes the induction base.

Induction step: Let n be a positive integer. Assume that Theorem 2.1c holds whenever $\dim E = n - 1$. We will now prove that Theorem 2.1c holds whenever $\dim E = n$.

So, let G be an $\mathbb{R}$-vector space satisfying $\dim G = n$. Let C be a polytope in

---

[98] *Proof.* Assume the opposite. Then, there exists some $i \in \{1, 2, \ldots, J\}$ such that $(sg + h)(x_i) = 0$. Consider this i. Since $i \in \{1, 2, \ldots, J\}$, we must be in one of the following two cases:

    Case 1: We have $i \in \{1, 2, \ldots, J - 1\}$.

    Case 2: We have $i = J$.

Let us first consider Case 1. In this case, $i \in \{1, 2, \ldots, J - 1\}$.

Our i satisfies $(sg + h)(x_i) = 0$, so that $0 = (sg + h)(x_i) = sg(x_i) + h(x_i)$, thus $sg(x_i) = -h(x_i)$. Since $g(x_i) \neq 0$, we can divide this by $g(x_i)$ and obtain

$$s \in \frac{-h(x_i)}{g(x_i)} \in \left\{ \frac{-h(x_1)}{g(x_1)}, \frac{-h(x_2)}{g(x_2)}, \ldots, \frac{-h(x_{J-1})}{g(x_{J-1})} \right\}$$
$$\subseteq \left\{ \frac{-h(x_1)}{g(x_1)}, \frac{-h(x_2)}{g(x_2)}, \ldots, \frac{-h(x_{J-1})}{g(x_{J-1})} \right\} \cup \{r\} = M,$$

contradicting $s \notin M$. Thus, we have obtained a contradiction in Case 1.

Let us now consider Case 2. In this case, $i = J$. Thus, $(sg + h)(x_i) = (sg + h)(x_J) = sg(x_J) + h(x_J)$. Compared with $(sg + h)(x_i) = 0$, this becomes $sg(x_J) + h(x_J) = 0$. Thus, $sg(x_J) = \underbrace{-h(x_J)}_{\neq 0} \neq -0 = 0$, so that $g(x_J) \neq 0$. Thus, $r = $

$$\begin{cases} \dfrac{-h(x_J)}{g(x_J)}, & \text{if } g(x_J) \neq 0; \\ 0, & \text{if } g(x_J) = 0 \end{cases} = \frac{-h(x_J)}{g(x_J)} \text{ (since } g(x_J) \neq 0).$$

But dividing $sg(x_J) = -h(x_J)$ by $g(x_J)$ (this is allowed since $g(x_J) \neq 0$), we obtain

$$s = \frac{-h(x_J)}{g(x_J)} = r \in \left\{ \frac{-h(x_1)}{g(x_1)}, \frac{-h(x_2)}{g(x_2)}, \ldots, \frac{-h(x_{J-1})}{g(x_{J-1})} \right\} \cup \{r\} = M,$$

contradicting $s \notin M$. Thus, we have obtained a contradiction in Case 2.

Hence, in each of the cases 1 and 2, we have obtained a contradiction. Thus, we always have a contradiction. Hence, our assumption was wrong, qed.

[99] In fact, in the case when $\dim E = 0$, the only polytopes in E are 0 and $\varnothing$. If $C = 0$, Assertion C1 holds and Assertion C2 does not; else, Assertion C2 holds and Assertion C1 does not. Thus, Theorem 2.1c holds in the case when $\dim E = 0$.

G. By the definition of a polytope, this shows that C is the convex hull of a finite set of vectors in G. In other words, there exist some $t \in \mathbb{N}$ and some vectors $x_1, x_2, \ldots, x_t$ in G such that $C = \operatorname{conv.hull} \{x_1, x_2, \ldots, x_t\}$. Consider this t and these $x_1, x_2, \ldots, x_t$.

We will now prove that exactly one of the following assertions holds:

Assertion $C_1 1$: We have $0 \in C$.

Assertion $C_1 2$: There exists an $f \in G^*$ such that every $x \in C$ satisfies $f(x) < 0$.

First, it is clear that the Assertions $C_1 1$ and $C_1 2$ cannot hold at the same time[100]. We will now show that at least one of these assertions holds. Our proof will proceed in several steps:[101]

Step 1: If some $i \in \{1, 2, \ldots, t\}$ satisfies $x_i = 0$, then at least one of Assertions $C_1 1$ and $C_1 2$ holds[102]. Hence, for the rest of this proof, we can WLOG assume that no $i \in \{1, 2, \ldots, t\}$ satisfies $x_i = 0$. Assume this.

Step 2: If $t = 0$, then at least one of Assertions $C_1 1$ and $C_1 2$ holds[103]. Hence, for the rest of this proof, we can WLOG assume that $t \neq 0$. Assume this.

Step 3: Every $i \in \{1, 2, \ldots, t\}$ satisfies $x_i \neq 0$ (since no $i \in \{1, 2, \ldots, t\}$ satisfies $x_i = 0$). Thus, Lemma 2.1d (applied to $k = \mathbb{R}$ and $E = G$) yields that there exists an $f \in G^*$ such that every $i \in \{1, 2, \ldots, t\}$ satisfies $f(x_i) \neq 0$. Denote this f by g. Thus, g is an element of G^* such that every $i \in \{1, 2, \ldots, t\}$ satisfies $g(x_i) \neq 0$. Now, $\operatorname{Im} g = \mathbb{R}$ [104].

Step 4: Now, let $F = \operatorname{Ker} g$. Then, F is an $\mathbb{R}$-vector subspace of E and satisfies $\dim F = n - 1$ [105].

---

[100] *Proof.* Assume the opposite. Then, the Assertions $C_1 1$ and $C_1 2$ hold at the same time. Since Assertion $C_1 2$ holds, there exists an $f \in G^*$ such that every $x \in C$ satisfies $f(x) < 0$. Consider this f. We know that every $x \in C$ satisfies $f(x) < 0$. Since $0 \in C$ (because Assertion $C_1 1$ holds), we can apply this to $x = 0$, and thus obtain $f(0) < 0$. But this contradicts $f(0) = 0$ (which is because f is linear). This contradiction shows that our assumption was wrong, qed.

[101] I have split this proof into several steps for the reader's convenience. These steps, however, are not self-contained; for example, Step 9 involves a case distinction, Step 10 handles its Case 1, and Steps 11 until 18 handle its Case 2.

[102] *Proof.* Assume that some $i \in \{1, 2, \ldots, t\}$ satisfies $x_i = 0$. Consider this i. Then, $0 = x_i \in \operatorname{conv.hull} \{x_1, x_2, \ldots, x_t\} = C$, so that Assertion $C_1 1$ holds. Thus, at least one of Assertions $C_1 1$ and $C_1 2$ holds, qed.

[103] *Proof.* Assume that $t = 0$. Then, $C = \operatorname{conv.hull} \underbrace{\{x_1, x_2, \ldots, x_t\}}_{\substack{= \varnothing \\ (\text{since } t = 0)}} = \operatorname{conv.hull} \varnothing = \varnothing$, so that Assertion $C_1 2$ holds (because it is vacuously true for $f = 0$). Thus, at least one of Assertions $C_1 1$ and $C_1 2$ holds.

[104] *Proof.* Every $i \in \{1, 2, \ldots, t\}$ satisfies $g(x_i) \neq 0$. Applied to $i = 1$, this yields $g(x_1) \neq 0$, so that $g \neq 0$, and thus $\operatorname{Im} g \neq 0$. Hence, $\dim(\operatorname{Im} g) \geq 1$. Since $1 = \dim \mathbb{R}$, this rewrites as $\dim(\operatorname{Im} g) \geq \dim \mathbb{R}$. Combined with $\operatorname{Im} g \subseteq \mathbb{R}$, this yields $\operatorname{Im} g = \mathbb{R}$, qed.

[105] *Proof.* By the homomorphism theorem, $G / (\operatorname{Ker} g) \cong \operatorname{Im} g$. Since $F = \operatorname{Ker} g$, we

---

Step 5: Let I^+ be the set $\{i \in \{1, 2, \ldots, t\} \mid g(x_i) > 0\}$, and let I^- be the set $\{i \in \{1, 2, \ldots, t\} \mid g(x_i) < 0\}$. Clearly, these sets I^+ and I^- are disjoint (since no $i \in \{1, 2, \ldots, t\}$ can satisfy $g(x_i) > 0$ and $g(x_i) < 0$ at the same time).

Since $I^+ = \{i \in \{1, 2, \ldots, t\} \mid g(x_i) > 0\}$, it is clear that

$$\text{every } i \in I^+ \text{ satisfies } g(x_i) > 0. \tag{89}$$

Since $I^- = \{i \in \{1, 2, \ldots, t\} \mid g(x_i) < 0\}$, it is clear that

$$\text{every } j \in I^- \text{ satisfies } g(x_j) < 0. \tag{90}$$

Every $i \in \{1, 2, \ldots, t\}$ satisfies $g(x_i) \neq 0$. Thus,

$$
\begin{aligned}
\{1, 2, \ldots, t\} &= \left\{ i \in \{1, 2, \ldots, t\} \mid \underbrace{g(x_i) \neq 0}_{\substack{\text{this is equivalent to} \\ (g(x_i) > 0 \text{ or } g(x_i) < 0)}} \right\} \\
&= \{i \in \{1, 2, \ldots, t\} \mid g(x_i) > 0 \text{ or } g(x_i) < 0\} \\
&= \underbrace{\{i \in \{1, 2, \ldots, t\} \mid g(x_i) > 0\}}_{=I^+} \cup \underbrace{\{i \in \{1, 2, \ldots, t\} \mid g(x_i) < 0\}}_{=I^-} \\
&= I^+ \cup I^-.
\end{aligned}
$$

Step 6: If $I^- = \varnothing$, then at least one of Assertions $C_1 1$ and $C_1 2$ holds[106]. Hence, for the rest of this proof, we can WLOG assume that $I^- \neq \varnothing$. Assume this.

Step 7: For every $i \in \{1, 2, \ldots, t\}$, let α_i be the element $\dfrac{1}{g(x_i)} \in \mathbb{R}$ (this is well-defined since $g(x_i) \neq 0$).

For every $(i, j) \in I^+ \times I^-$, we have $\alpha_i - \alpha_j \neq 0$ [107], so that $\dfrac{1}{\alpha_i - \alpha_j}$ is well-defined.

---

have $G/F = G/(\operatorname{Ker} g) \cong \operatorname{Im} g = \mathbb{R}$, so that $\dim(G/F) = \dim \mathbb{R} = 1$. Thus, $1 = \dim(G/F) = \underbrace{\dim G}_{=n} - \dim F = n - \dim F$. Hence, $\dim F = n - 1$, qed.

[106] *Proof.* Assume that $I^- = \varnothing$.

Let $i \in \{1, 2, \ldots, t\}$ be arbitrary. Then, $i \in \{1, 2, \ldots, t\} = I^+ \cup \underbrace{I^-}_{=\varnothing} = I^+$, so that $g(x_i) > 0$ (by (89)).

Now, forget that we fixed i. We thus have proven that every $i \in \{1, 2, \ldots, t\}$ satisfies $g(x_i) > 0$. Hence, every $i \in \{1, 2, \ldots, t\}$ satisfies $(-g)(x_i) = -\underbrace{g(x_i)}_{>0} < 0$.

Therefore, Proposition 2.0r (applied to $E = G$, $f = -g$ and $\delta = 0$) shows that every $x \in C$ satisfies $(-g)(x) < 0$ (because $C = \operatorname{conv.hull}\{x_1, x_2, \ldots, x_t\}$). Thus, there exists an $f \in G^*$ such that every $x \in C$ satisfies $f(x) < 0$ (namely, $f = -g$). Hence, Assertion $C_1 2$ holds. Thus, at least one of Assertions $C_1 1$ and $C_1 2$ holds, qed.

[107] *Proof.* Let $(i, j) \in I^+ \times I^-$. Then, $i \in I^+$ and $j \in I^-$. Thus, $g(x_i) > 0$ (by (89)), so

---

Now, let S denote the set

$$\left\{ \frac{1}{\alpha_i - \alpha_j} \left(\alpha_i x_i - \alpha_j x_j \right) \;\mid\; (i,j) \in I^+ \times I^- \right\}.$$

Since every $(i,j) \in I^+ \times I^-$ satisfies $\dfrac{1}{\alpha_i - \alpha_j} \left(\alpha_i x_i - \alpha_j x_j \right) \in \operatorname{Ker} g$ [108], we have

$$\left\{ \frac{1}{\alpha_i - \alpha_j} \left(\alpha_i x_i - \alpha_j x_j \right) \;\mid\; (i,j) \in I^+ \times I^- \right\} \subseteq \operatorname{Ker} g. \text{ Thus,}$$

$$S = \left\{ \frac{1}{\alpha_i - \alpha_j} \left(\alpha_i x_i - \alpha_j x_j \right) \;\mid\; (i,j) \in I^+ \times I^- \right\} \subseteq \operatorname{Ker} g = F.$$

Thus, $\operatorname{conv.hull} S$ is a polytope in F.

Since $S = \left\{ \dfrac{1}{\alpha_i - \alpha_j} \left(\alpha_i x_i - \alpha_j x_j \right) \;\mid\; (i,j) \in I^+ \times I^- \right\}$, we have

$$\frac{1}{\alpha_i - \alpha_j} \left(\alpha_i x_i - \alpha_j x_j \right) \in S \qquad \text{for every } (i,j) \in I^+ \times I^-. \qquad (91)$$

Step 8: Let us now see that $\operatorname{conv.hull} S \subseteq C$:

Every $(i,j) \in I^+ \times I^-$ satisfies $\dfrac{1}{\alpha_i - \alpha_j} \left(\alpha_i x_i - \alpha_j x_j \right) \in C$ [109]. Thus,

$$\left\{ \frac{1}{\alpha_i - \alpha_j} \left(\alpha_i x_i - \alpha_j x_j \right) \;\mid\; (i,j) \in I^+ \times I^- \right\} \subseteq C.$$

---

that $\dfrac{1}{g(x_i)} > 0$. Also, $g(x_j) < 0$ (by (90)), so that $\dfrac{1}{g(x_j)} < 0$. Now, since $\alpha_i = \dfrac{1}{g(x_i)}$ (by the definition of α_i) and $\alpha_j = \dfrac{1}{g(x_j)}$ (by the definition of α_j), we have $\alpha_i - \alpha_j = \underbrace{\dfrac{1}{g(x_i)}}_{>0} - \underbrace{\dfrac{1}{g(x_j)}}_{<0} > 0$, so that $\alpha_i - \alpha_j \neq 0$, qed.

[108] *Proof.* Let $(i,j) \in I^+ \times I^-$. Since $\alpha_i = \dfrac{1}{g(x_i)}$ (by the definition of α_i) and $\alpha_j = \dfrac{1}{g(x_j)}$ (by the definition of α_j), we have

$$g\left(\alpha_i x_i - \alpha_j x_j \right) = g\left(\frac{1}{g(x_i)} x_i - \frac{1}{g(x_j)} x_j \right) = \underbrace{\frac{1}{g(x_i)} g(x_i)}_{=1} - \underbrace{\frac{1}{g(x_j)} g(x_j)}_{=1}$$

$$\text{(since } g \text{ is } \mathbb{R}\text{-linear)}$$

$$= 1 - 1 = 0.$$

Thus, $\alpha_i x_i - \alpha_j x_j \in \operatorname{Ker} g$. Since $\operatorname{Ker} g$ is a k-vector space, this yields $\dfrac{1}{\alpha_i - \alpha_j} \left(\alpha_i x_i - \alpha_j x_j \right) \in \operatorname{Ker} g$, qed.

[109] *Proof.* Let $(i,j) \in I^+ \times I^-$. Then, $i \in I^+$ and $j \in I^-$. Thus, $g(x_i) > 0$ (by (89)), so that $\dfrac{1}{g(x_i)} > 0$. Also, $g(x_j) < 0$ (by (90)), so that $\dfrac{1}{g(x_j)} < 0$.

Thus,

$$S = \left\{ \frac{1}{\alpha_i - \alpha_j} \left(\alpha_i x_i - \alpha_j x_j \right) \mid (i, j) \in I^+ \times I^- \right\} \subseteq C.$$

Hence, C is a convex set containing S as a subset. But since every convex set containing S as a subset must also contain $\text{conv.hull}\, S$ as a subset (because $\text{conv.hull}\, S$ is the smallest convex set containing S as a subset), this yields that C contains $\text{conv.hull}\, S$ as a subset. In other words, $\text{conv.hull}\, S \subseteq C$.

Step 9: Now, since $\dim F = n - 1$, we can apply Theorem 2.1c to F and $\text{conv.hull}\, S$ instead of E and C (because we assumed that Theorem 2.1c holds whenever $\dim E = n - 1$). As a consequence, we obtain that exactly one of the following two assertions holds:

> *Assertion $C_2 1$:* We have $0 \in \text{conv.hull}\, S$.

> *Assertion $C_2 2$:* There exists an $f \in F^*$ such that every $x \in \text{conv.hull}\, S$ satisfies $f(x) < 0$.

Thus, we must be in one of the following two cases:

Case 1: Assertion $C_2 1$ holds.

Case 2: Assertion $C_2 2$ holds.

Step 10: Let us consider Case 1 first. In this case, Assertion $C_2 1$ holds. In other words, $0 \in \text{conv.hull}\, S$. Since $\text{conv.hull}\, S \subseteq C$, this yields $0 \in C$, so that Assertion $C_1 1$ holds. Thus, at least one of Assertions $C_1 1$ and $C_1 2$ holds in Case 1.

Now, since $\alpha_i = \dfrac{1}{g(x_i)}$ (by the definition of α_i) and $\alpha_j = \dfrac{1}{g(x_j)}$ (by the definition of α_j), we have $\alpha_i = \dfrac{1}{g(x_i)} > 0$ and $\alpha_j = \dfrac{1}{g(x_j)} < 0$. Thus, $\underbrace{\alpha_i}_{>0} - \underbrace{\alpha_j}_{<0} > 0$.

Since $\alpha_i > 0$ and $\alpha_i - \alpha_j > 0$, we have $\dfrac{\alpha_i}{\alpha_i - \alpha_j} > 0$. Since $-\alpha_j > 0$ (this is because $\alpha_j < 0$) and $\alpha_i - \alpha_j > 0$, we have $\dfrac{-\alpha_j}{\alpha_i - \alpha_j} > 0$.

But C is convex. Thus, every point on a segment connecting two points of C must lie in C.

Since $\dfrac{\alpha_i}{\alpha_i - \alpha_j} > 0$, $\dfrac{-\alpha_j}{\alpha_i - \alpha_j} > 0$ and $\dfrac{\alpha_i}{\alpha_i - \alpha_j} + \dfrac{-\alpha_j}{\alpha_i - \alpha_j} = \dfrac{\alpha_i}{\alpha_i - \alpha_j} - \dfrac{\alpha_j}{\alpha_i - \alpha_j} = \dfrac{\alpha_i - \alpha_j}{\alpha_i - \alpha_j} = 1$, it is clear that $\dfrac{\alpha_i}{\alpha_i - \alpha_j} x_i + \dfrac{-\alpha_j}{\alpha_i - \alpha_j} x_j$ is a point on the segment connecting x_i and x_j. Since $x_i \in C$ (because $x_i \in \{x_1, x_2, \ldots, x_t\} \subseteq \text{conv.hull}\, \{x_1, x_2, \ldots, x_t\} = C$) and $x_j \in C$ (because $x_j \in \{x_1, x_2, \ldots, x_t\} \subseteq \text{conv.hull}\, \{x_1, x_2, \ldots, x_t\} = C$), this yields that $\dfrac{\alpha_i}{\alpha_i - \alpha_j} x_i + \dfrac{-\alpha_j}{\alpha_i - \alpha_j} x_j$ is a point on a segment connecting two points of C. Thus, $\dfrac{\alpha_i}{\alpha_i - \alpha_j} x_i + \dfrac{-\alpha_j}{\alpha_i - \alpha_j} x_j \in C$ (since every point on a segment connecting two points of C must lie in C). Since $\dfrac{\alpha_i}{\alpha_i - \alpha_j} x_i + \dfrac{-\alpha_j}{\alpha_i - \alpha_j} x_j = \dfrac{1}{\alpha_i - \alpha_j} \left(\alpha_i x_i - \alpha_j x_j \right)$, this rewrites as $\dfrac{1}{\alpha_i - \alpha_j} \left(\alpha_i x_i - \alpha_j x_j \right) \in C$, qed.

Step 11: Now, let us consider Case 2. In this case, Assertion C$_2$2 holds. In other words, there exists an $f \in F^*$ such that every $x \in \text{conv . hull } S$ satisfies $f(x) < 0$. Denote this f by h. Then, every $x \in \text{conv . hull } S$ satisfies $h(x) < 0$.

Step 12:[110] A known fact from linear algebra states that if k is a field, P and R are two k-vector spaces, Q is a k-vector subspace of P, and $\xi : Q \to R$ is a k-linear map, then there exists a k-linear map $\eta : P \to R$ such that $\eta \mid_Q = \xi$. [111] Applying this fact to $k = \mathbb{R}$, $P = G$, $Q = F$, $R = \mathbb{R}$ and $\xi = h$, we conclude that there exists an $\mathbb{R}$-linear map $\eta : G \to \mathbb{R}$ such that $\eta \mid_F = h$. Denote this η by e. Thus, $e \mid_F = h$.

Now, e is an $\mathbb{R}$-linear map $G \to \mathbb{R}$, and

$$\text{every } x \in \text{conv . hull } S \text{ satisfies } e(x) < 0 \tag{92}$$

(since every $x \in \text{conv . hull } S$ satisfies $e(x) = \underbrace{(e \mid_F)}_{=h}(x) = h(x) < 0$).

Step 13: Let k be the element i of I^- minimizing $\alpha_i e(x_i)$ (this is well-defined since I^- is a finite set and $I^- \neq \varnothing$). Then, $k \in I^-$, and we have

$$\alpha_k e(x_k) \leq \alpha_i e(x_i) \qquad \text{for every } i \in I^-. \tag{93}$$

Step 14: Let us now prove that

$$\alpha_k e(x_k) > \alpha_i e(x_i) \qquad \text{for every } i \in I^+. \tag{94}$$

Proof of (94): Let $i \in I^+$. Since $i \in I^+$ and $k \in I^-$, we have $(i, k) \in I^+ \times I^-$. Thus, (91) (applied to (i, k) instead of (i, j)) yields

$$\frac{1}{\alpha_i - \alpha_k}(\alpha_i x_i - \alpha_k x_k) \in S \subseteq \text{conv . hull } S.$$

Hence, (92) (applied to $x = \dfrac{1}{\alpha_i - \alpha_k}(\alpha_i x_i - \alpha_k x_k)$) yields $e\left(\dfrac{1}{\alpha_i - \alpha_k}(\alpha_i x_i - \alpha_k x_k)\right) < 0$. Since $e\left(\dfrac{1}{\alpha_i - \alpha_k}(\alpha_i x_i - \alpha_k x_k)\right) = \dfrac{1}{\alpha_i - \alpha_k}e(\alpha_i x_i - \alpha_k x_k)$ (because e is $\mathbb{R}$-linear), this rewrites as

$$\frac{1}{\alpha_i - \alpha_k}e(\alpha_i x_i - \alpha_k x_k) < 0. \tag{95}$$

Since $i \in I^+$, we have $g(x_i) > 0$ (by (89)). Since $k \in I^-$, we have $g(x_k) < 0$ (by (90), applied to $j = k$). From $g(x_i) > 0$ and $g(x_k) < 0$, it follows that $g(x_i) g(x_k) < 0$.

---

[110]This is no longer Step 11, but we are still in Case 2.

[111]Advanced algebraists tend to state this fact in the following equivalent form: Every k-vector space is an injective k-module.

By the definition of α_i, we have $\alpha_i = \dfrac{1}{g(x_i)} > 0$ (since $g(x_i) > 0$). By the definition of α_k, we have $\alpha_k = \dfrac{1}{g(x_k)} < 0$ (since $g(x_k) < 0$). From $\alpha_i > 0$ and $\alpha_k < 0$, it follows that $\underbrace{\alpha_i}_{>0} - \underbrace{\alpha_k}_{<0} > 0$. Thus, we can multiply (98) with $\alpha_i - \alpha_k$, and obtain $e(\alpha_i x_i - \alpha_k x_k) < 0$. Since

$$e(\alpha_i x_i - \alpha_k x_k) = \alpha_i e(x_i) - \alpha_k e(x_k) \qquad \text{(since } e \text{ is } \mathbb{R}\text{-linear)} ,$$

this rewrites as $\alpha_i e(x_i) - \alpha_k e(x_k) < 0$. In other words, $\alpha_k e(x_k) > \alpha_i e(x_i)$. This proves (94).

Step 15: Let us slightly improve (94): It is easy to see that there exists a positive real δ such that

$$\text{every } i \in I^+ \text{ satisfies } \alpha_k(e(x_k) + \delta) > \alpha_i e(x_i). \tag{96}$$

[112] Consider this δ.

Step 16: We are now going to prove that

$$\text{every } i \in \{1, 2, \ldots, t\} \text{ satisfies } ((e(x_k) + \delta) g - g(x_k) e)(x_i) < 0. \tag{98}$$

Proof of (98): Let $i \in \{1, 2, \ldots, t\}$. Then, $i \in \{1, 2, \ldots, t\} = I^+ \cup I^-$. Hence, either $i \in I^+$ or $i \in I^-$. We thus must be in one of the following two subcases:

Subcase 2.1: We have $i \in I^+$.

---

[112] *Proof.* Every $i \in I^+$ satisfies $\alpha_k e(x_k) > \alpha_i e(x_i)$ (by (94)) and thus $\alpha_k e(x_k) - \alpha_i e(x_i) > 0$. In other words, $\{\alpha_k e(x_k) - \alpha_i e(x_i) \mid i \in I^+\}$ is a set of positive reals. Since this set is finite (because I^+ is finite), it is thus bounded from below by a positive real. In other words, there exists a positive real ε such that

$$\text{every } i \in I^+ \text{ satisfies } \alpha_k e(x_k) - \alpha_i e(x_i) \geq \varepsilon. \tag{97}$$

Consider this ε.

Let $\delta = \dfrac{-\varepsilon}{2\alpha_k}$. Then, $\alpha_k \delta = \dfrac{-\varepsilon}{2}$.

Since $k \in I^-$, we have $g(x_k) < 0$ (by (90), applied to $j = k$). By the definition of α_k, we have $\alpha_k = \dfrac{1}{g(x_k)} < 0$ (since $g(x_k) < 0$). Now, every $i \in I^+$ satisfies

$$\underbrace{\alpha_k(e(x_k) + \delta)}_{= \alpha_k e(x_k) + \alpha_k \delta} - \alpha_i e(x_i) = \alpha_k e(x_k) + \alpha_k \delta - \alpha_i e(x_i)$$

$$= \underbrace{\alpha_k e(x_k) - \alpha_i e(x_i)}_{\substack{\geq \varepsilon \\ \text{(by (97))}}} + \underbrace{\alpha_k \delta}_{= \frac{-\varepsilon}{2}} \geq \varepsilon + \frac{-\varepsilon}{2} = \frac{\varepsilon}{2} > 0$$

(since $\varepsilon > 0$). In other words, every $i \in I^+$ satisfies $\alpha_k(e(x_k) + \delta) > \alpha_i e(x_i)$.

We thus have proven that there exists a positive real δ such that every $i \in I^+$ satisfies $\alpha_k(e(x_k) + \delta) > \alpha_i e(x_i)$. Qed.

Subcase 2.2: We have $i \in I^-$.

Let us first consider Subcase 2.1. In this subcase, $i \in I^+$. Thus, $g(x_i) > 0$ (by (89)). Since $k \in I^-$, we have $g(x_k) < 0$ (by (90), applied to $j = k$). From $g(x_i) > 0$ and $g(x_k) < 0$, it follows that $g(x_i) g(x_k) < 0$.

By the definition of α_i, we have $\alpha_i = \dfrac{1}{g(x_i)} > 0$ (since $g(x_i) > 0$). By the definition of α_k, we have $\alpha_k = \dfrac{1}{g(x_k)} < 0$ (since $g(x_k) < 0$). From (96), we have $\alpha_k (e(x_k) + \delta) > \alpha_i e(x_i)$. Since $\alpha_i = \dfrac{1}{g(x_i)}$ and $\alpha_k = \dfrac{1}{g(x_k)}$, this rewrites as $\dfrac{1}{g(x_k)} (e(x_k) + \delta) > \dfrac{1}{g(x_i)} e(x_i)$. Multiplied by $g(x_i) g(x_k)$, this becomes $g(x_i)(e(x_k) + \delta) < g(x_k) e(x_i)$ (here, the sign has been switched since $g(x_i) g(x_k) < 0$). Thus,

$$((e(x_k) + \delta) g - g(x_k) e)(x_i) = \underbrace{(e(x_k) + \delta) g(x_i)}_{=g(x_i)(e(x_k)+\delta)<g(x_k)e(x_i)} - g(x_k) e(x_i)$$
$$< g(x_k) e(x_i) - g(x_k) e(x_i) = 0.$$

Thus, (98) is proven in Subcase 2.1.

Let us now consider Subcase 2.2. In this subcase, $i \in I^-$. Hence, (93) yields $\alpha_k e(x_k) \leq \alpha_i e(x_i)$.

Since $i \in I^-$, we have $g(x_i) < 0$ (by (90), applied to $j = i$). Since $k \in I^-$, we have $g(x_k) < 0$ (by (90), applied to $j = k$). From $g(x_i) < 0$ and $g(x_k) < 0$, it follows that $g(x_i) g(x_k) > 0$.

Since $\alpha_i = \dfrac{1}{g(x_i)}$ (by the definition of α_i) and $\alpha_k = \dfrac{1}{g(x_k)}$ (by the definition of α_k), the inequality $\alpha_k e(x_k) \leq \alpha_i e(x_i)$ (proven above) rewrites as $\dfrac{1}{g(x_k)} e(x_k) \leq \dfrac{1}{g(x_i)} e(x_i)$. Multiplied by $g(x_i) g(x_k)$, this becomes $g(x_i) e(x_k) \leq g(x_k) e(x_i)$ (here, the sign has not been switched since $g(x_i) g(x_k) > 0$). Thus,

$$((e(x_k) + \delta) g - g(x_k) e)(x_i) = \underbrace{(e(x_k) + \delta) g(x_i)}_{=e(x_k)g(x_i)+\delta g(x_i)} - g(x_k) e(x_i)$$
$$= \underbrace{e(x_k) g(x_i)}_{\substack{=g(x_i)e(x_k) \leq g(x_k)e(x_i)}} + \underbrace{\delta g(x_i)}_{\substack{<0 \\ (\text{since } \delta > 0 \\ \text{and } g(x_i) < 0)}} - g(x_k) e(x_i) < g(x_k) e(x_i) - g(x_k) e(x_i) = 0.$$

Thus, (98) is proven in Subcase 2.2.

We have thus proven (98) in each of the Subcases 2.1 and 2.2. Since these subcases cover all possibilities, this yields that (98) always holds. This completes the proof of (98).

Step 17: Now,

$$\text{every } x \in C \text{ satisfies } \left(\left(e\left(x_k \right) + \delta \right) g - g\left(x_k \right) e \right)(x) < 0. \tag{99}$$

Proof of (99): We know that every $i \in \{1, 2, \ldots, t\}$ satisfies $\left(\left(e\left(x_k \right) + \delta \right) g - g\left(x_k \right) e \right)(x_i) < 0$ (because of (98)). Thus, Proposition 2.0r (applied to G, $\left(e\left(x_k \right) + \delta \right) g - g\left(x_k \right) e$ and 0 instead of E, f and δ) shows that every $x \in C$ satisfies $\left(\left(e\left(x_k \right) + \delta \right) g - g\left(x_k \right) e \right)(x) < 0$. This proves (99).

Step 18: Due to (99), there exists an $f \in G^*$ such that every $x \in C$ satisfies $f(x) < 0$ (namely, $f = \left(e\left(x_k \right) + \delta \right) g - g\left(x_k \right) e$). In other words, Assertion $C_1 2$ holds. Thus, at least one of Assertions $C_1 1$ and $C_1 2$ holds in Case 2.

We have thus proven that, in each of the Cases 1 and 2, at least one of Assertions $C_1 1$ and $C_1 2$ holds. Since these cases cover all possibilities, this yields that, in every situation, at least one of Assertions $C_1 1$ and $C_1 2$ holds.

Altogether, we have shown the following two claims:

- Assertions $C_1 1$ and $C_1 2$ cannot hold at the same time.

- At least one of Assertions $C_1 1$ and $C_1 2$ holds.

Combining these claims, we conclude that exactly one of Assertions $C_1 1$ and $C_1 2$ holds.

Now, forget that we fixed G and C. We have thus proven the following result:

> **Result 1:** If G is an $\mathbb{R}$-vector space satisfying $\dim G = n$, and C is a polytope in G, then exactly one of the following assertions holds:
>
> *Assertion $C_1 1$:* We have $0 \in C$.
>
> *Assertion $C_1 2$:* There exists an $f \in G^*$ such that every $x \in C$ satisfies $f(x) < 0$.

If we rename G as E in Result 1, then this result takes the following form:

> **Result 2:** If E is an $\mathbb{R}$-vector space satisfying $\dim E = n$, and C is a polytope in E, then exactly one of the following assertions holds:
>
> *Assertion $C_3 1$:* We have $0 \in C$.
>
> *Assertion $C_3 2$:* There exists an $f \in E^*$ such that every $x \in C$ satisfies $f(x) < 0$.

Clearly, Result 2 is exactly the statement of Theorem 2.1c in the case when $\dim E = n$. Hence, we have proven that Theorem 2.1c holds in the case when $\dim E = n$. This completes the induction step, and thus the induction proof of Theorem 2.1c is complete. □

Second proof of Theorem 2.5c. We will prove Theorem 2.5c by strong induction over $|S|$:

Induction step: Let n be a nonnegative integer. Assume that Theorem 2.5c holds whenever $|S| < n$. We will now prove that Theorem 2.5c holds whenever $|S| = n$.

So, let E be a finite-dimensional $\mathbb{R}$-vector space. Let S be a finite subset of E such that $|S| = n$. Let $b \in E$. Then, we must prove that exactly one of the following two assertions holds:

> *Assertion $D_1 1$:* We have $b \in \operatorname{cone} S$.
>
> *Assertion $D_1 2$:* There exists an $f \in E^*$ such that $f(b) > 0$ and (every $x \in \operatorname{cone} S$ satisfies $f(x) \leq 0$).

Since $\operatorname{cone} S$ is a convex cone, it is clear that $\operatorname{cone} S$ is closed under multiplication by a nonnegative scalar (because convex cones are closed under multiplication by a nonnegative scalar). It is also clear that any linear combination of finitely many elements of $\operatorname{cone} S$ with nonnegative coefficients must lie in $\operatorname{cone} S$ (because $\operatorname{cone} S$ is a convex cone, and because any linear combination of finitely many elements of a convex cone with nonnegative coefficients must lie in this cone). In particular, the sum of any two elements of $\operatorname{cone} S$ must lie in $\operatorname{cone} S$.

First, it is clear that the Assertions $D_1 1$ and $D_1 2$ cannot hold at the same time[113]. We will now show that at least one of these assertions holds.

Let C be the polytope $\operatorname{conv.hull}(S \cup \{-b\})$ in E. According to Theorem 2.1c, exactly one of the following two assertions holds:

> *Assertion $D_2 1$:* We have $0 \in C$.
>
> *Assertion $D_2 2$:* There exists an $f \in E^*$ such that every $x \in C$ satisfies $f(x) < 0$.

Thus, we must be in one of the following two cases:

Case 1: Assertion $D_2 1$ holds.

Case 2: Assertion $D_2 2$ holds.

First, let us consider Case 1. In this case, Assertion $D_2 1$ holds. In other words, $0 \in C$. Thus, $0 \in C = \operatorname{conv.hull}(S \cup \{-b\})$. In other words, 0 is a convex combination of the elements of $S \cup \{-b\}$. In other words, there exist a family $(\lambda_s)_{s \in S \cup \{-b\}}$ of nonnegative reals such that $\sum\limits_{s \in S \cup \{-b\}} \lambda_s = 1$ and $\sum\limits_{s \in S \cup \{-b\}} \lambda_s s = 0$.

Consider this family $(\lambda_s)_{s \in S \cup \{-b\}}$.

---

[113] *Proof.* Assume the opposite. Then, the Assertions $D_1 1$ and $D_1 2$ hold at the same time. Since Assertion $D_1 2$ holds, there exists an $f \in E^*$ such that $f(b) > 0$ and (every $x \in \operatorname{cone} S$ satisfies $f(x) \leq 0$). Consider this f. We know that every $x \in \operatorname{cone} S$ satisfies $f(x) \leq 0$. Since $b \in \operatorname{cone} S$ (because Assertion $D_1 1$ holds), we can apply this to $x = b$, and thus obtain $f(b) \leq 0$. But this contradicts $f(b) > 0$. This contradiction shows that our assumption was wrong, qed.

We must be in one of the following two subcases:

Subcase 1.1: We have $\lambda_{-b} = 0$.

Subcase 1.2: We have $\lambda_{-b} \neq 0$.

Let us first consider Subcase 1.1. In this subcase, we have $\lambda_{-b} = 0$, so that

$$1 = \sum_{s \in S \cup \{-b\}} \lambda_s = \sum_{s \in S \setminus \{-b\}} \lambda_s + \underbrace{\sum_{s \in \{-b\}} \lambda_s}_{= \lambda_{-b} = 0}$$

$$\text{(since } S \cup \{-b\} \text{ is the union of the disjoint sets } S \setminus \{-b\} \text{ and } \{-b\})$$

$$= \sum_{s \in S \setminus \{-b\}} \lambda_s.$$

If every $s \in S \setminus \{-b\}$ would satisfy $\lambda_s = 0$, then we would have $\sum_{s \in S \setminus \{-b\}} \underbrace{\lambda_s}_{=0} = \sum_{s \in S \setminus \{-b\}} 0 = 0$, contradicting $\sum_{s \in S \setminus \{-b\}} \lambda_s = 1$. Thus, not every $s \in S \setminus \{-b\}$ satisfies $\lambda_s = 0$. In other words, there exists an $s \in S \setminus \{-b\}$ such that $\lambda_s \neq 0$. Let t be such an s. Then, $t \in S \setminus \{-b\}$ and $\lambda_t \neq 0$. Since λ_t is nonnegative and $\lambda_t \neq 0$, we have $\lambda_t > 0$.

We have

$$0 = \sum_{s \in S \cup \{-b\}} \lambda_s s = \sum_{s \in S \setminus \{-b\}} \lambda_s s + \underbrace{\sum_{s \in \{-b\}} \lambda_s s}_{\substack{= \lambda_{-b}(-b) = 0 \\ \text{(since } \lambda_{-b} = 0)}}$$

$$\text{(since } S \cup \{-b\} \text{ is the union of the disjoint sets } S \setminus \{-b\} \text{ and } \{-b\})$$

$$= \sum_{s \in S \setminus \{-b\}} \lambda_s s = \sum_{s \in (S \setminus \{-b\}) \setminus \{t\}} \lambda_s s + \lambda_t t,$$

so that

$$-\lambda_t t = \sum_{s \in (S \setminus \{-b\}) \setminus \{t\}} \lambda_s s \in \operatorname{cone} \underbrace{((S \setminus \{-b\}) \setminus \{t\})}_{\subseteq S}$$

$$\text{(since } \lambda_s \text{ is nonnegative for every } s \in (S \setminus \{-b\}) \setminus \{t\})$$

$$\subseteq \operatorname{cone} S. \tag{100}$$

Now, let us notice that

$$\text{every } \nu \in \mathbb{R} \text{ satisfies } \nu t \in \operatorname{cone} S. \tag{101}$$

[114] In other words,

$$t \mathbb{R} \subseteq \operatorname{cone} S. \tag{102}$$

---

[114] *Proof of (101):* Let $\nu \in \mathbb{R}$.

Let π be the canonical projection $E \to E \diagup (t\mathbb{R})$. Clearly,

$$|\pi(S \setminus \{t\})| \leq |S \setminus \{t\}| = \underbrace{|S|}_{=n} - 1 \qquad \text{(since } t \in S\text{)}$$

$$= n - 1 < n.$$

Hence, we can apply Theorem 2.5c to $E \diagup (t\mathbb{R})$, $\pi(S \setminus \{t\})$ and $\pi(b)$ instead of E, S and b (since we assumed that Theorem 2.5c holds whenever $|S| < n$), and conclude that exactly one of the following two assertions holds:

> *Assertion $D_3 1$:* We have $\pi(b) \in \operatorname{cone}(\pi(S \setminus \{t\}))$.

> *Assertion $D_3 2$:* There exists an $f \in (E \diagup (t\mathbb{R}))^*$ such that $f(\pi(b)) > 0$ and (every $x \in \operatorname{cone}(\pi(S \setminus \{t\}))$ satisfies $f(x) \leq 0$).

Thus, we must be in one of the following two subsubcases:
Subsubcase 1.1.1: Assertion $D_3 1$ holds.
Subsubcase 1.1.2: Assertion $D_3 2$ holds.
First, let us consider Subsubcase 1.1.1. In this subsubcase, Assertion $D_3 1$ holds. In other words, we have $\pi(b) \in \operatorname{cone}(\pi(S \setminus \{t\}))$. Thus, there exists a family $(\mu_w)_{w \in \pi(S \setminus \{t\})}$ of nonnegative reals such that $\pi(b) = \sum\limits_{w \in \pi(S \setminus \{t\})} \mu_w w$. Consider this family $(\mu_w)_{w \in \pi(S \setminus \{t\})}$.

For every $w \in \pi(S \setminus \{t\})$, let s_w be an (arbitrarily chosen) element of $S \setminus \{t\}$ satisfying $w = \pi(s_w)$. Such an s_w exists, since $w \in \pi(S \setminus \{t\})$. Now,

$$\pi(b) = \sum_{w \in \pi(S \setminus \{t\})} \mu_w \underbrace{w}_{=\pi(s_w)} = \sum_{w \in \pi(S \setminus \{t\})} \mu_w \pi(s_w) = \pi\left(\sum_{w \in \pi(S \setminus \{t\})} \mu_w s_w\right)$$

(since π is linear). Since π is linear, we have

$$\pi\left(b - \sum_{w \in \pi(S \setminus \{t\})} \mu_w s_w\right) = \pi(b) - \underbrace{\pi\left(\sum_{w \in \pi(S \setminus \{t\})} \mu_w s_w\right)}_{=\pi(b)} = \pi(b) - \pi(b) = 0.$$

We have $t \in S \subseteq \operatorname{cone} S$. Thus, if $\nu \geq 0$, then $\nu t \in \operatorname{cone} S$ (since $\operatorname{cone} S$ is closed under multiplication by a nonnegative scalar). Thus, if $\nu \geq 0$, then (101) clearly holds. Therefore, we can WLOG assume that $\nu \geq 0$ doesn't hold for the rest of this proof. Assume this.

So, we know that $\nu \geq 0$ doesn't hold. In other words, $\nu < 0$. Thus, $-\nu > 0$. Combined with $\lambda_t > 0$, this yields $\dfrac{-\nu}{\lambda_t} > 0$. Hence, $\dfrac{-\nu}{\lambda_t} \cdot (-\lambda_t t) \in \operatorname{cone} S$ (since $-\lambda_t t \in \operatorname{cone} S$ (by (100)), and since $\operatorname{cone} S$ is closed under multiplication by a nonnegative scalar). Since $\dfrac{-\nu}{\lambda_t} \cdot (-\lambda_t t) = \nu t$, this rewrites as $\nu t \in \operatorname{cone} S$. This proves (101).

[115] *Proof of (102):* Let $x \in t\mathbb{R}$ be arbitrary. Then, there exists a $\nu \in \mathbb{R}$ such that $x = \nu t$. Consider this ν. Then, $x = \nu t \in \operatorname{cone} S$ (by (101)). Now, forget that we fixed x. We have thus shown that every $x \in t\mathbb{R}$ satisfies $x \in \operatorname{cone} S$. In other words, $t\mathbb{R} \subseteq \operatorname{cone} S$. Thus, (102) is proven.

Thus,

$$b - \sum_{w \in \pi(S \setminus \{t\})} \mu_w s_w \in \operatorname{Ker} \pi = t\mathbb{R}$$

$$\text{(since } \pi \text{ is the canonical projection } E \to E \diagup (t\mathbb{R}))$$

$$\subseteq \operatorname{cone} S \qquad \text{(by (102))} . \tag{103}$$

But every $w \in \pi(S \setminus \{t\})$ satisfies $s_w \in S \setminus \{t\} \subseteq S \subseteq \operatorname{cone} S$. Therefore, $\sum_{w \in \pi(S \setminus \{t\})} \mu_w s_w \in \operatorname{cone} S$ (since any linear combination of finitely many elements of $\operatorname{cone} S$ with nonnegative coefficients must lie in $\operatorname{cone} S$). Combined with (103), this shows that $\sum_{w \in \pi(S \setminus \{t\})} \mu_w s_w$ and $b - \sum_{w \in \pi(S \setminus \{t\})} \mu_w s_w$ are two elements of $\operatorname{cone} S$.

Thus, their sum $\left(\sum_{w \in \pi(S \setminus \{t\})} \mu_w s_w \right) + \left(b - \sum_{w \in \pi(S \setminus \{t\})} \mu_w s_w \right)$ must also lie in $\operatorname{cone} S$ (because the sum of any two elements of $\operatorname{cone} S$ must lie in $\operatorname{cone} S$). In other words, $\operatorname{cone} S$ contains

$$\left(\sum_{w \in \pi(S \setminus \{t\})} \mu_w s_w \right) + \left(b - \sum_{w \in \pi(S \setminus \{t\})} \mu_w s_w \right) = b.$$

In other words, $b \in \operatorname{cone} S$. Thus, Assertion $D_1 1$ holds. Hence, at least one of Assertions $D_1 1$ and $D_1 2$ holds.

We have thus proven that at least one of Assertions $D_1 1$ and $D_1 2$ holds in Subsubcase 1.1.1.

Next, let us consider Subsubcase 1.1.2. In this subsubcase, Assertion $D_3 2$ holds. In other words, there exists an $f \in (E \diagup (t\mathbb{R}))^*$ such that $f(\pi(b)) > 0$ and (every $x \in \operatorname{cone}(\pi(S \setminus \{t\}))$ satisfies $f(x) \le 0$). Denote this f by h. Then, $h \in (E \diagup (t\mathbb{R}))^*$ satisfies $h(\pi(b)) > 0$ and (every $x \in \operatorname{cone}(\pi(S \setminus \{t\}))$ satisfies $h(x) \le 0$).

It is easy to see that

$$\text{every } s \in S \text{ satisfies } (h \circ \pi)(s) \le 0. \tag{104}$$

[116] From this, it is easy to see that

$$\text{every } x \in \operatorname{cone} S \text{ satisfies } (h \circ \pi)(x) \le 0. \tag{105}$$

---

[116] *Proof of (104):* Let $s \in S$.

Since π is the canonical projection $E \to E \diagup (t\mathbb{R})$, we have $\pi(t\mathbb{R}) = 0$, so that $\pi(t) = 0$ (since $t \in t\mathbb{R}$). Thus, $(h \circ \pi)(t) = h\left(\underbrace{\pi(t)}_{=0} \right) = h(0) = 0$ (since h is linear). Hence, if $s = t$, then (104) is proven. Therefore, we can WLOG assume that $s \ne t$ for the rest of this proof. Assume this.

Since $s \ne t$, we have $s \in S \setminus \{t\}$. Hence, $\pi(s) \in \pi(S \setminus \{t\}) \subseteq \operatorname{cone}(\pi(S \setminus \{t\}))$.

Recall that (every $x \in \operatorname{cone}(\pi(S \setminus \{t\}))$ satisfies $h(x) \le 0$). Applying this to $x = \pi(s)$, we obtain $h(\pi(s)) \le 0$. Hence, $(h \circ \pi)(s) = h(\pi(s)) \le 0$. This proves (104).

[117] Also, $(h \circ \pi)(b) = h(\pi(b)) > 0$. Altogether, we have thus shown that $(h \circ \pi)(b) > 0$ and (every $x \in \operatorname{cone} S$ satisfies $(h \circ \pi)(x) \leq 0$).

Thus, there exists an $f \in E^*$ such that $f(b) > 0$ and (every $x \in \operatorname{cone} S$ satisfies $f(x) \leq 0$) (namely, $f = h \circ \pi$). In other words, Assertion $D_1 2$ holds. Hence, at least one of Assertions $D_1 1$ and $D_1 2$ holds.

We have thus proven that at least one of Assertions $D_1 1$ and $D_1 2$ holds in Subsubcase 1.1.2.

Hence, in each of the Subsubcases 1.1.1 and 1.1.2, at least one of Assertions $D_1 1$ and $D_1 2$ holds. Since these Subsubcases 1.1.1 and 1.1.2 cover the whole Subcase 1.1, this yields that at least one of Assertions $D_1 1$ and $D_1 2$ holds in Subcase 1.1.

Next, let us consider Subcase 1.2. In this subcase, we have $\lambda_{-b} \neq 0$. Combined with the fact that λ_{-b} is nonnegative, this yields $\lambda_{-b} > 0$. Thus, $\dfrac{1}{\lambda_{-b}}$ exists and is > 0. Now,

$$0 = \sum_{s \in S \cup \{-b\}} \lambda_s s = \sum_{s \in S \setminus \{-b\}} \lambda_s s + \underbrace{\sum_{s \in \{-b\}} \lambda_s s}_{=\lambda_{-b}(-b)}$$

$$\qquad \text{(since } S \cup \{-b\} \text{ is the union of the disjoint sets } S \setminus \{-b\} \text{ and } \{-b\})$$

$$= \sum_{s \in S \setminus \{-b\}} \lambda_s s + \lambda_{-b}(-b) = \sum_{s \in S \setminus \{-b\}} \lambda_s s - \lambda_{-b} b,$$

---

[117] *Proof of (105):* Let $x \in \operatorname{cone} S$. Then,

$$x \in \operatorname{cone} S = \left\{ \sum_{s \in S} \nu_s s \ \mid \ (\nu_s)_{s \in S} \text{ is a family of nonnegative reals} \right\}$$

(by the definition of $\operatorname{cone} S$). Hence, there exists a family $(\nu_s)_{s \in S}$ of nonnegative reals such that $x = \sum_{s \in S} \nu_s s$. Consider this $(\nu_s)_{s \in S}$.

Since $x = \sum_{s \in S} \nu_s s$, we have

$$(h \circ \pi)(x) = (h \circ \pi)\left(\sum_{s \in S} \nu_s s\right) = \sum_{s \in S} \underbrace{\nu_s (h \circ \pi)(s)}_{\substack{\leq 0 \\ (\text{since } \nu_s \geq 0 \text{ and} \\ (h \circ \pi)(s) \leq 0 \text{ (by (104)))}}} \qquad (\text{since } h \circ \pi \text{ is } \mathbb{R}\text{-linear})$$

$$\leq \sum_{s \in S} 0 = 0.$$

This proves (105).

so that

$$\lambda_{-b} b = \sum_{s \in S \setminus \{-b\}} \lambda_s s \in \text{cone} \underbrace{(S \setminus \{-b\})}_{\subseteq S}$$

$$\text{(since } \lambda_s \text{ is nonnegative for every } s \in S \setminus \{-b\})$$

$$\subseteq \text{cone } S.$$

Since $\dfrac{1}{\lambda_{-b}} > 0$, this yields $\dfrac{1}{\lambda_{-b}} \lambda_{-b} b \in \text{cone } S$ (because cone S is closed under multiplication by a nonnegative scalar). Since $\dfrac{1}{\lambda_{-b}} \lambda_{-b} b = b$, this rewrites as $b \in \text{cone } S$. In other words, Assertion $D_1 1$ holds. Hence, at least one of Assertions $D_1 1$ and $D_1 2$ holds.

We have thus proven that at least one of Assertions $D_1 1$ and $D_1 2$ holds in Subcase 1.2.

Hence, in each of the Subcases 1.1 and 1.2, at least one of Assertions $D_1 1$ and $D_1 2$ holds. Since these Subcases 1.1 and 1.2 cover the whole Case 1, this yields that at least one of Assertions $D_1 1$ and $D_1 2$ holds in Case 1.

Finally, let us consider Case 2. In this case, Assertion $D_2 2$ holds. In other words, there exists an $f \in E^*$ such that every $x \in C$ satisfies $f(x) < 0$. Denote this f by h. Then,

$$\text{every } x \in C \text{ satisfies } h(x) < 0. \tag{106}$$

Now, it is easy to see that

$$\text{every } x \in \text{cone } S \text{ satisfies } h(x) \leq 0. \tag{107}$$

[118]

Also, $-b \in S \cup \{-b\} \subseteq \text{conv.hull}(S \cup \{-b\}) = C$, so that $h(-b) < 0$ (by (106)). Since h is linear, we have $-h(b) = h(-b) < 0$, so that $h(b) > 0$.

---

[118] *Proof of (107):* Let $x \in \text{cone } S$. Then, $x \in \text{cone } S = \left\{ \sum_{s \in S} \nu_s s \mid (\nu_s)_{s \in S} \text{ is a family of nonnegative reals} \right\}$. Thus, there exists a family $(\nu_s)_{s \in S}$ of nonnegative reals such that $x = \sum_{s \in S} \nu_s s$. Consider this family $(\nu_s)_{s \in S}$.

Every $s \in S$ satisfies $h(s) < 0$ (by (106), applied to $x = s$ (since $s \in S \subseteq S \cup \{-b\} \subseteq \text{conv.hull}(S \cup \{-b\}) = C$)) and $\nu_s \geq 0$, so that $\underbrace{\nu_s}_{\geq 0} \underbrace{h(s)}_{<0} \leq 0$. Now, since $x = \sum_{s \in S} \nu_s s$, we have

$$h(x) = h\left(\sum_{s \in S} \nu_s s \right) = \sum_{s \in S} \underbrace{\nu_s h(s)}_{\leq 0} \qquad \text{(since } h \text{ is } \mathbb{R}\text{-linear)}$$

$$\leq \sum_{s \in S} 0 = 0.$$

This proves (107).

Altogether, we now know that $h(b) > 0$ and (every $x \in \operatorname{cone} S$ satisfies $h(x) \leq 0$) (by (107)). Thus, there exists an $f \in E^*$ such that $f(b) > 0$ and (every $x \in \operatorname{cone} S$ satisfies $f(x) \leq 0$) (namely, $f = h$). In other words, Assertion $D_1 2$ holds. Hence, at least one of Assertions $D_1 1$ and $D_1 2$ holds.

We have thus proven that at least one of Assertions $D_1 1$ and $D_1 2$ holds in Case 2.

Hence, in each of the Cases 1 and 2, at least one of Assertions $D_1 1$ and $D_1 2$ holds. Since these Cases 1 and 2 cover all possibilities, this yields that, in every situation, at least one of Assertions $D_1 1$ and $D_1 2$ holds. Since we know that the Assertions $D_1 1$ and $D_1 2$ cannot hold at the same time, this yields that exactly one of Assertions $D_1 1$ and $D_1 2$ holds.

Now, forget that we fixed E, S and b. We have thus proven that if E is a finite-dimensional $\mathbb{R}$-vector space, S is a finite subset of E such that $|S| = n$, and b is an element of E, then exactly one of the following two assertions holds:

> *Assertion $D_1 1$:* We have $b \in \operatorname{cone} S$.
>
> *Assertion $D_1 2$:* There exists an $f \in E^*$ such that $f(b) > 0$ and (every $x \in \operatorname{cone} S$ satisfies $f(x) \leq 0$).

In other words, we have proven that Theorem 2.5c holds in the case when $|S| = n$. This completes the induction step, and thus the induction proof of Theorem 2.5c is complete. $\qquad\square$

References

[Schrij17] Alexander Schrijver, *A Course in Combinatorial Optimization*, version March 23rd, 2017.
`http://homepages.cwi.nl/~lex/files/dict.pdf`

[Schrij98] Alexander Schrijver, *Theory of Linear and Integer Programming*, Wiley, 1999.

[Laurit13] Niels Lauritzen, *Undergraduate Convexity: From Fourier and Motzkin to Kuhn and Tucker*, World Scientific 2013.

[Jacim11] Milojica Jaćimović, *Farkas' Lemma of Alternative*, The teaching of Mathematics, 2011, Vol. XIV, 2, pp. 77–86.
`http://elib.mi.sanu.ac.rs/files/journals/tm/27/tm1422.pdf`

[Bartl12] David Bartl, *A very short algebraic proof of the Farkas Lemma*, Math Meth Oper Res (2012) 75:101–104.
`https://doi.org/10.1007/s00186-011-0377-y`